Journal of Mathematical Analysis and Applications **263,** 343–354 (2001)
doi:10.1006/jmaa.2001.7535, available online at http://www.idealibrary.com on IDEAL®

A New Accuracy Criterion for Approximate Proximal Point Algorithms[1]

Deren Han and Bingsheng He

Department of Mathematics, Nanjing University, Nanjing 210093, People's Republic of China

Submitted by Harold P. Benson

Received July 7, 1999

In this paper, we give a new accuracy criterion for approximate proximal point algorithms. The criterion depends on the current iterate and is easy to verify. Under the suggested enforceable accuracy restriction, the convergence analysis is quite easy to follow.

Key Words: proximal point algorithms; monotone operators; inexact methods.

1. INTRODUCTION

A set $T \subset R^n \times R^n$ with the property

$$(x, y), (x', y') \in T \Rightarrow \langle x - x', y - y' \rangle \geq 0,$$

is called a monotone operator on R^n, where $\langle \cdot, \cdot \rangle$ denotes the inner product on R^n. T is maximal if (considered as a graph) it is not strictly contained in any other monotone operator on R^n. In this paper, we consider the central problem associated with T: Find $z \in R^n$, such that $0 \in T(z)$, i.e., to find one of the roots of T. Here $T(\cdot)$ is defined as $T(x) = \{y \mid (x, y) \in T\}$.

The theory of maximal monotone operators provides a powerful general framework for the study of convex programming and variational inequalities; see [2, 3, 15], for example. A classical method to solve this problem is the proximal point algorithm, which, starting with any vector $x^0 \in R^n$, iteratively updates x^{k+1} conforming to the following recursion

$$x^{k+1} + c_k T(x^{k+1}) \ni x^k, \tag{1}$$

[1] The research was supported by NSFC Grant 19971040.

0022-247X/01 $35.00

where $\{c_k\}_{k=0}^{\infty} \subset [c, \infty)$, $c > 0$, is a sequence of scalars. However, as pointed out in [13], the ideal form of the method is often impractical, since in many cases, solving problem (1) exactly is either impossible or as difficult as solving the original problem $0 \in T(z)$. On the other hand, there seems to be little justification of the effort required to solve the problem accurately when the iterate is far from the solution point. In [20], Rockafellar gave an inexact variant of the method

$$x^{k+1} + c_k T(x^{k+1}) \ni x^k + e^{k+1}, \tag{2}$$

where $\{e^{k+1}\}$ is regarded as an error sequence. This method is called an *inexact proximal point algorithm*. It was shown that if $e^k \to 0$ quickly enough such that

$$\sum_{k=1}^{\infty} \|e^k\| < +\infty,$$

then $x^k \to z \in R^n$ with $0 \in T(z)$.

Because of its relaxed accuracy requirement, the inexact proximal point algorithm is more practical than the exact one. Thus, it has been studied widely and various forms of the method have been developed [3, 8, 10, 18]. In most of these papers, the condition that the error term being summable is an essential condition for the convergence of the method. In [20] and some sequel papers (e.g., [5]), the accuracy criterion is

$$\|e^{k+1}\| \le \eta_k \|x^{k+1} - x^k\| \qquad \text{with } \sum_{k=0}^{\infty} \eta_k < +\infty. \tag{3}$$

Recently, Eckstein [13] extended the method to Bregman-function-based inexact proximal methods and proved that the sequence $\{x^k\}$ generated by the algorithm converges to a root of T under the conditions

$$\sum_{k=1}^{\infty} \|e^k\| < +\infty \qquad \text{and} \qquad \sum_{k=1}^{\infty} \langle e^k, x^k \rangle \text{ exists and is finite} \tag{4}$$

(see Eqs. (18) and (19) in [13]). Condition (4) is an assumption on the whole generated sequence $\{x^k\}$ and the error term sequence $\{e^k\}$, and thus seems to be slightly stronger, but it can be checked and enforced in practice more easily than those that existed earlier. On the other hand, more recently, He [14] gave another inexact criterion in the study of monotone general variational inequalities, which involves a relation between the error term and the residual function.

In this paper, similar to He [14], we give the following accuracy criterion

$$\|e^{k+1}\| \le \eta_k \|x^{k+1} - x^k\| \qquad \text{with } \sum_{k=0}^{\infty} \eta_k^2 < +\infty, \tag{5}$$

to recursion (2) and study the resulting convergence properties. It is clear that the accuracy criterion (5) is weaker than the one in [20] (see (3)).

We note that da Silva e Silva *et al*. [9] and Solodov and Svaiter [21–23] recently proposed some new accuracy criteria for proximal point algorithms. Their criteria, rather than requiring inequality (5), require only that $\sup_{k \geq 0} \eta_k < 1$. Thus, their results are in some sense stronger than ours. However, in [21–23], this comes at the cost of adding an additional projection or "extragradient" step to the algorithm, and the applicable portion of [9] applies only to convex minimization.

Throughout this paper, we assume that the roots set of T, denoted by Z, is nonempty.

2. PRELIMINARIES

In this section, we summarize some basic properties and related definitions of the monotone operator T. As is the custom, we regard T as the graph of a point-to-set mapping. The domain of the mapping T is

$$\operatorname{dom} T = \{x \in R^n |, \exists y \in R^n, (x, y) \in T\} = \{x \in R^n \mid T(x) \neq \emptyset\}.$$

We say T has full domain if $\operatorname{dom} T = R^n$. The range or image of T is

$$\operatorname{im} T = \{y \mid x \in R^n, \{x, y\} \in T\}.$$

For all real numbers c, we let

$$cT = \{(x, cy) \mid (x, y) \in T\},$$

and for all operators $A, B \in R^n \times R^n$, we define $A + B$ via

$$A + B = \{(x, y + z) \mid (x, y) \in A, (x, z) \in B\}.$$

The inverse of T, denoted by T^{-1}, is

$$T^{-1} = \{(y, x) \mid (x, y) \in T\}.$$

T is maximal monotone if and only if T^{-1} is maximal monotone. Given any positive scalar c and operator T, $J_c = (I + cT)^{-1}$ is called a resolvent of T, where I denotes the identity mapping on R^n. T is said to be firmly nonexpansive if

$$\|y' - y\|^2 \leq \langle x' - x, y' - y \rangle, \qquad \forall (x, y), (x', y') \in T.$$

In the rest of this section, we quote some preliminaries for sequences $\{x^k\}$ and $\{e^k\}$ conforming to recursion (2). First, it is important to ask if the sequences $\{x^k\}_{k=0}^{\infty}$, $\{e^k\}_{k=1}^{\infty}$ exist. Lemma 1 gives a positive answer to this question.

LEMMA 1. *Let c be any positive scalar. An operator T on R^n is monotone if and only if its resolvent $J_c = (I + cT)^{-1}$ is firmly nonexpansive. Furthermore, T is maximal monotone if and only if J_c is firmly nonexpansive and* $\operatorname{dom} J_c = R^n$.

Proof. See [12, Theorem 2], for example. ∎

From [2], we know that if T is a maximal monotone operator, then it is a closed set in $R^n \times R^n$. Hence, for the problem under consideration, the sequences $\{x^k\}_{k=0}^{\infty}$, $\{e^k\}_{k=1}^{\infty}$ conforming to recursion (2) exist. In Lemma 2 we will list a few inequalities associated with recursion (2). The results are special cases of known results (Lemma 2 of [13]). For completeness, we have included the proofs, which are short.

LEMMA 2. *Let $\{x^k\}$ and $\{e^k\}$ be sequences that conform to recursion* (2). *Then for any $x^* \in Z$ (root of T) and all $k \geq 0$ we have*

$$\langle x^k - x^{k+1} + e^{k+1}, x^{k+1} - x^* \rangle \geq 0 \tag{6}$$

and

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - \|x^{k+1} - x^k\|^2 + 2\langle e^{k+1}, x^{k+1} - x^* \rangle. \tag{7}$$

Proof. The proof is similar to that in Eckstein [13]. It follows from (2) that

$$\frac{1}{c_k}(x^k - x^{k+1} + e^{k+1}) \in T(x^{k+1}).$$

Since x^* is a root of T, $0 \in T(x^*)$, and T is monotone, we have

$$\left\langle \frac{1}{c_k}(x^k - x^{k+1} + e^{k+1}) - 0, x^{k+1} - x^* \right\rangle \geq 0.$$

The first assertion is obtained from the assumption that c_k is a positive scalar. Furthermore, using $\|u + v\|^2 = \|u\|^2 - \|v\|^2 + 2\langle v, u + v \rangle$ and (6) we get

$$\begin{aligned}
\|x^{k+1} - x^*\|^2 &= \|x^k - x^*\|^2 - \|x^{k+1} - x^k\|^2 + 2\langle x^{k+1} - x^k, x^{k+1} - x^* \rangle \\
&= \|x^k - x^*\|^2 - \|x^{k+1} - x^k\|^2 + 2\langle e^{k+1}, x^{k+1} - x^* \rangle \\
&\quad - 2\langle x^k - x^{k+1} + e^{k+1}, x^{k+1} - x^* \rangle \\
(\text{use } (6)) &\leq \|x^k - x^*\|^2 - \|x^{k+1} - x^k\|^2 + 2\langle e^{k+1}, x^{k+1} - x^* \rangle.
\end{aligned}$$

This completes the proof. ∎

3. MAIN RESULTS

Now, we begin to investigate the convergence properties of recursion (2) under the accuracy criterion

$$\|e^{k+1}\| \leq \eta_k \|x^{k+1} - x^k\| \qquad \text{with } \sum_{k=0}^{\infty} \eta_k^2 < +\infty.$$

Note that in the exact proximal point algorithm (1), x^k is a root of T if and only if $x^{k+1} = x^k$. Hence, roughly speaking, we can see the distance $\|x^{k+1} - x^k\|$ as an "error bound," which measures how much x^k fails to be in the roots set of T. If $\|x^{k+1} - x^k\|$ is small enough, it follows from Eq. (1) that x^{k+1} is an acceptable approximate solution of the original problem. Hence, it is reasonable to give an accuracy criterion as in (5) that depends on the distance $\|x^{k+1} - x^k\|$. In the following we will prove that the sequence $\{x^k\}$ is *weakly contractive* and the *error bound* will converge to zero.

THEOREM 1. *Let $\{x^k\}$ and $\{e^k\}$ be sequences generated by the inexact proximal point algorithm* (2) *under the proposed accuracy criterion* (5). *Then there exists an integer $k_0 \geq 0$, such that for all $k \geq k_0$*

$$\|x^{k+1} - x^*\|^2 \leq \left(1 + \frac{2\eta_k^2}{1 - 2\eta_k^2}\right)\|x^k - x^*\|^2 - \frac{1}{2}\|x^{k+1} - x^k\|^2. \tag{8}$$

Furthermore, $\{x^k\}$ is a bounded sequence and

$$\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0. \tag{9}$$

Proof. Let x^* be any root of T. For $\eta_k > 0$, using the Cauchy–Schwarz inequality we have

$$2\langle e^{k+1}, x^{k+1} - x^* \rangle \leq \frac{1}{2\eta_k^2}\|e^{k+1}\|^2 + 2\eta_k^2\|x^{k+1} - x^*\|^2. \tag{10}$$

Since $\eta_k \to 0$, there exists $k_0 \geq 0$, such that for all $k \geq k_0$, $1 - 2\eta_k^2 > 0$. Substituting (10) in (7) we obtain

$$\begin{aligned}\|x^{k+1} - x^*\|^2 &\leq \left(1 + \frac{2\eta_k^2}{1 - 2\eta_k^2}\right)\|x^k - x^*\|^2 - \frac{1}{2(1 - 2\eta_k^2)}\|x^{k+1} - x^k\|^2 \\ &\leq \left(1 + \frac{2\eta_k^2}{1 - 2\eta_k^2}\right)\|x^k - x^*\|^2 - \frac{1}{2}\|x^{k+1} - x^k\|^2.\end{aligned}$$

The first part of the theorem is obtained and thus

$$\|x^{k+1} - x^*\|^2 \le \left(1 + \frac{2\eta_k^2}{1 - 2\eta_k^2}\right)\|x^k - x^*\|^2, \qquad \forall k \ge k_0. \tag{11}$$

Since $\sum_{k=0}^{\infty} \eta_k^2 < +\infty$, it follows that

$$C_S := \sum_{k=k_0}^{\infty} \frac{2\eta_k^2}{1 - 2\eta_k^2} < +\infty \qquad \text{and}$$

$$C_P := \prod_{k=k_0}^{\infty} \left(1 + \frac{2\eta_k^2}{1 - 2\eta_k^2}\right) < +\infty,$$

and thus $\{x^k\}$ is bounded. Also from (8) we have

$$\begin{aligned}
&\frac{1}{2} \sum_{k=k_0}^{\infty} \|x^{k+1} - x^k\|^2 \\
&\quad \le \sum_{k=k_0}^{\infty} \left(\|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2\right) + \sum_{k=k_0}^{\infty} \frac{2\eta_k^2}{1 - 2\eta_k^2}\|x^k - x^*\|^2 \\
&\quad \le \|x^{k_0} - x^*\|^2 + \sum_{k=k_0}^{\infty} \frac{2\eta_k^2}{1 - 2\eta_k^2}\left(\sup_{k_0 \le k < \infty}\|x^k - x^*\|^2\right) \\
&\quad \le \left(1 + C_S C_p\right)\|x^{k_0} - x^*\|^2 \\
&\quad < +\infty.
\end{aligned}$$

It follows that

$$\lim_{k \to +\infty} \|x^{k+1} - x^k\| = 0$$

and the proof is complete. ∎

We can obtain the convergence of $\{x^k\}$ from the weak contraction (8).

THEOREM 2. *Let T be a maximal monotone operator on R^n, and $\{x^k\}$ and $\{e^k\}$ be sequences generated by the inexact proximal point algorithm* (2) *under the proposed accuracy criterion* (5). *Then $\{x^k\}$ converges to some x^∞ with $0 \in T(x^\infty)$.*

Proof. From Theorem 1, $\{x^k\}$ is bounded, so that it has at least a cluster point. Let x^∞ be a cluster point of $\{x^k\}$ and the subsequence $\{x^{k_j}\}$ converges to x^∞. Define

$$y^{k+1} = \frac{1}{c_k}\left(x^k - x^{k+1} - e^{k+1}\right).$$

Then $y^{k_j+1} \in T(x^{k_j+1})$. Using $x^{k_j} \to x^\infty$, $\lim_{k\to\infty}\|x^{k+1} - x^k\| = 0$, and $e^k \to 0$, we have

$$\lim_{j\to\infty} y^{k_j+1} = \lim_{j\to\infty} \frac{1}{c_{k_j}}\left(x^{k_j} - x^{k_j+1} + e^{k_j+1}\right) = 0.$$

Because T is maximal, it is a closed set in $R^n \times R^n$. Therefore

$$\lim_{j\to\infty}\left(x^{k_j}, y^{k_j}\right) = (x^\infty, 0) \in T,$$

and x^∞ is a root of T. Note that the inequality (8) (in Theorem 1) is true for all roots of T. Hence we have

$$\|x^{k+1} - x^\infty\|^2 \leq \left(1 + \frac{2\eta_k^2}{1 - 2\eta_k^2}\right)\|x^k - x^\infty\|^2, \qquad \forall k \geq k_0. \tag{12}$$

Since $\{x^{k_j}\} \to x^\infty$ and $\prod_{k=k_0}^{\infty}(1 + 2\eta_k^2/(1 - 2\eta_k^2)) < +\infty$, for any given $\varepsilon > 0$, there is an $l > 0$, such that

$$\|x^{k_l} - x^\infty\| < \frac{\varepsilon}{2} \qquad \text{and} \qquad \sqrt{\prod_{k=k_l}^{\infty}\left(1 + \frac{2\eta_k^2}{1 - 2\eta_k^2}\right)} < 2. \tag{13}$$

Therefore, for any $k \geq k_l$, it follows from (12) and (13) that

$$\|x^k - x^\infty\| \leq \sqrt{\prod_{t=k_l}^{k-1}\left(1 + \frac{2\eta_t^2}{1 - 2\eta_t^2}\right)}\left(\|x^{k_l} - x^\infty\|\right) < \varepsilon$$

and the sequence $\{x^k\}$ converges to x^∞. ∎

4. EXTENSION TO BREGMAN-FUNCTION-BASED PROXIMAL ALGORITHMS

Much recent research has focused on *nonlinear* generalizations of recursion (1) based on Bregman functions [13]. Suppose h is a strictly convex function, continuously differentiable on some open set S. The Bregman distance between x and y is defined via the "D-function"

$$D_h(x, y) = h(x) - h(y) - \langle \nabla h(y), x - y\rangle, \tag{14}$$

where $x \in \bar{S}$ and $y \in S$. From the strict convexity of h, one can prove that $D_h(x, y) \geq 0$, and $D_h(x, y) = 0$ if and only if $x = y$. If $h(x) = \frac{1}{2}\|x\|^2$, then $D_h(x, y) = \frac{1}{2}\|x - y\|^2$. In the following, we will use a class of functions that are presented as

$$h(x) = h_0(x) + \frac{\mu}{2}\|x\|^2,$$

where h_0 is a Bregman function and $\mu > 0$. It is easy to see that h satisfies the conditions of the definition of a Bregman function (see [13, 17], for example), so h is also a Bregman function. Thus, for all $x \in \bar{S}$, $y \in S$, we have

$$\begin{aligned} D_h(x, y) &= h(x) - h(y) - \langle \nabla h(y), x - y \rangle \\ &= h_0(x) - h_0(y) - \langle \nabla h_0(y), x - y \rangle + \frac{\mu}{2}\|x\|^2 \\ &\quad - \frac{\mu}{2}\|y\|^2 - \langle \mu y, x - y \rangle \\ &= D_{h_0}(x, y) + \frac{\mu}{2}\|x - y\|^2 \\ &\geq \frac{\mu}{2}\|x - y\|^2. \end{aligned}$$

It follows therefore, for any fixed $x^* \in Z$, that there is a constant $C > 0$ such that

$$\|x - x^*\|^2 \leq C \cdot D_h(x^*, x), \qquad \forall x \in \{x^k\}. \tag{15}$$

An alternative to Eq. (1) is the recursion

$$\nabla h(x^{k+1}) + c_k T(x^{k+1}) \ni \nabla h(x^k); \tag{1'}$$

see [4, 7, 8], for example. Because of the practical difficulty in computing the exact solutions of Eq. (1′), Eckstein suggested a natural generalization by taking a simpler and more practically verifiable approach than [3, 4, 16, 24], i.e.,

$$\nabla h(x^{k+1}) + c_k T(x^{k+1}) \ni \nabla h(x^k) + e^{k+1}. \tag{2'}$$

For this problem, instead of condition (5), we can take

$$\|e^{k+1}\| \leq \eta_k D_h^{1/2}(x^{k+1}, x^k) \qquad \text{with } \sum_{k=0}^{\infty} \eta_k^2 < +\infty \tag{5'}$$

as the approximate criterion corresponding to recursion (2′). Eckstein proved (see Lemma 2 in [13]) that the sequences $\{x^k\}$ and $\{e^k\}$ conforming to (2′) satisfy

$$D_h(x^*, x^{k+1}) \leq D_h(x^*, x^k) - D_h(x^{k+1}, x^k) + \langle e^{k+1}, x^{k+1} - x^* \rangle. \tag{7'}$$

Note that this is a similar result to (7) in the last section. Using the Cauchy–Schwarz inequality, (5′) and (15) we get

$$\begin{aligned}\langle e^{k+1}, x^{k+1} - x^* \rangle &\le \frac{1}{2\eta_k^2}\|e^{k+1}\|^2 + \eta_k^2\|x^{k+1} - x^*\|^2 \\ &\le \frac{1}{2}D_h(x^{k+1}, x^k) + C\eta_k^2 D_h(x^*, x^{k+1}). \qquad (10')\end{aligned}$$

Since $\eta_k \to 0$, there exists $k_0 \ge 0$, such that for all $k \ge k_0$, $1 - C\eta_k^2 > 0$. Based on (7′) and (10′), using the same technique as in last section, we can prove the following theorem:

THEOREM 3. *Let $\{x^k\}$ and $\{e^k\}$ be sequences generated by the generalized proximal point algorithm* (2′) *under the proposed accuracy criterion* (5′). *Then there exists an integer $k_0 \ge 0$, such that for all $k \ge k_0$*

$$D_h(x^*, x^{k+1}) \le \left(1 + \frac{C\eta_k^2}{1 - C\eta_k^2}\right) D_h(x^*, x^k) - \frac{1}{2}D_h(x^{k+1}, x^k), \quad (8')$$

and

$$\lim_{k\to\infty} D_h(x^{k+1}, x^k) = 0. \qquad (9')$$

Proof. It follows from (7′) and (10′) that

$$D_h(x^*, x^{k+1}) \le D_h(x^*, x^k) - \frac{1}{2}D_h(x^{k+1}, x^k) + C\eta_k^2 D_h(x^*, x^{k+1}).$$

Since $\eta_k \to 0$, there exists $k_0 \ge 0$, such that for all $k \ge k_0$, $1 - C\eta_k^2 > 0$. Therefore,

$$\begin{aligned}&D_h(x^*, x^{k+1}) \\ &\quad\le \left(1 + \frac{C\eta_k^2}{1 - C\eta_k^2}\right) D_h(x^*, x^k) - \frac{1}{2(1 - C\eta_k^2)} D_h(x^{k+1}, x^k) \\ &\quad\le \left(1 + \frac{C\eta_k^2}{1 - C\eta_k^2}\right) D_h(x^*, x^k) - \frac{1}{2}D_h(x^{k+1}, x^k).\end{aligned}$$

Inequality (8′) follows immediately and thus

$$D_h(x^*, x^{k+1}) \le \left(1 + \frac{C\eta_k^2}{1 - C\eta_k^2}\right) D_h(x^*, x^k), \qquad \forall k \ge k_0. \quad (11')$$

Since $\sum_{k=0}^{\infty} \eta_k^2 < +\infty$, it follows that

$$C'_S := \sum_{k=k_0}^{\infty} \frac{C\eta_k^2}{1 - C\eta_k^2} < +\infty \qquad \text{and}$$

$$C'_P := \prod_{k=k_0}^{\infty} \left(1 + \frac{C\eta_k^2}{1 - C\eta_k^2}\right) < +\infty.$$

From (11′), we have

$$D_h(x^*, x^{k+1}) \le C'_P D_h(x^*, x^{k_0}) < +\infty,$$

and thus $\{x^k\}$ is bounded. Also from (8′) we have

$$\begin{aligned}
&\frac{1}{2} \sum_{k=k_0}^{\infty} D_h(x^{k+1}, x^k) \\
&\quad \le \sum_{k=k_0}^{\infty} \left(D_h(x^*, x^k) - D_h(x^*, x^{k+1})\right) + \sum_{k=k_0}^{\infty} \frac{C\eta_k^2}{1 - C\eta_k^2} D_h(x^*, x^k) \\
&\quad \le D_h(x^*, x^{k_0}) + \sum_{k=k_0}^{\infty} \frac{C\eta_k^2}{1 - C\eta_k^2} \left(\sup_{k_0 \le k < \infty} D_h(x^*, x^k)\right) \\
&\quad \le \left(1 + C'_S C'_p\right) D_h(x^*, x^{k_0}) \\
&\quad < +\infty.
\end{aligned}$$

It follows that

$$\lim_{k \to +\infty} D_h(x^{k+1}, x^k) = 0,$$

and the proof is complete. ∎

Furthermore, since h is strictly convex, by using the same technique as in the last section, we can prove that $\{x^k\}$ converges to x^∞, a root of T.

5. CONCLUDING REMARKS

In this paper, we suggested a new accuracy criterion for approximate proximal point algorithms. The accuracy condition is easy to verify and to extend to Bregman-function-based proximal point algorithms. However, we would like to point out that the convergence analysis is based on the assumption that the roots set of T is nonempty. Note that T may have no root even if T is maximal monotone, that is, the roots set Z may be empty. If Z is empty, the sequence $\{x^k\}$ conforming to (2) (resp. (2′)) is unbounded; see [1, 6, 11, 20], for example.

ACKNOWLEDGMENTS

The authors thank the two anonymous referees and the associate editor for their constructive suggestions which led to the improvement of the paper.

REFERENCES

1. S. Aizicovici and Y. Q. Chen, Note on the topological degree of the subdifferential of a lower semi-continuous convex function, *Proc. Amer. Math. Soc.* **126** (1998), 2905–2908.
2. H. Brézis, "Opérateurs Maximaux Monotone et Semi-Groups de Contractions dans les Espaces de Hilbert," North-Holland, Amsterdam, 1973.
3. R. S. Burachik, A. N. Iusem, and B. F. Svaiter, Enlargement of monotone operators with applications to variational inequalities, *Set-Valued Anal.* **5** (1997), 159–180.
4. Y. Censor and S. A. Zenios, The proximal minimization algorithm with D-functions, *J. Optim. Theory Appl.* **73** (1992), 451–464.
5. G. Chen and M. Teboulle, A proximal-based decomposition method for convex minimization problems, *Math. Programming* **64** (1994), 81–101.
6. Y. Q. Chen, A generalization of the zero point theorem of maximal monotone mappings, *Acta Math. Sinica* **38** (1995), 831–836.
7. G. Cohen, Auxiliary problem principle and decomposition of optimization problems, *J. Optim. Theory Appl.* **32** (1980), 277–305.
8. G. Cohen, Auxiliary problem principle extended to variational inequalities, *J. Optim. Theory Appl.* **59** (1988), 325–333.
9. P. J. da Silva e Silva, J. Eckstein, and C. Humes, "Rescaling and Stepsize Selection in Proximal Methods Using Separable Generalized Distance," RUTCOR Research Report RRR 35-99, Rutgers University, 1999.
10. R. S. Dembo, S. C. Eisenstat, and T. Steihaug, Inexact Newton methods, *SIAM J. Numer. Anal.* **19** (1982), 400–408.
11. J. Eckstein, Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming, *Math. Oper. Res.* **18** (1993), 202–226.
12. J. Eckstein and D. P. Bertsekas, On the Douglas–Rachford splitting method and the proximal points algorithm for maximal monotone operators, *Math. Programming* **55** (1992), 293–318.
13. J. Eckstein, Approximate iterations in Bregman-function-based proximal algorithms, *Math. Programming* **83** (1998), 113–123.
14. B. S. He, Inexact implicit methods for monotone general variational inequalities, *Math. Programming* **86** (1999), 199–217.
15. M. R. Hestenes, Multiplier and gradient methods, *J. Optim. Theory Appl.* **4** (1969), 303–320.
16. K. C. Kiwiel, Proximal minimization methods with generalized Bregman functions, *SIAM J. Control Optim.* **35** (1997), 1142–1168.
17. M. Kyono and M. Fukushima, Nonlinear proximal decomposition method for convex programming, *J. Optim. Theory Appl.* **106** (2000), 357–372.

18. J. S. Pang, Inexact Newton methods for the nonlinear complementarity problem, *Math. Programming* **36** (1986), 54–71.
19. B. Polyak, "Introduction to Optimization," Optimization Software Inc., New York, 1987.
20. R. T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM J. Control Optim*. **14** (1976), 877–898.
21. M. V. Solodov and B. F. Svaiter, A hybrid projection-proximal point algorithm, *J. Convex Anal*. **6** (1999), 59–70.
22. M. V. Solodov and B. F. Svaiter, A hybrid approximate extragradient-proximal point algorithm using the enlargement of a maximal monotone operator, *Set-Valued Anal*. **7** (1999), 323–345.
23. M. V. Solodov and B. F. Svaiter, An inexact hybrid generalized proximal point algorithm and some new result on the theory of Bregman functions, *Math. Oper. Res*. **25** (2000), 214–230.
24. M. Teboulle, Convergence of proximal-like algorithms, *SIAM J. Optim*. **7** (1997), 1069–1083.