An Infeasible Interior Proximal Method for the Variational Inequality Problem

R. S. Burachik[*]
School of Mathematics and Statistics
University of South Australia
Mawson Lakes, S.A. 5095, Australia
regina.burachik@unisa.edu.au

J. O. Lopes[†]
Universidade Federal do Piauï¿½ CCN-UFPI,CP 64.000,
Teresina, PI, CEP 64.000-000, Brazil.
jurandir@ufpi.br

G. J. P. Da Silva[‡]
Universidade Federal de Goiás, IME/UFG, CP 131,
Goiï¿½ia, GO, CEP 74.001-970, Brazil.
geci@mat.ufg.br

April 25, 2006

Abstract

In this paper, we propose an infeasible interior proximal method for solving variational inequality problems with maximal monotone operators and linear constraints. The interior proximal method proposed by Auslender, Teboulle and Ben-Tiba [3] is a proximal method using a

[*]Partially supported by PRONEX–Optimization.
[†]CNPq
[‡]PACDT/CNPq

distance-like barrier function and it has a global convergence property under mild assumptions. However this method is applicable only to problems whose feasible region has nonempty interior. The algorithm proposed in this paper is applicable to problems whose feasible region may have empty interior. Moreover, a new kind of inexact scheme is used. We present a full convergence analysis for our algorithm.
Key words. maximal monotone operators, outer approximation algorithm, interior point method, global convergence.

AMS subject classifications.

1 Introduction

sec:1

Let $C \subset I\!R^n$ be a closed and convex set, and $T\colon I\!R^n \rightrightarrows I\!R^n$ be a maximal monotone point-to-set operator. We consider the *variational inequality problem associated with* T *and* C: Find $\overline{x}$ such that there exists $\overline{v} \in T(\overline{x})$ satisfying

$$\langle \overline{v}, y - \overline{x} \rangle \geq 0, \ \text{for all } y \in C. \tag{1.1}$$ e1

This problem will be denoted by $VIP(T, C)$. In the particular case in which T is the subdifferential of $f\colon I\!R^n \to I\!R \cup \{\infty\}$, where f is proper, convex and lower semicontinuous, (1.1) reduces to the *constrained convex optimization problem*: Find $\overline{x}$ such that

$$f(\overline{x}) \leq f(y), \ \text{for all } y \in C. \tag{1.2}$$ e2

We will be concerned in this work with C a polyhedral set on $I\!R^n$ defined by

$$C := \{x \in I\!R^n |\ Ax \leq b\}, \tag{1.3}$$ c2

where A is an $m \times n$ real matrix, $b \in I\!R^m$ and $m \geq n$. Well-known methods for solving $VIP(T, C)$ are the so-called *generalized proximal* schemes, which involve a regularization term that incorporates the constraint set C in such a way that all the subproblems have solutions in the interior of C. For this reason, these methods are also called *interior proximal methods*. Examples of these regularizing functionals are the Bregman distances (see, e.g. [1, 8, 12, 13, 19, 24]), φ-divergences ([25, 5, 14, 17, 18, 26, 27]) and *log-quadratic* regularizations ([3, 4]). Being *interior* point methods, it is a basic assumption that the topological interior of C is nonempty. Otherwise, the iterates are not well-defined. However, a set C as above may usually have empty interior. In order to solve problem (1.2) for an arbitrary set $C \neq \emptyset$ of the kind given in (1.3), Yamasita et al.[28] devised an interior-point scheme in which the subproblems deal with a constraint set C^k given by

$$C^k := \{x \in I\!R^n |\ Ax \leq b + \delta^k\}, \tag{1.4}$$ ck

where the vectors δ^k have positive coordinates and are such that $\sum_1^\infty \|\delta^k\| < \infty$. So, if $C \neq \emptyset$, it holds $C \subset int\, C^k$ and hence a regularizing functional D_k can be associated to the set C^k. More precisely, the subproblems find an approximate solution $x^k \in int\, C^k$ of the inclusion

$$0 \in \lambda_k \partial_{\varepsilon_k} f(x^k) + \nabla_1 D_k(x^k, x^{k-1}),$$

where $\lambda_k > 0$ and $\partial_\varepsilon f$ is the ε-subdifferential of f[6] and D_k is the regularization functional proposed by Auslender, Teboulle and Ben-Tiba [3, 4]. Yamasita et al. allow an error e^k in the inclusion above, and they prove convergence under summability assumptions on the “error” sequences $\{\varepsilon_k\}$, $\{e^k\}$ and $\{\delta^k\}$. We want to extend the above scheme to the more general problem (1.1), so we are concerned with iterations of the kind: Find an approximate solution $x^k \in int\, C^k$ of

$$0 \in \lambda_k T^{\varepsilon_k}(x^k) + \nabla_1 D_k(x^k, x^{k-1}),$$

where $\lambda_k > 0$ and T^ε is an enlargement of the operator T [10, 9]. In our scheme, we will require no summability assumption on the parameters $\{\varepsilon_k\}$ and $\{e^k\}$ (the latter sequence controlling the error in the inclusion above). Instead, we define a criterium which can be checked at each iteration. However, we still need here the summability assumptions on $\{\delta_k\}$ for obtaining convergence results. Our relative error analysis is inspired in the one given by Burachik and Svaiter in [11], which yields a more pratical framework. Additionally, for the proposed algorithm we establish global convergence with assumptions weaker than those used in [28].

The paper is organized as follows. In Section 2 we give some basic definitions and properties of the family of regularizations, as well as some known results on monotone operators. In the same section, the extension T^ε is reviewed, together with its elementary features. In Section 3, we describe the algorithm, prove its well-definedness and give its inexact version. We finish this section with the convergence proof.

2 Basic assumptions and properties

sec:basic

A point-to-set valued map $T: I\!R^n \rightrightarrows I\!R^n$ is an operator which associates with each point $x \in I\!R^n$ a set (possibly empty) $T(x) \subset I\!R^n$. The domain and the graph of a point-to-set valued map T are defined as:

$$Dom\ T := \{x \in I\!R^n |\ T(x) \neq \emptyset\},$$

$$G(T) := \{(x, v) \in I\!R^n \times I\!R^n |\ x \in Dom\ T,\ v \in T(x)\}.$$

A point-to-set operator T is said to be *monotone* if

$$\langle u - v, x - y \rangle \geq 0, \quad \forall\ u \in T(x),\ v \in T(y).$$

A monotone operator T is said to be *maximal* when its graph is not properly contained in the graph of any other monotone operator. The well-known result below has proved in [23, Theorem 1]. Denote by $ir\,A$ the relative interior of the set A.

pro:2.2

Proposition 2.1 *Let T_1, T_2 maximal monotone operators. If $ir\,D(T_1) \cap ir\,D(T_2) \neq \emptyset$, then $T_1 + T_2$ is maximal monotone.*

The theorem bellow is an essential tool for our analysis. We denote by f_∞ the asymptotic function [2, Definition 2.5.1] associated with the function $f : I\!R^n \to I\!R \cup \{+\infty\}$.

teo:A

Theorem 2.2 *([3, Proposition 3.1]) Let $f: I\!R^n \to I\!R \cup \{+\infty\}$ be a closed proper convex function with $dom(f)$ open. Assume that f is differentiable on $dom(f)$ and such that $f_\infty(d) = +\infty \ \ \forall\ d \neq 0$. Let A be an $m \times n$ matrix with $m \geq n$ and $rank = n$, $\tilde{b} \in I\!R^m$ with $(\tilde{b} - A(I\!R^n)) \cap dom(f) \neq \emptyset$, and set $h(x) := f(\tilde{b} - Ax)$. Let $\tilde{T}: I\!R^n \rightrightarrows I\!R^n$ be a maximal monotone operator such that $D(\tilde{T}) \cap dom(h) \neq \emptyset$ and set*

$$U(x) := \begin{cases} \tilde{T}(x) + \nabla h(x) & \text{if } x \in D(\tilde{T}) \cap D(\nabla h), \\ \emptyset & \text{otherwise.} \end{cases}$$

Then $\nabla h(x)$ is onto. Moreover, there exists x solution of equation

$$0 \in U(x),$$

which is unique if f is strictly convex on its domain.

We describe below the family of regularizations we will use. From now on, the function $\varphi\colon I\!R_+ \to (-\infty, \infty]$ is given by

$$\varphi(t) := \mu h(t) + (\nu/2)(t-1)^2, \tag{2.1}$$

where h is a closed and proper convex function satisfying the following additional properties:

(1) h is twice continuously differentiable on $int(dom\ h) = (0, +\infty)$,

(2) h is strictly convex on its domain,

(3) $\lim_{t\to 0^+} h^{'}(t) = -\infty$,

(4) $h(1) = h^{'}(1) = 0$ and $h^{''}(1) > 0$, and

(5) for $t > 0$

$$h^{''}(1)\left(1 - \frac{1}{t}\right) \le h^{'}(t) \le h^{''}(1)(t-1). \tag{2.2}$$

re

Items (1) − (4) and (1) − (5) were used in [4] to define, respectively, the families Φ and Φ_2. The positive parameters μ, ν shall verify the following inequality

$$\nu > \mu h^{''}(1) > 0. \tag{2.3}$$

ni

Note that conditions above imply

$$\mu h^{''}(1)\left(1 - \frac{1}{t}\right) + \nu(t-1) \le \varphi^{'}(t) < \nu h^{''}(1)(t-1), \tag{2.4}$$

des varphi

therefore $\lim\limits_{t\to\infty} \varphi^{'}(t) = +\infty$.

The generalized distance induced by φ, is denoted by $d_\varphi(x, y)$ and defined as:

$$d_\varphi(x, y) := \sum_{i=1}^{N} y_i^2 \varphi(x_i/y_i), \tag{2.5}$$

dfi

for $x, y \in I\!R^n_{++} := \{z \in I\!R^n \mid z_i > 0\ \forall\ i = 1, \ldots, n\}$. Since $\lim\limits_{t\to\infty} \varphi^{'}(t) = +\infty$ follows that $[d_\varphi(\cdot, y^{k-1})]_\infty(d) = +\infty,\ \forall\ d \neq 0$ and denoting by ∇_1 the gradient with respect to the first variable, it holds that $[\nabla_1 d_\varphi(x, y)]_i = y_i \varphi'(x_i/y_i)$ for all $i = 1, \ldots, n$.

The following lemma has a crucial role in the convergence analysis. Its first part has been established in [3]. Define

$$\theta := \nu + \rho\mu, \quad \tau := \nu - \rho\mu \text{ and } \rho := h''(1). \tag{2.6}$$ tau

le:cru **Lemma 2.3** *For all* $w, z \in \mathbb{R}^n_{++}$ *and* $v \in \mathbb{R}^n_+ := \{z \in \mathbb{R}^n \mid z_i \geq 0 \ \forall\ i = 1, \ldots, n\}$*, it holds that*

(i) $\langle \nabla_1 d_\varphi(w, z), w - v\rangle \geq \frac{\theta}{2}(\|w - v\|^2 - \|z - v\|^2) + \frac{\tau}{2}\|w - z\|^2$;

(ii) $\langle v, \nabla_1 d_\varphi(w, z)\rangle \leq \theta\|v\|\|w - z\|$.

Proof. For part (i), see [3, Lemma 3.4]. We proceed to prove (ii). Since $\varphi(t) = \mu h(t) + \frac{\nu}{2}(t-1)^2$, we have that $\varphi'(t) = \mu h'(t) + \nu(t-1)$. By (2.2) and (2.3) we get $\varphi'(t) \leq (\nu + \rho\mu)(t - 1)$. Letting $t = \frac{w_i}{z_i}$ and multiplying both sides by $v_i z_i$ yield

$$v_i z_i \varphi'(\frac{w_i}{z_i}) \leq \theta v_i z_i (\frac{w_i}{z_i} - 1),$$

for all $i = 1, \ldots, n$. Therefore,

$$\langle v, \nabla_1 d_\varphi(w, z)\rangle \leq \theta\langle v, w - z\rangle.$$

Using the Cauchy-Schwartz inequality in the expression above, we get (ii). ■

The result below is known as Hoffman's lemma[15].

le:ho **Lemma 2.4** *Let* $C = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ *and* $C^k = \{x \in \mathbb{R}^n \mid Ax \leq b + \delta^k\}$ *where* A *is matrix* $m \times n$ *with* $m \geq n$ *and* $b, \delta^k \in \mathbb{R}^m$*. Given* $x^k \in C^k$ *there exists a constant* $\alpha > 0$ *such that*

$$dist(x^k, C) := \inf_{y \in C} \|y - x^k\| = \|p^k - x^k\| \leq \alpha\|\delta^k\|$$

where p^k *is projection of* x^k *in* C*.*

We recall next two technical results on nonnegative sequences of real numbers. The first one was taken from [21] and the second from [20].

le:se **Lemma 2.5** *Let* $\{\sigma_k\}$ *and* $\{\beta_k\}$ *be nonnegative sequences of real numbers satisfying:*

(i) $\sigma_{k+1} \leq \sigma_k + \beta_k$;

le:se1

(ii) $\sum_{k=1}^{\infty} \beta_k < \infty$.

Then the sequence $\{\sigma_k\}$ *converges.*

Lemma 2.6 *Let* $\{\lambda_k\}$ *be a sequence of positive numbers, and* $\{a_k\}$ *be a sequence of real numbers. Let* $\sigma_k := \sum_{j=1}^{k} \lambda_j$ *and* $b_k := {\sigma_k}^{-1} \sum_{j=1}^{k} \lambda_j a_j$. *If* $\sigma_k \to \infty$, *then*

(i) $\liminf_{k\to\infty} a_k \le \liminf_{k\to\infty} b_k \le \limsup_{k\to\infty} b_k \le \limsup_{k\to\infty} a_k$;

(ii) *If* $\lim_{k\to\infty} a_k = a < \infty$, *then* $\lim_{k\to\infty} b_k = a$.

In our analysis, we will relax the inclusion $v^k \in T(x^k)$, by means of an ε-extension of the operator T [9]: Given T a monotone operator, define

$$T^{\varepsilon}(x) := \{v \in I\!R^N \mid \langle v - w, x - y\rangle \ge -\varepsilon \ \forall y \in I\!R^N,\, w \in T(y)\}. \qquad (2.7)$$

teps:def

This extension has many useful properties, similar to the ε-subdifferential of a proper closed convex function f. Indeed, when $T = \partial f$, we have (see [9])

$$\partial_{\varepsilon} f(x) \subseteq T^{\varepsilon}(x).$$

For an arbitrary maximal monotone operator T, the relation

$$T^0(x) = T(x)$$

holds trivially. Furthermore, for $\varepsilon' \ge \varepsilon \ge 0$, we have

$$T^{\varepsilon}(x) \subset T^{\varepsilon'}(x).$$

In particular, for each $\varepsilon \ge 0$,

$$T(x) \subset T^{\varepsilon}(x).$$

3 The Algorithm

sec:alg

In this section, we propose an infeasible interior proximal method for the solution of $VIP(T,C)$ (1.1). To state formally our algorithm, we consider

$$C^k = \{x \in I\!R^n \mid\, Ax \le b + \delta^k\} \text{ where } \delta^k \in I\!R^m_{++} \text{ and } \sum_{k=1}^{\infty} \|\delta^k\| < \infty,$$

which is considered a perturbation of the original the constraint set C. Moreover, if $C \neq \emptyset$, then $C \subset int\, C^k \neq \emptyset$ for all k. Since $\delta^k \to 0$ as $k \to \infty$, the sequence $\{C^k\}$ converges to the set C. Now, if a_i denotes the row i of the matrix A, for each $x \in C^k$ we consider $y^k(x) = (y_1^k(x), y_2^k(x), ..., y_m^k(x))^T$, where $y_i^k(x) = b_i + \delta_i^k - \langle a_i, x\rangle$ with $i = 1, 2, ..., n$. Therefore, we have the function $D_k : int\, C^k \times int\, C^k \to I\!R$ defined by

$$D_k(x, z) = d_\varphi(y^k(x), y^k(z)). \tag{3.1}$$

¿From the definition of d_φ, for each $x^k \in int\, C^k$, $x^{k-1} \in int\, C^{k-1}$, we have

$$\nabla_1 D_k(x^k, x^{k-1}) = -A^T \nabla_1 d_\varphi \left(y^k(x^k), y^{k-1}(x^{k-1})\right). \tag{3.2}$$

gradDk

In the method proposed in [28] for the convex optimization problem (1.2) with C defined as in (1.3), the exact algorithm of the iteration k is given by: For $\lambda_k > 0$, $\delta^k > 0$ and $(x^{k-1}, y^{k-1}) \in int\, C^{k-1} \times I\!R^m_{++}$, find $(x, y) \in int\, C^k \times I\!R^m_{++}$ and $u \in I\!R^n$ such that

$$\begin{cases} u \in \partial f(x), & \\ \lambda_k u + \nabla_1 D_k(x, x^{k-1}) & = \; 0, \\ y - (b - Ax) & = \; \delta^k, \end{cases}$$

where $y \in I\!R^m_{++}$ can be seen as a slack variable associated to $x \in int\, C^k$.

The corresponding *inexact* iteration is given by:

$$\begin{cases} \widetilde{u} \in \partial_{\varepsilon_k} f(\widetilde{x}), & \\ \lambda_k \widetilde{u} + \nabla_1 D_k(\widetilde{x}, x^{k-1}) & = \; e^k, \\ \widetilde{y} - (b - A\widetilde{x}) & = \; \delta^k. \end{cases} \tag{3.3}$$

yama

Following this approach, the exact version of our algorithm is obtained replacing ∂f by an arbitrary maximal monotone operator T. Namely, given $\lambda_k > 0$, $\delta^k > 0$ and $(x^{k-1}, y^{k-1}) \in int\, C^{k-1} \times I\!R^m_{++}$, find $(x, y) \in int\, C^k \times I\!R^m_{++}$ and $u \in I\!R^n$ such that

$$\begin{cases} u \in T(x), & \\ \lambda_k u + \nabla_1 D_k(x, x^{k-1}) & = \; 0, \\ y - (b - Ax) & = \; \delta^k. \end{cases} \tag{3.4}$$

eq:6.1a

It is important to guarantee the existence of $(x^k, y^k) \in int\, C^k \times I\!R^m_{++}$ satisfying (3.4). In fact, the next proposition shows that there exists a unique pair $(x^k, y^k) \in int\, C^k \times I\!R^m_{++}$ satisfying (3.4) under the following two assumptions:

(H_1) $ir\, C \cap ir\, D(T) \neq \emptyset$;

(H_2) rank(A)=n (and therefore, A injective).

pro:5.1 **Proposition 3.1** *Suppose that (H_1) and (H_2) hold. For every $\lambda_k > 0$, $\delta^k > 0$ and $(x^{k-1}, y^{k-1}) \in int\, C^k \times I\!R^m_{++}$, there exists a unique pair $(x^k, y^k) \in int\, C^k \times I\!R^m_{++}$ satisfying (3.4).*

Proof. Define the operator $\widetilde{T}^k(x) := T(x) + N_{C^k}(x) + {\lambda_k}^{-1}\nabla h(x)$, where $h := D_k(\cdot, x^{k-1})$. We will prove that we are in the conditions of Theorem 2.2 for $\widetilde{T} := T + N_{C^k}$, $f(\cdot) := d_\varphi(\cdot, y^{k-1})$ and $\tilde{b} := b + \delta^k$. Indeed, the operator $T + N_{C^k}$ is maximal monotone by (H_1) and the fact that $C \subseteq C^k$ (we are using here Proposition 2.1(ii)). The function $d_\varphi(\cdot, y^{k-1})$ is by definition convex, proper and differentiable on its (open) domain $I\!R^m_{++}$ and $[d_\varphi(\cdot, y^{k-1})]_\infty(d) = +\infty,\ \forall\ d \neq 0$.

By (H_2), A has maximal rank. We claim that $(b+\delta^k - A(I\!R^n)) \cap dom(d_\varphi) \neq \emptyset$. Indeed, fix $x \in C$. It holds that

$$b + \delta^k - Ax \geq \delta^k > 0, \tag{3.5}$$ ult

and therefore $b + \delta^k - Ax \in I\!R^n_{++} = dom(d_\varphi)(\cdot, y^{k-1})$.

The only hypothesis that remains to be checked is: $D(\tilde{T}) \cap dom(h) \neq \emptyset$, where $dom(h) = int\, C^k$. Indeed, by (H_1) and by definition of the C^k we get

$$\emptyset \neq C \cap D(T) \subset int\, C^k \cap D(T) \subset D(\tilde{T}).$$

Hence $\emptyset \neq C \cap D(T) \subset D(\tilde{T}) \cap int\, C^k = D(\tilde{T}) \cap dom(h)$. So the hypotheses of Theorem 2.2 are satisfied and therefore there exists x^* a solution of the equation

$$0 \in T(x) + N_{C^k}(x) + {\lambda_k}^{-1}\nabla_1 D_k(x, x^{k-1}). \tag{3.6}$$ zero

This solution is unique, because $d_\varphi(\cdot, y^{k-1})$ is strictly convex on its domain.

In this way, there exists $u^k \in T(x^k)$, $v^k \in N_{C^k}(x^k)$ and $z^k = \nabla_1 D_k(x^k, x^{k-1}) = \nabla_1 d_\varphi(b + \delta^k - A(x^k), y^{k-1})$, such that

$$0 = u^k + v^k + z^k. \tag{3.7}$$ zero1

Making $b+\delta^k - Ax^k =: y^k$ we have that y^k is also unique. Since $y^k \in I\!R^m_{++}$, it holds that $x^k \in int\, C^k$, thus $v^k = 0$. Hence by (3.7) there exists a unique pair $(x^k, y^k) \in int\, C^k \times I\!R^m_{++}$ satisfying

$$\left\{\begin{array}{lcl} u^k \in T(x^k), & & \\ u^k + \lambda^{-1}\nabla_1 D_k(x^k, x^{k-1}) & = & 0, \\ y^k - (b - Ax^k) & = & \delta^k, \end{array}\right.$$

which completes the proof. ∎

To deal with approximations, we will relax the inclusion and the equation of the exact system (3.4) in a way similar to (3.3):

$$\left\{\begin{array}{lcl} \widetilde{u} \in T^{\varepsilon_k}(\widetilde{x}), & & \\ \lambda_k \widetilde{u} + \nabla_1 D_k(\widetilde{x}, x^{k-1}) & = & e^k, \\ \widetilde{y} - (b - A\widetilde{x}) & = & \delta^k, \end{array}\right. \qquad (3.8)$$ eq:6.2

where T^ε is the enlargement of T given in (2.7).

In the exact solution, we have $\varepsilon_k = 0$ and $e^k = 0$. An approximate solution should have ε_k and e^k "small ".

Our aim is to use a relative error criteria as the one used in [11] to control the size of ε_k and e^k. The intuitive idea is to perform an extragradient step from x^{k-1} to x, using the direction $\tilde{u}$ (see (3.9)), and then check whether the "error terms" of the iteration, given by $\varepsilon_k + \langle \tilde{u}, \tilde{x} - x\rangle$ and $\|\tilde{y} - y\|$ are small enough with respect to the previous step.

def:reg **Definition 3.2** . *Let $\sigma \in [0, 1)$ and $\gamma > 0$. We say that $(\widetilde{x}, \widetilde{y}, \widetilde{u}, \varepsilon_k)$ in (3.8) is an* approximated solution *of system (3.4) with tolerance σ and γ if for (x, y) such that*

$$\left\{\begin{array}{lcl} \lambda_k \widetilde{u} + \nabla_1 D_k(x, x^{k-1}) & = & 0, \\ y - (b - Ax) & = & \delta^k. \end{array}\right. \qquad (3.9)$$ eq:6.3

is holds that

$$\lambda_k(\varepsilon_k + \langle \widetilde{u}, \widetilde{x} - x\rangle) \le \sigma\frac{\tau}{2}\|y - y^{k-1}\|^2, \qquad (3.10)$$ eq:6.4

$$\|\widetilde{y} - y\| \le \gamma\|y - y^{k-1}\|. \qquad (3.11)$$ eq:6.5

ob:6.1 **Remark 3.3** .

(i) *Since the domain of $d_\varphi(\cdot, y^{k-1})$ is $I\!R^m_{++}$, for $\widetilde{x}, \widetilde{u}, \varepsilon_k$ and x as in Definition 3.2, it holds that*

$$\widetilde{x}, x \in int\, C^k.$$

(ii) *If (x, y, u) verifies (3.4), then $(x, y, u, 0)$ is an approximated solution of system (3.4) with tolerance σ, γ for any $\sigma \in [0, 1)$ and $\gamma > 0$. It is clear that in this case $e^k = 0$. Conversely, if $(\widetilde{x}, \widetilde{y}, \widetilde{u}, \varepsilon_k)$ is an approximated solution of system (3.4) with tolerance $\sigma = 0$ and $\gamma > 0$ arbitrary, then we must have $\varepsilon_k = 0$ and $(\widetilde{x}, \widetilde{y}, \widetilde{u})$ satisfying (3.4). Indeed, since $\gamma > 0$ is arbitrary, we get $y = \widetilde{y}$. Using the fact that A is one-to-one, we get $x = \widetilde{x}$. From the fact that $\sigma = 0$, we conclude that $\varepsilon_k = 0$.*

(iii) *If (H_1) and (H_2) hold, by Proposition 3.1 the system (3.8) with $e^k = 0$ and $\varepsilon_k = 0$ has a solution. By (ii), this solution generates an approximated solution.*

We describe below formally our algorithm, which we call *Extragradient Algorithm.*

Extragradient Algorithm-EA
Initialize: Take $\overline{\lambda} > 0, \sigma \in [0, 1), \gamma > 0, x^0 \in I\!R^n$ and $y^0 \in I\!R^m_{++}$ such that $\delta^0 := y^0 - (b - Ax^0) \in I\!R^m_{++}$.
Iteration: For k=1,2,...,
Step 1. Take λ_k with $\overline{\lambda} \leq \lambda_k$ and $0 < \delta^k < \delta^{k-1}$. Find $(\widetilde{x}^k, \widetilde{y}^k, \widetilde{u}^k, \varepsilon_k)$ an approximated solution of system (3.4) with tolerance σ, γ.
Step 2. Compute (x^k, y^k) such that

$$\begin{cases} \lambda_k \widetilde{u}^k + \nabla_1 D_k(x^k, x^{k-1}) & = & 0, \\ y^k - (b - Ax^k) & = & \delta^k. \end{cases} \tag{3.12}$$

Step 3. Set $k := k + 1$, and return to Step 1.

alg.3

3.1 Convergence analysis

sec:conv

In this section, we prove convergence of the Algorithm above. From now on $\{x^k\}, \{\widetilde{x}^k\}, \{\widetilde{y}^k\}, \{y^k\}, \{\widetilde{u}^k\}, \{\varepsilon_k\}, \{\lambda_k\}$ and $\{\delta_k\}$ are sequences generated by **EA** with approximating criteria (3.10)-(3.11). The main result we shall prove is that the sequence $\{x^k\}$ converges to a solution of $VIP(T, C)$.

The next proposition is essential for the convergence analysis, to show this we need the following further assumptions

(H_3) The solution set S of $VIP(T, C)$ is nonempty.

pro:con1 **Proposition 3.4** . *Suppose that (H_3) holds and let $\overline{x} \in S$ and $\overline{u} \in T(\overline{x})$. Define $\overline{y} := b - A\overline{x}$. Then, for $k = 1, 2, ...$,*

$$\begin{aligned} \|y^k - \overline{y}\|^2 \;\; &\leq \;\; \|y^{k-1} - \overline{y}\|^2 - \tfrac{\tau}{\theta}(1-\sigma)\|y^k - y^{k-1}\|^2 \\ &\quad +2\|\delta^k\| \, \|y^k - y^{k-1}\| + \alpha\tfrac{2}{\theta}\lambda_k\|\overline{u}\| \, \|\delta^k\|, \end{aligned} \tag{3.13}$$ eq:con

where θ, τ are as in (2.6) *and α is as in* Lemma 2.4

Proof. Fix $k > 0$ and take $\widetilde{u}^k \in T^{\varepsilon_k}(\widetilde{x}^k)$. For all $(x, u) \in G(T)$ we have that

$$\lambda_k\langle x - \widetilde{x}^k, u - \widetilde{u}^k\rangle \geq -\lambda_k\varepsilon_k.$$

Therefore,

$$\begin{aligned} \lambda_k\langle x - \widetilde{x}^k, u\rangle \;\; &\geq \;\; \lambda_k\langle x - \widetilde{x}^k, \widetilde{u}^k\rangle - \lambda_k\varepsilon_k \\ &= \;\; \lambda_k\langle x - x^k, \widetilde{u}^k\rangle + \lambda_k\langle x^k - \widetilde{x}^k, \widetilde{u}^k\rangle - \lambda_k\varepsilon_k. \end{aligned} \tag{3.14}$$

Using (3.9) and (3.10) in the inequality above, we get

$$\lambda_k\langle x - \widetilde{x}^k, u\rangle \geq \langle x - x^k, -\nabla_1 D_k(x^k, x^{k-1})\rangle - \sigma\frac{\tau}{2}\|y^k - y^{k-1}\|^2. \tag{3.15}$$ eq:6.6

Now, using (3.2) we have

$$\begin{aligned} \langle x - x^k, -\nabla_1 D_k(x^k, x^{k-1})\rangle \;\; &= \;\; \langle x - x^k, A^T\nabla_1 d_\varphi(y^k, y^{k-1})\rangle \\ &= \;\; \langle A(x - x^k), \nabla_1 d_\varphi(y^k, y^{k-1})\rangle \\ &= \;\; \langle y^k - y, \nabla_1 d_\varphi(y^k, y^{k-1})\rangle - \langle\delta^k, \nabla_1 d_\varphi(y^k, y^{k-1})\rangle, \end{aligned}$$

where $y = b - Ax$. Combining the equality above with (3.15), we get

$$\lambda_k\langle x - \widetilde{x}^k, u\rangle \geq \langle y^k - y, \nabla_1 d_\varphi(y^k, y^{k-1})\rangle - \langle\delta^k, \nabla_1 d_\varphi(y^k, y^{k-1})\rangle - \sigma\frac{\tau}{2}\|y^k - y^{k-1}\|^2.$$

Applying Lemma 2.3 in this inequality yields

$$\lambda_k\langle x - \widetilde{x}^k, u\rangle \geq \frac{\theta}{2}(\|y^k - y\|^2 - \|y^{k-1} - y\|^2) + \frac{\tau}{2}(1-\sigma)\|y^k - y^{k-1}\|^2 - \theta\|\delta^k\| \, \|y^k - y^{k-1}\|. \tag{3.16}$$ eq:5.2

The inequality above is valid in particular for $(x, u) := (\overline{x}, \overline{u})$ with $\overline{x} \in S$ and $\overline{y}$ such that $\overline{y} = b - A\overline{x}$. Therefore,

$$\lambda_k\langle \overline{x} - \widetilde{x}^k, \overline{u}\rangle \geq \frac{\theta}{2}(\|y^k - \overline{y}\|^2 - \|y^{k-1} - \overline{y}\|^2) + \frac{\tau}{2}(1-\sigma)\|y^k - y^{k-1}\|^2 - \theta\|\delta^k\| \, \|y^k - y^{k-1}\|. \tag{3.17}$$ eq:5.3

On the other hand, for $\overline{x} \in S$, there exists $\overline{u} \in T(\overline{x})$ such that

$$\langle \overline{x} - x, \overline{u} \rangle \leq 0 \; \forall \; x \in C.$$

Let p^k be the projection of $\widetilde{x}^k$ onto C. Since $p^k \in C$, we have that

$$\langle \overline{x} - p^k, \overline{u} \rangle \leq 0,$$

and therefore

$$\langle \overline{x} - \widetilde{x}^k, \overline{u} \rangle \leq \langle p^k - \widetilde{x}^k, \overline{u} \rangle.$$

Using the Cauchy-Schwarz inequality and multiplying by $\lambda_k > 0$, we get

$$\lambda_k \langle \overline{x} - \widetilde{x}^k, \overline{u} \rangle \leq \lambda_k \|\overline{u}\| \|\widetilde{x}^k - p^k\|.$$

By Lemma 2.4 we conclude that

$$\lambda_k \langle \overline{x} - \widetilde{x}^k, \overline{u} \rangle \leq \lambda_k \alpha \|\overline{u}\| \|\delta^k\|, \tag{3.18}$$ eq:5.4

for some $\alpha > 0$. Combining (3.17) and (3.18), we get (3.13). ∎

The next corollary guarantees boundedness of the sequence $\{y^k - y^{k-1}\}$.

cor:lim **Corollary 3.5** *Suppose that (H_3) holds, then the sequence $\{\|y^k - y^{k-1}\|\}$ is bounded.*

Proof. Assume the sequence $\{\|y^k - y^{k-1}\|\}$ is unbounded. Then there is a subsequence $\{\|y^k - y^{k-1}\|\}_{k \in K}$ such that $\|y^k - y^{k-1}\| \to \infty$ for $k \in K$, whereas the complementary subsequence $\{\|y^k - y^{k-1}\|\}_{k \notin K}$ is bounded (note that this complementary subsequence could be finite or even empty). From (3.13), we have

$$\|y^k - y^{k-1}\| \left[\frac{\tau}{\theta}(1-\sigma)\|y^k - y^{k-1}\| - 2\|\delta^k\|\right] \leq \|y^{k-1} - \overline{y}\|^2 - \|y^k - \overline{y}\|^2 + \alpha \frac{2}{\theta} \lambda_k \|\overline{u}\| \|\delta^k\|. \tag{3.19}$$ eq:con2

Summing the inequalities (3.13) over $k = 1, 2, ..., n$ gives

$$\sum_{\substack{k = 1, ..., n \\ k \notin K}} \|y^k - y^{k-1}\| \left[\frac{\tau}{\theta}(1-\sigma)\|y^k - y^{k-1}\| - \frac{1}{2}\|\delta^k\|\right] +$$

$$\sum_{\substack{k=1,\ldots,n \\ k\in K}} \|y^k - y^{k-1}\| \left[\frac{\tau}{\theta}(1-\sigma)\|y^k - y^{k-1}\| - \frac{1}{2}\|\delta^k\|\right] \leq$$

$$\|y^0 - \overline{y}\|^2 - \|y^n - \overline{y}\|^2 + \alpha\frac{\tau}{\theta}\overline{\lambda}\|\overline{u}\| \sum_{k=1}^{n} \|\|\delta^k\| \leq \|y^0 - \overline{y}\|^2 + \alpha\frac{\tau}{\theta}\overline{\lambda}\|\overline{u}\| \sum_{k=1}^{n} \|\|\delta^k\|.$$

Setting

$$a_n = \sum_{\substack{k=1,\ldots,n \\ k\notin K}} \|y^k - y^{k-1}\| \left[\frac{\tau}{\theta}(1-\sigma)\|y^k - y^{k-1}\| - \frac{1}{2}\|\delta^k\|\right],$$

$$b_n = \sum_{\substack{k=1,\ldots,n \\ k\in K}} \|y^k - y^{k-1}\| \left[\frac{\tau}{\theta}(1-\sigma)\|y^k - y^{k-1}\| - \frac{1}{2}\|\delta^k\|\right]$$

and

$$c_n = \|y^0 - \overline{y}\|^2 + \alpha\frac{\tau}{\theta}\overline{\lambda}\|\overline{u}\| \sum_{k=1}^{n} \|\delta^k\|,$$

it follows from the $\sum_{k=1}^{n} \|\delta^k\| < \infty$ that $\lim\limits_{n\to\infty} c_n < \infty$. We will show that $\{a_n\}$ is bounded below and $\lim\limits_{n\to\infty} b_n = \infty$, which is a contradiction and this will complete the proof. Since, $\{\|y^k - y^{k-1}\|\}_{k\notin K}$ is bounded, there is L such that $\|y^k - y^{k-1}\| \leq L$ for all $k \notin K$, then

$$-\frac{1}{2}L\|\delta^k\| \leq -\frac{1}{2}\|y^k - y^{k-1}\|\,\|\delta^k\| \leq \frac{\tau}{\theta}(1-\sigma)\|y^k - y^{k-1}\|^2 - \frac{1}{2}\|y^k - y^{k-1}\|\,\|\delta^k\|,$$

summing the inequalities, we have $-\frac{1}{2}L \sum\limits_{\substack{k=1,\ldots,n \\ k\notin K}} \|\delta^k\| \leq a_n$, it follows that the sequence $\{a_n\}$ is bounded below because $\sum_{k=1}^{n} \|\delta^k\| < \infty$.

Since in K the sequence $\{\|y^k - y^{k-1}\|\}$ is unbounded and $\{\|\delta^k\|\}$ converges to zero, there exists an $k_0 \in K$ such that

$$\frac{\tau}{\theta}(1-\sigma)\|y^k - y^{k-1}\| - \frac{1}{2}\|\delta^k\| \geq L > 0,$$

therefore,

$$\infty = L \sum_{\substack{k > k_0 \\ k \in K}}^{\infty} \|y^k - y^{k-1}\| \leq \sum_{\substack{k > k_0 \\ k \in K}}^{\infty} \|y^k - y^{k-1}\| \left[\frac{\tau}{\theta}(1-\sigma)\|y^k - y^{k-1}\| - \frac{1}{2}\|\delta^k\|\right],$$

it follows that $\lim_{n\to\infty} b_n = \infty$. ■

cor:4.1 **Corollary 3.6** *Suppose that (H_3) holds. Then, for $\overline{x}, \overline{u}, \overline{y}$ as in Proposition 3.4, it holds that*

(i) $\{\|y^k - \overline{y}\|\}$ *converges (and hence $\{y^k\}$ is bounded);*

(ii) $\lim_k \|y^k - y^{k-1}\| = 0$*;*

(iii) $\{\|A(x^k - \overline{x})\|\}$ *converges (hence $\{\|x^k - \overline{x}\|\}$ converges and $\{x^k\}$ is bounded);*

(iv) $\lim_k \|\widetilde{x}^k - x^k\| = 0$*;*

(v) $\{\widetilde{x}^k\}$ *is bounded.*

Proof. (i) From (3.13) we have that

$$\|y^k - \overline{y}\|^2 \leq \|y^{k-1} - \overline{y}\|^2 + 2\|\delta^k\|\|y^k - y^{k-1}\| + \alpha\frac{2}{\theta}\lambda_k\|\overline{u}\|\|\delta^k\| \;\forall\; k. \quad (3.20)$$ eq:5.b

Define

$$\sigma_{k+1} := \|y^k - \overline{y}\|^2 \text{ and } \beta_k := 2\|\delta^k\|\|y^k - y^{k-1}\| + \alpha\frac{2}{\theta}\lambda_k\|\overline{u}\|\|\delta^k\|.$$

Since $\{\|y^k - y^{k-1}\|\}$ is bounded by Corollary 3.5 and $\sum_{k=1}^{\infty}\|\delta^k\| < \infty$, then $\sum_{k=1}^{\infty}\|\beta^k\| < \infty$. Therefore the sequences $\{\sigma_k\}$ and $\{\beta_k\}$ are in the conditions of Lemma 2.5. This implies that $\{\|y^k - \overline{y}\|\}$ converges and therefore $\{y^k\}$ is bounded.
(ii) It follows from (i) and Proposition 3.4 that

$$\sum_{k=1}^{\infty}\|y^k - y^{k-1}\|^2 < \infty,$$

therefore $\lim_k \|y^k - y^{k-1}\| = 0$
(iii) Since $y^k - \overline{y} = A(\overline{x} - x^k) + \delta^k$, we get that

$$\|y^k - \overline{y}\| - \|\delta^k\| \leq \|A(\overline{x} - x^k)\| \leq \|y^k - \overline{y}\| + \|\delta^k\|.$$

Being $\{\|y^k - \overline{y}\|\}$ convergent and $\{\|\delta^k\|\}$ convergent to zero, we conclude from the expression above that $\{\|A(\overline{x} - x^k)\|\}$ is also convergent. By (H_2), the function $u \longrightarrow \|u\|_A := \|Au\|$ is a norm in $I\!\!R^n$, and then it follows that $\{\|x^k - \overline{x}\|\}$ converges and therefore $\{x^k\}$ is bounded.
(iv) From (ii) and (3.11), it follows that $\lim_{k\to\infty} \|\widetilde{y}^k - y^k\| = 0$. Therefore $\lim_{k\to\infty} \|A(\widetilde{x}^k - x^k)\| = 0$. Again, the assumptions on A imply that $\lim_{k\to\infty} \|\widetilde{x}^k - x^k\| = 0$.
(v) Follows from (iii) and (iv). ∎

We will show below that the sequence $\{x^k\}$ converges to a solution of $VIP(T, C)$. Denote by $\mathcal{A}cc(z^k)$ the set of accumulation points of the sequence $\{z^k\}$.

teo:6.1 **Theorem 3.7** . *Suppose that (H_1)-(H_3) hold. Then $\{x^k\}$ converges to an element of S.*

Proof. By Corollary 3.6 (iii) and (iv) then $\mathcal{A}cc(\widetilde{x}^k) = \mathcal{A}cc(x^k) \neq \emptyset$. We prove first that every element of $\mathcal{A}cc(\widetilde{x}^k) = \mathcal{A}cc(x^k)$ is a solution of $VIP(T, C)$. Indeed, by (3.16), for all $(x, u) \in G(T)$ it holds

$$\begin{array}{rcl} \langle x - \widetilde{x}^k, u\rangle & \geq & (\lambda_k)^{-1}[\frac{\theta}{2}(\|y^k - y\|^2 - \|y^{k-1} - y\|^2) \\ & & +(1-\sigma)\frac{\tau}{2}\|y^k - y^{k-1}\|^2 - \theta\|\delta^k\|\,\|y^k - y^{k-1}\|]. \end{array} \quad (3.21)$$ eq:5.6

Using Corollary 3.6 (ii) and (iii), we have that $\{x^k\}$ and $\{y^k\}$ are bounded and $\lim_k \|y^k - y^{k-1}\| = 0$. These facts, together with the identity

$$\|y^k - y\|^2 - \|y^{k-1} - y\|^2 = \|y^k - y^{k-1}\|^2 + 2\langle y^k - y^{k-1}, y^{k-1} - y\rangle,$$

yield

$$\lim_{k\to\infty} \|y^k - y\|^2 - \|y^{k-1} - y\|^2 = 0.$$

Let $\{\widetilde{x}^{k_j}\} \subseteq \{\widetilde{x}^k\}$ be a subsequence converging to x^*, we have that

$$\langle x - x^*, u\rangle = \lim_j \langle x - \widetilde{x}^{k_j}, u\rangle \geq \liminf_k \langle x - \widetilde{x}^k, u\rangle. \quad (3.22)$$ jj

Using the above inequality and the fact that $\lambda_k \geq \overline{\lambda} > 0$, we obtain the following expression by taking limits for $k \to \infty$ in (3.21):

$$\langle x - x^*, u \rangle \geq 0 \;\; \forall \, (x, u) \in G(T). \tag{3.23}$$

eq:5.7

By definition, $y^{k_j} = b + \delta^{k_j} - Ax^{k_j}$ with $y^{k_j} > 0$. We know that $\{y^{k_j}\}$ converges to $y^* = b - Ax^*$, with $y^* \geq 0$. Therefore $Ax^* \leq b$. Equivalently, $x^* \in C$. By definition of N_C, we have

$$\langle x - x^*, w \rangle \geq 0 \;\; \forall \, (x, w) \in G(N_C). \tag{3.24}$$

eq:5.7a

Combining (3.23) and (3.24), we conclude that

$$\langle x - x^*, u + w \rangle \geq 0 \;\; \forall \, (x, u + w) \in G(T + N_C).$$

By (H_1) and Proposition 2.1, $T + N_C$ is maximal monotone. Then the above inequality implies that $0 \in (T + N_C)(x^*)$, i.e, $x^* \in S$. Recall that x^* is also an accumulation point of $\{x^k\}$. Using Corollary 3.6(iii), we have that the sequence $\{\|x^* - x^k\|\}$ is convergent. Since it has a subsequence that converges to zero , the whole sequence $\{\|x^* - x^k\|\}$ must converge to zero. This completes the proof. ■

References

AUS95 [1] Auslender, A. and Haddou, M., *An interior proximal method for convex linearly constrained problems and its extension to variational inequalities*, Mathematical Programming, **71** (1995), pp. 77-100.

nder-Teboule-2003 [2] Auslender, A. and Teboulle, M., *Asymptotic cones and functions in Optimization and Variational Inequalities*, Springer Monographs in Mathematics Springer-Verlag, New York, 2003.

AUS99a [3] Auslender, A. and Teboulle M. and Ben-Tiba S., *A logarithmic-quadratic proximal method for variational inequalities*, Computational Optimization and Applications, **12** (1999), pp. 31-40.

AUS99b [4] Auslender, A. and Teboulle M. and Ben-Tiba S., *Interior proximal proximal and multiplier methods based on second order homogeneous kernels*, Mathematics of Operations Research, **24** (1999), pp. 645-668.

Ben-Tal [5] Ben-Tal, A. and Zibulevsky, M., *Penalty-Barrier Methods for convex programming problems*, SIAM Journal on Optimization, **7** (1997), pp. 347-366.

BR-65 [6] Brøndsted, A. and Rockafellar, R.T., *On the subdifferentiability of convex functions*, Proceedings of the American Mathematical Society, **16** (1965), pp. 605-611.

BRO76 [7] Browder, F. E., *Nonlinear operators and nonlinear equations of evolution in Banach spaces*, Proceedings of Symposia in Pure Mathematics, Journal of the Americam Mathematical Society, **18**, 1976.

BUR98 [8] Burachik, R. S. and Iusem, A. N., *A generalized proximal point method for the variational inequality problem in a Hilbert space*, SIAM Journal on Optimization, **8** (1998), pp. 197-216.

BUR97 [9] Burachik, R. S., Iusem, A. N. and Svaiter, B. F., *Enlargements of maximal monotone operators with applications to variational inequalities.* Set Valued Analysis, **5** (1997), pp. 159-180.

bss [10] Burachik, R.S., Sagastizábal, C.A. and Svaiter, B.F., *Epsilon-Enlargement of Maximal Monotone Operators with Application to Variational Inequalities*, in Reformulation - Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, Fukushima, M. and Qi, L. (editors), Kluwer Academic Publishers, 1997.

BUR01i [11] Burachik, R. S. and Svaiter, B. F., *A relative error tolerance for a family of generalized proximal point methods*, Mathematics of Operations Research, **26** (2001), pp. 816-831.

CEN98 [12] Censor, Y., Iusem, A. and Zenios S., *An interior point method with Bregman functions for the variational inequality problem with paramonotone operators*, Mathematical Programming, **83** (1998), pp. 113-123.

eckstein [13] Eckstein, J., *Approximate iterations in Bregman-function-based proximal algorithms*, Mathematical Programming, **81** (1998), pp. 373-400.

eggermont [14] Eggermont, P., *Multiplicatively iterative algorithms for convex programming*, Linear Algebra and its Applications, **130** (1990), pp. 25-42.

HOF52 [15] Hoffman, A. J., *On approximate solutions of systems of linear inequalities*, J. of the National Bureau of Standards, **49** (1952), pp. 263-265.

IUS94 [16] Iusem, A. N., *On some properties of paramonotone operators*, Math. Oper. Res., **19**, (1994), pp. 790-814.

IUS98 [17] Iusem, A. N., Svaiter, B. F. and Teboulle M., *Entropy-like proximal methods in convex programming*, Journal Convex Anal., **5**, (1998), pp. 269-278.

iusem-teboulle [18] Iusem, A.N. and Teboulle, M., *Convergence rate analysis of non-quadratic proximal and augmented Lagrangian methods for convex and linear programming*, Math. of Oper. Res., **20** (1995), pp. 657-677.

kiwiel [19] Kiwiel, K. C., *Proximal minimization methods with generalized Bregman functions*, SIAM Journal on Control and Optimization, **35** (1997), pp. 1142-1168.

LEM95 [20] Lemaire, B., *On the convergence of some iterative methods for convex minimization.* In Recent Developments in Optimization, R. Durier and C. Michelot (eds), Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, **429** (1995), pp. 252-268.

POL87 [21] Polyak, B. T., *Introduction to optimization*, Optimization Software INc., New York, 1987.

PAS78 [22] Pascali, D. and Sburlan, S., *Nonlinear mappings of monotone type*, Ed. Academiei, Bucaresti, Romenia, (1978).

ROC70ii [23] Rockafellar, R. T., *On the maximality of sums of nonlinear monotone operators*, Transactions of the American Mathematical Society , **149** (1970), pp. 75-88.

ss2 [24] Solodov, M.V. and Svaiter, B.F., *An Inexact Hybrid Generalized Proximal Point Algorithm and some New Results on the Theory of Bregman Functions*. Mathematics of Operations Research, **25** (2000), pp. 214-230.

TEB92 [25] Teboulle, M., *Entropic proximal mappings with applications to nonlinear programming*, Mathematics of Operations Research, **1** (1992), pp. 670-690.

TEB97 [26] Teboulle, M., *Convergence of proximal-like algorithms.* SIAM J. Optim., **7** (1997) pp. 1069-1083.

tseng-bertsekas [27] Tseng, P. and Bertsekas D., *On the convergence of exponential multiplier method for convex programming*, Mathematical Programming, **60** (1993), pp. 1-19.

YAM01 [28] Yamashita, N., Kanzon, C., Morimoto, T. and Fukushima, M., *An infeasible interior proximal method for convex programing problems with linear constraints*, J. Nonlinear Convex Analysis , **2** (2001), pp. 139-156.