J Glob Optim (2008) 40:529–543
DOI 10.1007/s10898-006-9123-z

ORIGINAL PAPER

A new criterion for the inexact logarithmic-quadratic proximal method and its derived hybrid methods

Xiao-Ming Yuan

Received: 31 October 2006 / Accepted: 2 November 2006 / Published online: 16 February 2007

Abstract To solve nonlinear complementarity problems, the inexact logarithmic-quadratic proximal (LQP) method solves a system of nonlinear equations (*LQP system*) approximately at each iteration. Therefore, the efficiencies of inexact-type LQP methods depend greatly on the involved inexact criteria used to solve the *LQP systems*. This paper relaxes inexact criteria of existing inexact-type LQP methods and thus makes it easier to solve the *LQP system* approximately. Based on the approximate solutions of the *LQP systems*, a descent method, and a prediction–correction method are presented. Convergence of the new methods are proved under mild assumptions. Numerical experiments for solving traffic equilibrium problems demonstrate that the new methods are more efficient than some existing methods and thus verify that the new inexact criterion is attractive in practice.

Keywords Nonlinear complementarity problems · Logarithmic-quadratic proximal · Descent · Prediction–correction · Inexact criterion

1 Introduction

The finite-dimensional nonlinear complementarity problem (NCP) is to determine a vector $x \in \mathcal{R}^n$ such that

$$x \geq 0, \qquad F(x) \geq 0 \qquad \text{and} \quad x^T F(x) = 0, \tag{1.1}$$

where F is a mapping from R^n into itself. Throughout we assume that F is continuous and monotone and that the solution set of (1.1) denoted by $\mathcal{X}^*$ is nonempty. We refer to, e.g., [10,16], for NCP's various applications arising in operation research,

X.-M Yuan (✉)
Department of Management Science, Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai, 200052, China
e-mail: xmyuan@sjtu.edu.cn

transportation research, engineering, economic equilibrium, mathematical programming, etc.

The well known proximal point algorithm (PPA), which was presented originally by Martinet [15] for finding roots of maximal monotone operators, is applicable for solving NCP. In particular, let x^k be the current approximation of a solution of (1.1), then PPA generates the new iterate x^{k+1} by solving the following auxiliary NCP:

$$x \geq 0, \quad c_k F(x) + (x - x^k) \geq 0 \qquad \text{and} \qquad x^T\big(c_k F(x) + (x - x^k)\big) = 0, \tag{1.2}$$

where $c_k \in [c, \infty)$ is the proximal parameter and $c > 0$. Compared to the monotone NCP (1.1), (1.2) is easier to handle since it is a strong monotone NCP. For more developments on PPA, we refer to, e.g., [1,4,7–9,11,17,18].

The logarithmic-quadratic proximal (LQP) method presented in [2] improves the PPA by replacing the linear term $x - x^k$ with

$$x - (1-\mu)x^k - \mu X_k^2 x^{-1},$$

where $\mu \in (0, 1)$ is a given constant, $X_k = \text{diag}(x_1^k, x_2^k, \ldots, x_n^k)$ and x^{-1} is a n-vector whose jth element is $1/x_j$. At the kth iteration, solving NCP by the LQP method is equivalent to finding the positive solution of the following system of nonlinear equations (*LQP system* for convenience)

$$c_k F(x) + x - (1-\mu)x^k - \mu X_k^2 x^{-1} = 0. \tag{1.3}$$

Since in general solving the *LQP system* is much easier than solving the auxiliary NCP (1.2), the LQP method provides a very powerful approach to solving NCP. Note that exact solutions of the *LQP system* cannot be obtained trivially. Therefore, the more practical version of the LQP method is the inexact LQP method presented also in [2], which solves the *LQP system* approximately in the following sense: find $x^{k+1} \in \mathcal{R}^n_{++}$ and $\xi^k \in \mathcal{R}^n$ such that

$$c_k F(x) + x - (1-\mu)x^k - \mu X_k^2 x^{-1} = \xi^k \tag{1.4}$$

and

$$\sum_{k=0}^{\infty} \|\xi^k\| < +\infty \qquad \text{and} \qquad \sum_{k=0}^{\infty} \langle \xi^k, x^k \rangle < +\infty. \tag{1.5}$$

Note that (1.5) implies that the involved *LQP systems* need to be solved with increasing accuracies, which probably results in computational difficulties. Therefore, it is worthy to investigate extensively how to relax existing inexact criteria used to solve the involved *LQP system*. The first contribution to overcome this drawback was due to [5], in which the authors improved (1.5) with relative errors and thus presented a meaningful modification of the inexact LQP method. The attractive characteristic of the new method is that the relative errors for solving the *LQP system* approximately, which is denoted by

$$\frac{\|\xi^k\|}{\|x^k - x^{k+1}\|}$$

can be fixed on a constant. Recently, Xu and Bnouhachem [19] improved the inexact LQP method in the sense that the restriction on ξ^k is relaxed to

$$\|\xi^k\| \leq \eta\sqrt{1-\mu^2}\|x^k - x^{k+1}\| \quad \text{with} \quad \eta \in (0, 1), \tag{1.6}$$

which implies that the relative errors of solving the *LQP systems* approximately can be fixed at $\eta\sqrt{1-\mu^2}$. Based on (1.6), a descent direction of $\|x-x^*\|^2$ at the point x^k was constructed and thus a descent algorithm was presented in [19]. The consequent improvement in this direction is the prediction–correction approach presented in [21], which uses the inexact LQP method under (1.6) to predict the new iterate and then uses PPA to correct the prediction. For most recent developments on inexact-type LQP methods, we refer to [13,14].

For solving the involved *LQP systems* more efficiently, this paper presents a new inexact criterion (see (2.2)), which allows the *LQP system* to be solved under very relaxed restriction and thus improves the existing inexact-type LQP methods. In addition, approximate solutions of the *LQP system* are used to construct a descent method and a prediction–correction approach analogous to [19,21], respectively. Both of the new methods are easy to implement. Numerical applications to traffic equilibrium problems demonstrate that the new methods are very efficient, and thus verify that the new inexact criterion is practical.

The rest of the paper is organized as follows. The descent method and some of its contractive propositions are presented in Sect. 2. In Sect. 3, we present the prediction–correction method and some of its theoretical results. Section 4 focuses on how to choose the optimal step sizes and concerns convergence of the new methods. Some numerical results are reported in Sect. 5. Finally, some conclusions are drawn in Sect. 6.

2 The derived descent method

We first give the descent method and then prove some contractive propositions.

Let $P_{\mathcal{R}^n_+}$ denote the projection onto $\mathcal{R}^n_+$ under Euclidean norm:

$$P_{\mathcal{R}^n_+}(v) = \operatorname{argmin}\{\|u-v\| \ : \ u \in \mathcal{R}^n_+\}.$$

Given constants $c>0$, $\eta\in(0,1)$, $r>\eta$, $\mu\in(0,1)$, and $\gamma\in(0,2)$ and starting from $x^0\in\mathcal{R}^n_{++}$, the new iterate x^{k+1} is generated by

Step 0 Choose $c_k \ge c$.

Step 1 Inexact LQP step

Find $\tilde{x}^k \in \mathcal{R}^n_+$ and $\xi^k \in \mathcal{R}^n$ such that

$$c_k F(\tilde{x}^k) + \tilde{x}^k - (1-\mu)x^k - \mu X_k^2(\tilde{x}^k)^{-1} = \xi^k \tag{2.1}$$

with the inexact criterion:

$$\frac{|(\xi^k)^T(x^k-\tilde{x}^k)|}{\|x^k-\tilde{x}^k\|^2} \le \eta \qquad \text{and} \qquad \frac{\|\xi^k\|}{\|x^k-\tilde{x}^k\|} \le r(1+\mu); \tag{2.2}$$

Step 2 Descent step

$$x_d^{k+1}(\alpha_k) = P_{\mathcal{R}^n_+}[x^k - \alpha_k d(x^k,\tilde{x}^k,\xi^k,\mu)], \tag{2.3}$$

where

$$d(x^k,\tilde{x}^k,\xi^k,\mu) = (x^k-\tilde{x}^k) + \tfrac{1}{1+\mu}\xi^k, \tag{2.4}$$

$$\varphi(x^k,\tilde{x}^k,\xi^k,\mu) = \tfrac{1}{1+\mu}\|x^k-\tilde{x}^k\|^2 + \tfrac{1}{1+\mu}(x^k-\tilde{x}^k)^T\xi^k \tag{2.5}$$

and

$$\alpha_k = \gamma \frac{\varphi(x^k, \tilde{x}^k, \xi^k, \mu)}{\|d(x^k, \tilde{x}^k, \xi^k, \mu)\|^2}. \tag{2.6}$$

Remark 2.1 *As pointed in [13], the inexact LQP step (2.1) is implementable. For example, in the case that F is Lipschitz continuous with constant L on R^n_+, we particularly take $\xi^k = c_k(F(\tilde{x}^k) - F(x^k))$. Then the criterion (2.2) can be satisfied via choosing a suitable $c_k > 0$ and the inexact LQP system (2.1) reduces to*

$$c_k F(x^k) + \tilde{x}^k - (1-\mu)x^k - \mu X_k^2 (\tilde{x}^k)^{-1} = 0,$$

whose positive solution can be computed explicitly:

$$\tilde{x}^k_j = \frac{(1-\mu)x^k_j - c_k F_j(x^k) + \sqrt{[(1-\mu)x^k_j - c_k F_j(x^k)]^2 + 4\mu (x^k_j)^2}}{2}.$$

Remark 2.2 *Note that convergence of the new method can also be guaranteed under the stricter inexact criterion $\|\xi^k\| \le \eta \|x^k - \tilde{x}^k\|$, which implies that the relative errors of solving the involved LQP systems approximately can only be fixed at $\eta \in (0, 1)$. Note that r can take any positive value greater than η. Therefore, the inexact criterion (2.2) allows the involved LQP systems to be solved approximately under very relaxed criterion and thus reduces the computational load considerably.*

Remark 2.3 *The new derived descent method differs from the hybrid method in [19] mainly in that the inexact criterion (1.6) is relaxed to (2.2). This relaxation is particularly preferable in practice, which will be verified by the numerical reports.*

Remark 2.4 *The theoretical reasons of choosing the optimal step size α_k and why $\gamma \in (0, 2)$ will be discussed later.*

We first list some fundamental lemmas that are useful in the consequent analysis. The first lemma provides some basic inequalities of projection onto $\mathcal{R}^n_+$ without proof, see, e.g., [22].

Lemma 2.1 *Let $P_{\mathcal{R}^n_+}$ denote the projection onto $\mathcal{R}^n_+$ under Euclidean norm. Then we have the following fundamental inequalities:*

$$\{z - P_{\mathcal{R}^n_+}(z)\}^T \{x - P_{\mathcal{R}^n_+}(z)\} \le 0, \quad \forall x \in \mathcal{R}^n_+, \ \forall z \in \mathcal{R}^n, \tag{2.7}$$

$$\|P_{\mathcal{R}^n_+}(y) - P_{\mathcal{R}^n_+}(z)\| \le \|y - z\|, \quad \forall y, z \in \mathcal{R}^n, \tag{2.8}$$

$$\|P_{\mathcal{R}^n_+}(y) - x\|^2 \le \|y - x\|^2 - \|y - P_{\mathcal{R}^n_+}(y)\|^2, \quad \forall x \in \mathcal{R}^n_+, \ \forall y \in \mathcal{R}^n. \tag{2.9}$$

The following is a basic lemma in the analysis of the LQP method and its variants. We omit the proof, which can be found in, e.g., [2,5,13].

Lemma 2.2 *For given $x^k \in \mathcal{R}^n_+$ and $c_k > 0$, let $\tilde{x}^k$ and ξ^k be obtained by (2.1), then for any $x \in \mathcal{R}^n_+$ we have*

$$(\tilde{x}^k - x)^T(\xi^k - c_k F(\tilde{x}^k)) \ge \tfrac{1+\mu}{2}\left(\|\tilde{x}^k - x\|^2 - \|x^k - x\|^2\right) + \tfrac{1-\mu}{2}\|x^k - \tilde{x}^k\|^2 \tag{2.10}$$

and

$$(\widetilde{x}^k - x)^T(\xi^k - c_k F(\tilde{x}^k)) \geq (x^k - \widetilde{x}^k)^T \left((1+\mu)x - (\mu x^k + \widetilde{x}^k)\right). \quad (2.11)$$

The following theorem provides the theoretical reason of designing the descent method (2.3).

Theorem 2.1 *Let x^* be any solution of* (1.1). *For given $x^k \in \mathcal{R}^n_+$ and $c_k > 0$, let $\widetilde{x}^k$ and ξ^k be obtained by* (2.1) and (2.2)*; let $d(x^k, \tilde{x}^k, \xi^k, \mu)$ and $\varphi(x^k, \tilde{x}^k, \xi^k, \mu)$ be defined by* (2.4) *and* (2.5), *respectively, then it holds*

$$(x^k - x^*)^\top d(x^k, \tilde{x}^k, \xi^k, \mu) \geq \varphi(x^k, \tilde{x}^k, \xi^k, \mu) \geq \frac{1-\eta}{1+\mu}\|x^k - \widetilde{x}^k\|^2. \quad (2.12)$$

Proof Since F is monotone and x^* is a solution, we get

$$(\widetilde{x}^k - x^*)^\top F(\widetilde{x}^k) \geq (\widetilde{x}^k - x^*)^\top F(x^*) \geq 0.$$

Then it follows from (2.10) that

$$(\widetilde{x}^k - x^*)^\top \xi^k \geq \frac{1+\mu}{2}\left(\|\widetilde{x}^k - x^*\|^2 - \|x^k - x^*\|^2\right) + \frac{1-\mu}{2}\|x^k - \widetilde{x}^k\|^2 + c_k(\widetilde{x}^k - x^*)^\top F(\widetilde{x}^k),$$

which implies that

$$\frac{1}{1+\mu}(x^k - x^*)^\top \xi^k \geq \frac{1}{2}\left(\|\widetilde{x}^k - x^*\|^2 - \|x^k - x^*\|^2\right) + \frac{1-\mu}{2(1+\mu)}\|x^k - \widetilde{x}^k\|^2 + \frac{1}{1+\mu}(x^k - \tilde{x}^k)^\top \xi^k. \quad (2.13)$$

Note the following identity

$$(x^k - x^*)^\top (x^k - \tilde{x}^k) = \frac{1}{2}\left(\|x^k - x^*\|^2 - \|\widetilde{x}^k - x^*\|^2\right) + \frac{1}{2}\|x^k - \widetilde{x}^k\|^2. \quad (2.14)$$

Therefore from (2.4), (2.5) (2.13), and (2.14), we have

$$\begin{aligned}(x^k - x^*)^\top d(x^k, \tilde{x}^k, \xi^k, \mu) &= (x^k - x^*)^\top (x^k - \tilde{x}^k) + \frac{1}{1+\mu}(x^k - x^*)^\top \xi^k \\ &\geq \frac{1}{(1+\mu)}\|x^k - \widetilde{x}^k\|^2 + \frac{1}{1+\mu}(x^k - \tilde{x}^k)^\top \xi^k \\ &= \varphi(x^k, \tilde{x}^k, \xi^k, \mu),\end{aligned}$$

which is the first inequality in (2.12).

Since (2.2) implies that

$$(x^k - \tilde{x}^k)^T(\xi^k) \geq -\eta\|x^k - \tilde{x}^k\|^2$$

the second inequality in (2.12) is obvious.

Theorem 2.1 shows that $d(x^k, \tilde{x}^k, \xi^k, \mu)$ is a descent direction of $\|x - x^*\|^2$ at $x = x^k$. Therefore, it is nature to design the descent method (2.3). Obviously, it is worthy to investigate the strategy of choosing the optimal step size along the descent direction from computational points of view. For this purpose, we denote

$$x_d^{k+1}(\alpha) := P_{\mathcal{R}^n_+}[x^k - \alpha d(x^k, \tilde{x}^k, \xi^k, \mu)]. \quad (2.15)$$

For any solution of (1.1) x^*, then

$$\Theta_d^k(\alpha) := \|x^k - x^*\|^2 - \|x_d^{k+1}(\alpha) - x^*\|^2 \tag{2.16}$$

measures the progress of the iterate generated by the descent method (2.3) taking α as the step size. The following theorem motivates the strategy of choosing α_k in the form of (2.6). □

Theorem 2.2 *Let x^* be any solution of (1.1). For given $x^k \in \mathcal{R}^n_+$ and $c_k > 0$, let $\tilde{x}^k$ and ξ^k be obtained by (2.1) and (2.2); let $d(x^k, \tilde{x}^k, \xi^k, \mu)$ and $\varphi(x^k, \tilde{x}^k, \xi^k, \mu)$ be defined by (2.4) and (2.5), respectively. If the step size in the descent step (2.3) is α, i.e., the new iterate is $x_d^{k+1}(\alpha)$ defined by (2.15); and $\Theta_d^k(\alpha)$ is defined by (2.16), then we have*

$$\Theta_d^k(\alpha) \geq \Lambda^k(\alpha) := 2\alpha\varphi(x^k, \tilde{x}^k, \xi^k, \mu) - \alpha^2\|d(x^k, \tilde{x}^k, \xi^k, \mu)\|^2. \tag{2.17}$$

Proof

$$\begin{aligned}\|x^{k+1}(\alpha) - x^*\|^2 &= \left\|P_{R^n_+}[x^k - \alpha d(x^k, \tilde{x}^k, \xi^k, \mu)] - x^*\right\|^2 \\ &\leq \|x^k - x^* - \alpha d(x^k, \tilde{x}^k, \xi^k, \mu)\|^2 \\ &= \|x^k - x^*\|^2 - 2\alpha(x^k - x^*)^\top d(x^k, \tilde{x}^k, \xi^k, \mu) + \alpha^2\|d(x^k, \tilde{x}^k, \xi^k, \mu)\|^2 \\ &\leq \|x^k - x^*\|^2 - 2\alpha\varphi_k(x^k, \tilde{x}^k, \xi^k, \mu) + \alpha^2\|d(x^k, \tilde{x}^k, \xi^k, \mu)\|^2, \end{aligned} \tag{2.18}$$

where the inequalities are guaranteed by (2.8) and (2.12), respectively. Therefore, the assertion (2.17) is proved.

3 The derived prediction–correction method

Given constants $c > 0$, $\eta \in (0, 1)$, $r > \eta$, $\mu \in (0, 1)$, and $\gamma \in (0, 2)$ and starting from $x^0 \in \mathcal{R}^n_{++}$, each iteration consists of the following steps:
Step 0 Choose $c_k \geq c$.
Step 1 Inexact LQP step (prediction step)
Solving the *LQP system* approximately in the sense of (2.1) and (2.2) and thus obtain the predictor $\tilde{x}^k \in \mathcal{R}^n_+$.
Step 2 PPA step (correction step)

$$x_{\text{pc}}^{k+1}(\alpha_k) = P_{\mathcal{R}^n_+}[x^k - \alpha_k \frac{c_k}{1+\mu} F(\tilde{x}^k)], \tag{3.1}$$

where $d(x^k, \tilde{x}^k, \xi^k, \mu)$, $\varphi(x^k, \tilde{x}^k, \xi^k, \mu)$, and α_k are defined by (2.4)–(2.6), respectively.

Remark 3.5 *It is well known [6] that solving (1.2) is equivalent to solving the following equation:*

$$x = P_{\mathcal{R}^n_+}[x^k - c_k F(x)]. \tag{3.2}$$

Note that (3.2) is a nonsmooth equation and the new iterate x^{k+1} cannot be computed directly via (3.2) since it is an implicit scheme. This difficulty makes straightforward applications of PPA impractical in many cases. The hybrid method in [21] provides a prediction–correction approach to make PPA implementable. In particular, it solves the LQP system (2.1) under (1.6) to obtain the predictor $\tilde{x}^k \in \mathcal{R}^n_+$ and then computes the new iterate x^{k+1} by (3.1). The new derived prediction–correction method improves *the method in* [21] *with the practical characteristic that* (1.6) *is replaced by the relaxed criterion* (2.2)*. The numerical results to be reported demonstrate that this improvement reduces computational efforts considerably.*

Similarly, we denote

$$x_{\rm pc}^{k+1}(\alpha) := P_{\mathcal{R}_+^n}\left[x^k - \alpha\left(\frac{c_k}{1+\mu}F(\tilde{x}^k)\right)\right]. \tag{3.3}$$

For any solution of (1.1) x^*, then

$$\Theta_{\rm pc}^k(\alpha) := \|x^k - x^*\|^2 - \|x_{\rm pc}^{k+1}(\alpha) - x^*\|^2 \tag{3.4}$$

measures the progress of the iterate generated by the prediction–correction method (3.1) taking α as the step size.

Theorem 3.1 *Let x^* be any solution of* (1.1)*. For given $x^k \in \mathcal{R}_+^n$ and $c_k > 0$, let $\tilde{x}^k$ and ξ^k be obtained by* (2.1) *and* (2.2)*; let $d(x^k, \tilde{x}^k, \xi^k, \mu)$ and $\varphi(x^k, \tilde{x}^k, \xi^k, \mu)$ be defined by* (2.4) *and* (2.5)*, respectively. If the step size in the prediction–correction step* (3.1) *is α, i.e., the new iterate is $x_{\rm pc}^{k+1}(\alpha)$ defined by* (3.3)*; $\Theta_{\rm pc}^k(\alpha)$ is defined by* (3.4) *and $\Lambda^k(\alpha)$ is defined in* (2.17)*. Then we have*

$$\Theta_{\rm pc}^k(\alpha) \geq \Lambda^k(\alpha). \tag{3.5}$$

Proof We omit the proof since it is analogous to that of Theorem 3.1 in [21].

4 Convergence

In this section, convergence of the new methods are proved under mild assumptions. First, we explain the reason of choosing the step size α_k in the new methods (2.3) and (3.1).

Note that $\Lambda^k(\alpha)$ is referred to the *profit-function* since it is a lower-bound of the progress obtained by the new iterates. Theorems 2.2 and 3.1 motivate us to maximize the profit-function $\Lambda^k(\alpha)$ to accelerate convergence of the new methods. Note that $\Lambda^k(\alpha)$ is a quadratic function of α and it reaches its maximum at

$$\alpha^* = \frac{\varphi(x^k, \tilde{x}^k, \xi^k, \mu)}{\|d(x^k, \tilde{x}^k, \xi^k, \mu)\|^2} \quad \text{with} \quad \Lambda^k(\alpha^*) = \alpha^*\varphi(x^k, \tilde{x}^k, \xi^k, \mu). \tag{4.1}$$

The following lemma concerns the nice proposition of α^* and justifies the theoretical reason of choosing α^*.

Lemma 4.1 *Let μ, η, r be given in the algorithm. Let α^* be defined by* (4.1) *and $\delta \geq \max\{r, 1\}$ be a constant. Then we have*

$$\alpha^* \geq \frac{1-\eta}{(1+\delta^2)(1+\mu)}. \tag{4.2}$$

Proof **Case 1** $(x^k - \tilde{x}^k)^T\xi^k \leq 0$.

It follows from the definition of $d(x^k, \tilde{x}^k, \xi^k, \mu)$, (2.2), and $\delta \geq \max\{r, 1\}$ that

$$\|d(x^k, \tilde{x}^k, \xi^k, \mu)\|^2 \leq \|x^k - \tilde{x}^k\|^2 + \frac{1}{(1+\mu)^2}\|\xi^k\|^2 \leq (1+r^2)\|x^k - \tilde{x}^k\|^2 \leq (1+\delta^2)\|x^k - \tilde{x}^k\|^2.$$

Note that

$$\varphi(x^k,\tilde{x}^k,\xi^k,\mu) = \tfrac{1}{1+\mu}\|x^k-\tilde{x}^k\|^2 + \tfrac{1}{1+\mu}(x^k-\tilde{x}^k)^T\xi^k \geq \frac{1-\eta}{1+\mu}\|x^k-\tilde{x}^k\|^2. \quad (4.3)$$

Hence, the assertion (4.2) is proved in this case.

Case 2 $(x^k-\tilde{x}^k)^T\xi^k > 0$.

It follows from the definition of $d(x^k,\tilde{x}^k,\xi^k,\mu)$, $\varphi(x^k,\tilde{x}^k,\xi^k,\mu)$, (2.2), $\mu \in (0,1)$, and $\delta \geq \max\{r,1\}$ that

$$\begin{aligned}(1+\mu)\varphi(x^k,\tilde{x}^k,\xi^k,\mu) &= \|x^k-\tilde{x}^k\|^2 + (x^k-\tilde{x}^k)^T\xi^k \\ &\geq \|x^k-\tilde{x}^k\|^2 + (x^k-\tilde{x}^k)^T(\frac{\xi^k}{1+\mu}) \\ &\geq \frac{1}{1+\delta^2}\|x^k-\tilde{x}^k\|^2 + \frac{2}{1+\delta^2}(x^k-\tilde{x}^k)^T(\frac{\xi^k}{1+\mu}) \\ &+ \frac{\delta^2}{1+\delta^2}\|x^k-\tilde{x}^k\|^2 \\ &\geq \frac{1}{1+\delta^2}\|x^k-\tilde{x}^k\|^2 + \frac{2}{1+\delta^2}(x^k-\tilde{x}^k)^T(\frac{\xi^k}{1+\mu}) \\ &+ \frac{1}{1+\delta^2}\|\frac{\xi^k}{1+\mu}\|^2 \\ &= \frac{1}{1+\delta^2}\|d(x^k,\tilde{x}^k,\xi^k,\mu)\|^2. \end{aligned} \quad (4.4)$$

Therefore, we have

$$\alpha^* \geq \frac{1}{(1+\delta^2)(1+\mu)},$$

which implies the assertion (4.2) immediately.

Lemma 4.1 shows that the step size α_k of the new methods is bounded away from zero, which contributes much to the satisfactory efficiencies of the new methods.

From numerical point of view, it is necessary to attach a relax factor to the optimal step size α^* obtained theoretically to achieve faster convergence. The following corollary concerns how to choose the relax factor.

Theorem 4.1 *Let x^* be any solution of* (1.1). *For given $x^k \in \mathcal{R}^n_+$ and $c_k > 0$, let $\tilde{x}^k$ be obtained by* (2.1) *and* (2.2)*; let γ be a positive constant and α^* be defined by* (4.1). *If the step size $\alpha_k = \gamma\alpha^*$ in* (2.3) *and* (3.1). *Then there exists a constant $\rho > 0$, such that*

$$\Lambda^k(\alpha_k) \geq \rho\|x^k-\tilde{x}^k\|^2, \qquad \forall k > 0. \quad (4.5)$$

Proof It follows from (4.1)–(4.3) that

$$\Lambda(\alpha^*) = \alpha^*\varphi(x^k,\tilde{x}^k,\xi^k,\mu) \geq \frac{1-\eta}{(1+\delta^2)(1+\mu)^2}\|x^k-\tilde{x}^k\|^2. \quad (4.6)$$

Note that

$$\begin{aligned}\Lambda(\gamma\alpha^*) &\overset{(2.17)}{=} 2\gamma\alpha^*\varphi(x^k,\tilde{x}^k,\xi^k,\mu) - (\gamma^2\alpha^*)(\alpha^*\|d(x^k,\tilde{x}^k,\xi^k,\mu)\|^2)\\ &\overset{(4.1)}{=} (2\gamma\alpha^* - \gamma^2\alpha^*)\varphi(x^k,\tilde{x}^k,\xi^k,\mu)\\ &= \gamma(2-\gamma)\Lambda(\alpha^*)\\ &\overset{(4.6)}{\geq} \frac{\gamma(2-\gamma)(1-\eta)}{(1+\delta^2)(1+\mu)^2}\|x^k - \tilde{x}^k\|^2. \end{aligned} \tag{4.7}$$

Thus the assertion (4.5) is proved with $\rho := \frac{\gamma(2-\gamma)(1-\eta)}{(1+\delta^2)(1+\mu)^2} > 0$.

Theorem 4.1 shows theoretically that any $\gamma \in (0,2)$ guarantees that the new iterates generated by (2.3) and (3.1) make progress to a solution of (1.1). However, from numerical experiments, $\gamma \in [1,2)$ is much preferable since it leads to better numerical performances. Therefore, in practical computation, we choose

$$\alpha_k = \gamma\alpha^* = \gamma\frac{\varphi(x^k,\tilde{x}^k,\xi^k,\mu)}{\|d(x^k,\tilde{x}^k,\xi^k,\mu)\|^2}$$

with $\gamma \in [1,2)$ as the step size in both (2.3) and (3.1).

Note that Theorem 4.1 also shows that

$$\|x_d^{k+1}(\alpha_k) - x^*\|^2 \leq \|x^k - x^*\|^2 - \rho\|x^k - \tilde{x}^k\|^2$$

and

$$\|x_{pc}^{k+1}(\alpha_k) - x^*\|^2 \leq \|x^k - x^*\|^2 - \rho\|x^k - \tilde{x}^k\|^2,$$

which implies that the sequence $\{x^k\}$ generated by either the descent method (2.3) or the prediction–correction method (3.1) is Féjer monotone with respect to the solution set of (1.1), see, e.g., [3]. Therefore, the following corollary is concluded immediately from Theorem 4.1.

Corollary 4.1 Let x^* be any solution of (1.1). The sequences $\{x^k\}$ and $\{\tilde{x}^k\}$ are generated by the descent method (2.3) (or the prediction–correction method (3.1)), then we have

(1) The sequence $\{x^k\}$ is bounded.
(2) The sequence $\{\|x^k - x^*\|\}$ is nonincreasing.
(3) $\mathrm{Lim}_{k\to\infty}\|x^k - \tilde{x}^k\| = 0$.
(4) The sequence $\{\tilde{x}^k\}$ is bounded.

Now we are ready to prove convergence of the new methods.

Theorem 4.2 *The sequences $\{x^k\}$ generated by either the descent method* (2.3) *or the prediction–correction method* (3.1) *converges to some x^∞ which is a solution of* (1.1).

Proof It follows from (2.11) that

$$(x-\tilde{x}^k)^T(c_kF(\tilde{x}^k)) \geq (x^k-\tilde{x}^k)^T\left((1+\mu)x - (\mu x^k + \tilde{x}^k)\right) + (x-\tilde{x}^k)^T\xi^k, \quad \forall x \in R_+^n.$$

Since $\lim_{k\to\infty}\|x^k - \tilde{x}^k\| = 0$ (see (3) of Corollary. 4.1), it follows from (2.2) that $\lim_{k\to\infty}\|\xi^k\| = 0$. Note that both $\{x^k\}$ and $\{\tilde{x}^k\}$ are bounded and $c_k \geq c > 0$, we have

$$\liminf_{k\to\infty}(x-\tilde{x}^k)^T F(\tilde{x}^k) \geq 0, \quad \forall x \in R_+^n.$$

Since $\{\tilde{x}^k\}$ is bounded (see (4) of Corollary. 4.1), it has at least a cluster point. Let x^∞ be a cluster point of $\{\tilde{x}^k\}$ and the subsequence $\{\tilde{x}^{k_j}\}$ converges to x^∞. It follows that

$$\liminf_{j\to\infty}(x - \tilde{x}^{k_j})^T F(\tilde{x}^{k_j}) \geq 0, \quad \forall x \in R^n_+$$

and consequently

$$(x - x^\infty)^T F(x^\infty) \geq 0, \quad \forall x \in R^n_+,$$

which means that x^∞ is a solution of (1.1). Hence, from (2) of Corollary. 4.1, we have

$$\|x^{k+1} - x^\infty\|^2 \leq \|x^k - x^\infty\|^2, \quad \forall k \geq 0. \tag{4.8}$$

Since $\tilde{x}^{k_j} \to x^\infty (j \to \infty)$ and $x^k - \tilde{x}^k \to 0 (k \to \infty)$, for any given $\varepsilon > 0$, there exists an $l > 0$ such that

$$\|\tilde{x}^{k_l} - x^\infty\| < \varepsilon/2 \qquad \text{and} \quad \|x^{k_l} - \tilde{x}^{k_l}\| < \varepsilon/2. \tag{4.9}$$

Therefore, for any $k \geq k_l$, it follows from (4.8) and (4.9)that

$$\|x^k - x^\infty\| \leq \|x^{k_l} - x^\infty\| \leq \|x^{k_l} - \tilde{x}^{k_l}\| + \|\tilde{x}^{k_l} - x^\infty\| \leq \varepsilon.$$

This implies that the sequence $\{x^k\}$ converges to x^∞, which is a solution of (1.1).

5 Numerical experiments

In this section, we apply the new derived descent method and prediction–correction method to a traffic equilibrium problem, which is a classical and important problem in transportation, see, e.g., [12,20]. The satisfactory numerical results verify the theoretical assertions in aforementioned sections in computational senses and thus demonstrate that the inexact criterion (2.2) is preferable to (1.6) in practice. In particular, we compare the numerical results of the new derived methods to the hybrid methods in [19,21], respectively.

We first illustrate the traffic equilibrium problem. Consider a network $[N, L]$ of nodes N and directed links L, which consists of a finite sequence of connecting links with a certain orientation. Let a, b, etc., denote the links; p, q, etc., denote the paths; ω denote an origin/destination (O/D) pair of nodes of the network; P_ω denote the set of all paths connecting the O/D pair ω; x_p represent the traffic flow on path p; d_ω denote the traffic demand between O/D pair ω, which must satisfy

$$d_\omega = \sum_{p \in P_\omega} x_p,$$

where $x_p \geq 0, \forall p$; f_a denotes the link load on link a, which must satisfy the following conservation of flow equation

$$f_a = \sum_{p \in P} \delta_{ap} x_p,$$

where

$$\delta_{ap} = \begin{cases} 1, & \text{if } a \text{ is contained in path } p; \\ 0, & \text{otherwise.} \end{cases}$$

Let A be the path-arc incidence matrix of the given problem and $f = \{f_a, a \in L\}$ be the vector of the link load. Since x is the path-flow, f is given by

$$f = A^T x.$$

In addition, let $t = \{t_a, a \in L\}$ be the vector of link costs, with t_a denoting the user cost of traversing link a which is given by

$$t_a(f_a) = t_a^0\left[1 + 0.15\left(\frac{f_a}{C_a}\right)^4\right], \tag{5.1}$$

where t_a^0 is the free-flow travel cost on link a and C_a is the designed capacity of link a. Then t is a mapping of the path-flow x and its mathematical form is

$$t(x) := t(f) = t(A^T x).$$

Note that the travel cost on the path p denoted by θ_p is

$$\theta_p = \sum_{a \in L} \delta_{ap} t_a(f_a).$$

Let P denote the set of all the paths concerned. Let $\theta = \{\theta_p, p \in P\}$ be the vector of (path) travel cost. Therefore, for a given link travel cost vector t, θ is a mapping of the path-flow x, which is given by

$$\theta(x) := At(x) = At(A^T x).$$

Associated with every O/D pair ω, there is a travel disutility λ_ω, which is defined as following

$$\lambda_\omega(d) = -m_\omega \log(d_\omega) + q_\omega. \tag{5.2}$$

Note that both the path costs and the travel disutilities are functions of the flow pattern x.

The traffic network equilibrium problem is to seek the path flow pattern x^*, which induces a demand pattern $d^* = d(x^*)$, for every O/D pair ω and each path $p \in P_\omega$,

$$F_p(x) = \theta_p(x) - \lambda_\omega(d(x)).$$

The problem can be reduced to a monotone NCP in the space of path-flow pattern x: Find $x \in \mathcal{R}^n$ such that

$$x \geq 0, \qquad F(x) \geq 0 \qquad \text{and} \quad x^T F(x) = 0. \tag{5.3}$$

In particular, we test the example studied in [12,20]. The network is depicted in Fig. 1. The free-flow travel cost and the designed capacity of links in (5.1) are given in Table 1, the O/D pairs and the coefficient m and q in the disutility function (5.2) are given in Table 5. For this example, there are together 12 paths for the four given O/D pairs as listed in Table 5.

All codes are written by Matlab 7.0 and run on an IBM T42 laptop. The stopping criterion is

$$\|\min\{x, F(x)\}\|_\infty \leq \varepsilon.$$

We test the above traffic equilibrium problem with different ε and compare the new derived methods to the hybrid methods in [19,21]. The initial point is $x^0 = (0.05, 0.05, \ldots, 0.05)^T$; $\mu = 0.01$, and $c_0 = 1$, as in [21]. Let $(\eta, \gamma) = (0.95, 1.95)$ as

Fig. 1 The network used for the numerical test

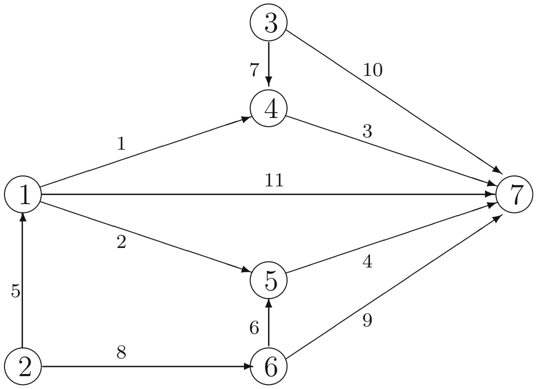

Table 1 The free-flow travel cost and the designed capacity of links in (5.1)

Link	Free-flow Travel time t_a^0	Capacity C_a	Link	Free-flow Travel time t_a^0	Capacity C_a
1	6	200	7	5	150
2	5	200	8	10	150
3	6	200	9	11	200
4	16	200	10	11	200
5	6	100	11	15	200
6	1	100	–	–	–

in [21] when we apply the hybrid methods in [19,21]. Let $(r, \eta, \gamma) = (1.5, 0.82, 1.99)$ when the new descent method is applied and $(r, \eta, \gamma) = (1.5, 0.81, 1.99)$ when the new prediction–correction method is applied. Finally, how to choose the proximal parameter c_k deserves further illustrations. Although theoretically $c_k \in [c, \infty)$ guarantees convergence of the algorithms, it is necessary to choose c_k self-adaptively during the implementations for achieving satisfactory numerical performance. We adopt the strategy in [13] in our numerical experiments. In particular, let

$$\kappa_1 := \frac{|(\xi^k)^T(x^k - \tilde{x}^k)|}{\|x^k - \tilde{x}^k\|^2} \qquad \text{and} \quad \kappa_2 := \frac{\|\xi^k\|}{\|x^k - \tilde{x}^k\|}$$

then

$$c_{k+1} := \begin{cases} c_k * 0.8/\kappa_2, \text{ if } \kappa_1 > \eta \quad \text{or} \quad \kappa_2 > r(1+\mu), \\ c_k * 0.7/\kappa_2, \text{ if } \kappa_2 \le 0.1 \\ c_k, \qquad\qquad \text{otherwise.} \end{cases}$$

The respective iterative numbers (No. Iter.) and computational times (CPU (s.)) are reported in the following tables (Tables 3, 4).

For the case that $\varepsilon = 10^{-8}$, the optimal path flow and link flow are given in Tables 5 and 6, respectively.

The above numerical experiments show that both the descend method (2.3) and the prediction–correction method (3.1) method solve the traffic equilibrium problem very efficiently. But for this particular traffic equilibrium problem, the iterative numbers and computational time of (3.1) are no greater than those of (2.3), therefore it is preferable slightly. In addition, the comparisons to the hybrid methods in

Table 2 The O/D pairs and the coefficient m and q in (5.2)

No. of the pair	O/D pair	m_ω	q_ω
1	(1,7)	25	25 log 600
2	(2,7)	33	33 log 500
3	(3,7)	20	20 log 500
4	(6,7)	20	20 log 400

Table 3 Comparison of the new descent method to the method in [19]

ε	The method in [19]		The new descent method	
	No. Iter.	CPU (s)	No. Iter.	CPU (s)
10^{-3}	122	0.110	82	0.040
10^{-4}	152	0.125	104	0.050
10^{-5}	184	0.144	126	0.081
10^{-6}	213	0.160	144	0.090
10^{-7}	244	0.160	163	0.090
10^{-8}	275	0.188	185	0.125

Table 4 Comparison of the new PC method to the method in [21]

ε	The method in [21]		The new PC method	
	No. Iter.	CPU (s)	No. Iter.	CPU (s)
10^{-3}	113	0.110	80	0.040
10^{-4}	143	0.121	100	0.050
10^{-5}	174	0.130	119	0.072
10^{-6}	203	0.140	134	0.080
10^{-7}	235	0.160	151	0.090
10^{-8}	263	0.180	171	0.110

Table 5 The optimal path flow

O/D pairs	Path No.	Link on the path	Optimal path-flow
	1	(1,3)	165.3145
O/D pair (1,7)	2	(2,4)	0
	3	(11)	138.5735
	4	(5,1,3)	82.5281
	5	(5,2,4)	0
O/D pair (2,7)	6	(5,11)	55.7871
	7	(8,6,4)	0
	8	(8,9)	87.0260
O/D pair (3,7)	9	(7,3)	19.7549
	10	(10)	229.9747
O/D pair (6,7)	11	(9)	178.5600
	12	(6,4)	0

Table 6 The optimal link flow

Link No.	Link flow	Link No.	Link flow	Link No.	Link flow	Link No.	Link flow
1	247.8426	4	0	7	19.7549	10	229.9747
2	0	5	138.3152	8	87.0260	11	194.3606
3	267.5974	6	0	9	265.5860	–	–

[19,21] demonstrate that the new inexact criterion (2.2) is more practical than (1.6) in computational senses.

6 Conclusion

This paper contributes a new practical inexact criterion for solving NCP by inexact-type LQP methods. Based on the reduced approximate solutions of the involved *LQP system*, two new efficient methods are derived. Both of these methods are very easy to implement and the involved computational loads are very tiny. Numerical applications to some traffic equilibrium problems demonstrate that the new methods are more efficient than some existing methods and thus verify that the new inexact criterion is attractive in practice. How to design other efficient LQP-based numerical algorithms for variational inequalities and NCP are worthy of further extensive investigations.

Acknowledgement This author was supported by an internal grant of Antai College of Economics and Management, Shanghai Jiao Tong University.

References

1. Auslender, A., Haddou, M.: An interior proximal point method for convex linearly constrained problems and its extension to variational inequalities. Math. Program. **71**, 77–100 (1995)
2. Auslender, A., Teboulle, M., Ben-Tiba, S.: A logarithmic-quadratic proximal method for variational inequalities. Comput. Optim. Appl. **12**, 31–40 (1999)
3. Bauschke, H.H., Combettes, P.L.: A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces. Math. Oper. Res. **26**(2), 248–264 (2001)
4. Burachik, R.S., Iusem, A.N.: A generalized proximal point alogrithm for the variational inequality problem in a Hilbert space. SIAM J. Optim. **8**, 197–216 (1998)
5. Burachik, R.S., Svaiter, B.F.: A relative error tolerance for a family of generalized proximal point methods. Math. Oper. Res. **26**(4), 816–831 (2001)
6. Eaves, B.C.: On the basic theorem of complementarity. Math. Program. **1**, 68–75 (1971)
7. Eckstein, J.: Approximating iterations in Bregman-function-based proximal algorithms. Math. Program. **83**, 113–123 (1998)
8. Güler, O.: On the convergence of the proximal point algorithm for convex minimization. SIAM J. Control Optim. **29**, 403–419 (1991)
9. Han, D.R., He, B.S.: A new accuarcy criterion for approximate proximal point algorithms. J. Math. Anal. Appl. **263**, 343–353 (2001)
10. Harker, P.T., Pang, J.S.: Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications. Math. Program. **48**, 161–220 (1990)
11. He, B.S., Liao, L.Z., Yang, Z.H.: A new approximate proximal point algorithm for maximal monotone operator. Sci. China Ser. A **46**(2), 200–206 (2003)
12. He, B.S., Yang, Z.H., Yuan, X.M.: An approximate proximal-extragradient type method for monotone variational inequalities. J. Math. Anal. Appl. **300**(2), 362–374 (2004)

13. He, B.S., Liao, L.Z., Yuan, X.M.: A LQP method-based interior prediction-correction for nonliner complementarity problems. J. Comput. Math. **24**(1), 33–44 (2006)
14. He, B.S., Xu, Y., Yuan, X.M.: A Logarithmic-quadratic proximal prediction-correction method for structured monotone variational inequalities. Comput. Optim. Appl. **35**, 19–46 (2006)
15. Martinet, B.: Regularization d'inéquations variationnelles par approximations succesives. Rev. Fr. Inform. Rech. Opérat. **2**, 154–159 (1970)
16. Moré, J.J.: Global methods for nonlinear complementarity problems. Math. Oper. Res. **21**, 589–614 (1996)
17. Rockafellar, R.T.: Monotone operators and the proximal point algorithm. SIAM J. Optim. **126**, 877–898 (1976)
18. Teboulle, M.: Convergence of proximal-like algorithms. SIAM J. Optim. **7**, 1069–1083 (1997)
19. Xu, Y., Bnouhachem, A.: A new logarithmic-quadratic proximal prediction–correction method for NCP. manuscript (2006)
20. Yang, H., Bell, M.G.H.: Traffic restraint, road pricing and network equilibrium. Transport. Res. B **31**, 303–314 (1997)
21. Yuan, X.M.: The prediction–correction approach to nonlinear complementarity problems. Eur. J. Oper. Res. **176**(3), 1357–1370 (2007)
22. Zarantonello, E.H. (1971) Projection on covex sets in Hilbert space and sepctral theory. In: Zarantonello, E.H. (eds) Contributation to Nonlinear Functional Analysis. Academic Press, New York