JOURNAL OF DECISION MAKING AND HEALTHCARE
Volume 2 (2025), No. 2, 105–113
https://doi.org/10.69829/jdmh-025-0202-ta02
Tulipa Opera Scholarum

RELAXED PROXIMAL POINT ALGORITHM FOR MONOTONE INCLUSION PROBLEM BEYOND MONOTONICITY

GRACE NNENNAYA OGWO[1] AND YEKINI SHEHU[1,*]

[1] *School of Mathematical Sciences, Zhejiang Normal University, Jinhua 321004, People's Republic of China*

Dedicated to Professor Hari Mohan Srivastava on the Occasion of His 85th Birthday

ABSTRACT. In this paper, we introduce a relaxed proximal point algorithm for solving the generalized inclusion problem involving a maximally comonotone operator in the framework of real Hilbert spaces. While the classical theory for the proximal point algorithm relies heavily on monotonicity assumptions, we demonstrate how comonotonicity (a weaker yet structurally rich property) provides sufficient conditions for convergence. By exploiting the fundamental connection between comonotonicity and averaged resolvents, we establish weak convergence of our proposed method under standard assumptions.

Keywords. Proximal operator, Maximally monotone operator, Nonexpansive mapping, Weak convergence.

1. INTRODUCTION

Let $\mathscr{H}$ denote the Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\|\cdot\|$. An important problem in applied mathematics and optimization is the inclusion problem defined as: Find $x \in \mathscr{H}$ such that

$$0 \in Bx, \tag{1.1}$$

where $B : \mathscr{H} \rightrightarrows \mathscr{H}$ is a maximal monotone operator. We denote the solution set of (1.1) by $\text{Zer}(B) := B^{-1}(0)$ and assume it to be nonempty. The inclusion problem covers several important applications in scientific fields such as image processing, signal processing, economics, game theory, computer vision, statistics (see [2, 4, 5, 9, 19, 20, 23, 26, 29, 30] and other references therein).

An important algorithm for approximating (1.1) is the proximal point algorithm (PPA) first suggested by Martinet [24] for solving variational problems and further generalized by Rockafellar [26]. The PPA is defined as:

$$x_{n+1} := J_\lambda^B x_n, \tag{1.2}$$

where J_λ^B is the resolvent of a maximal monotone operator B and $\lambda > 0$. Rockafellar [26] proved that the sequence $\{x_n\}$ generated by (1.2) converges weakly to a zero of B when $\text{Zer}(B) \neq \emptyset$. Recall that the resolvent operator of a set-valued operator is defined by

$$J_\lambda^B = (I + \lambda B)^{-1}.$$

Note that the resolvent J_λ^B of a monotone operator is always single-valued. An important relation between B and J_λ^B is that $B^{-1}(0)$ coincides with the fixed point set of J_λ^B which is not dependent on the choice of $\lambda > 0$. If B is monotone, then J_λ^B is firmly nonexpansive. The PPA plays a significant theoretical and algorithmic role in several areas of scientific computing, such as optimization, image processing, among others (see [2, 4, 9, 30] and other references therein). Many well-known algorithms

*Corresponding author.
E-mail address: graceogwo@zjnu.edu.cn (G. N. Ogwo), yekini.shehu@zjnu.edu.cn (Y. Shehu)
Accepted: June 04, 2025.

can be considered as PPA for solving maximally monotone operators. These include the augmented Lagrangian method, the proximal method of multipliers [27], the Douglas-Rachford splitting method [14, 22], the alternating direction method of multipliers [16, 17], and the primal-dual hybrid gradient method [7, 8].

Although PPA's convergence is well-established for monotone operators, many applications involve nonmonotone mappings. To solve functions that are not necessarily convex, one needs to weaken the monotonicity of the operator B. To address this, Combettes and Pennanen [10], Iusem et al. [18] and Pennanen [25] relaxed the monotonicity assumption to hypomonotonicity. However, hypomonotonicity may still be restrictive for certain problems. To relax the hypomonotonicity condition, Bauschke et al. [1] studied the ρ-comonotonicity assumption on B and related it to the resolvent J_λ^B properties. One of the major results is that the resolvent is an averaged mapping whenever the operator is ρ-comonotone with $\rho > -\frac{1}{2}$. Their method heavily relies on the conically nonexpansive operators and the notions of ρ-monotonicity (ρ-comonotonicity).

Motivated by the works mentioned above, we propose and study a proximal point algorithm for finding the zero of a maximally ρ-comonotone operator in the framework of the real Hilbert space. Although the proximal point algorithm has been studied for ρ-comonotone operators, our work leverages the critical connection between the ρ-comonotonicity of an operator B and the averagedness of its resolvent J_λ^B as established in [1]. This connection allows us to systematically adapt classical proximal point algorithm convergence proofs (originally developed for monotone operators) to the ρ-comonotone case. Moreover, while Bauschke et al. [1] characterized resolvent properties, we exploit these to propose a novel proximal point algorithm variant with relaxed parameters, achieving weak convergence under broader conditions.

This paper is organized as follows. In Section 2, we present some basic definitions, concepts, lemmas, and results required in the convergence analysis of our proposed algorithm. In Section 3, we present our algorithm and the convergence result. In Section 4, we present a summary of our results.

2. Preliminaries

Throughout this paper, we denote the real Hilbert space by $\mathscr{H}$ and the identity operator by I. We denote the weak convergence of a sequence $\{x_n\}$ to a point x by $x_n \rightharpoonup x$. Let $B : \mathscr{H} \rightrightarrows \mathscr{H}$ be a set-valued operator. The domain, range and graph of B is defined as

$$D(B) := \{x \in \mathscr{H} : B(x) \neq \emptyset\},$$

$$\text{ran}(B) := \bigcup\{B(z) : z \in D(B)\},$$

and

$$\text{gra}(B) := \{(x, y) \in \mathscr{H} \times \mathscr{H} : x \in D(B), y \in B(x)\},$$

respectively. The inverse of B, denoted by B^{-1} is defined such that

$$x \in B^{-1}(y) \iff y \in B(x).$$

An operator is called monotone if for all $x, y \in D(B)$,

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

A monotone operator B is called maximally monotone if it has no proper monotone extension, or equivalently, if

$$\text{ran}(I + \lambda B) = \mathscr{H}, \forall \lambda > 0.$$

Definition 2.1. Let $B : \mathscr{H} \rightrightarrows \mathscr{H}$ and let $\rho \in \mathbb{R}$.

(i) B is ρ-monotone if $(\forall\, (x, u) \in \text{gra}(B))$ $(\forall\, (y, v) \in \text{gra}(B))$, we have

$$\langle x - y, u - v\rangle \geq \rho\|x - y\|^2.$$

(ii) B is maximally ρ-monotone if B is ρ-monotone and there is no ρ-monotone operator A such that gra(A) properly contains gra(B), i.e, for every $(x,u) \in \mathscr{H} \times \mathscr{H}$,

$$(x,u) \in \text{gra}(B) \iff (\forall\,(y,v) \in \text{gra}(B)), \langle x-y, u-v\rangle \geq \rho\|x-y\|^2.$$

(iii) B is ρ-comonotone if $(\forall\,(x,u) \in \text{gra}(B))$ $(\forall\,(y,v) \in \text{gra}(B))$, we have

$$\langle x-y, u-v\rangle \geq \rho\|u-v\|^2.$$

(iv) B is maximally ρ-comonotone if B is ρ-comonotone and there is no ρ-comonotone operator A such that gra(A) properly contains gra(B), i.e for every $(x,u) \in \mathscr{H} \times \mathscr{H}$,

$$(x,u) \in \text{gra}(B) \iff (\forall\,(y,v) \in \text{gra}(B)), \langle x-y, u-v\rangle \geq \rho\|u-v\|^2.$$

Remark 2.2.

(i) If $\rho = 0$, then ρ-monotonicity and ρ-comonotonicity of B reduce to the monotonicity of B.

(ii) When $\rho < 0$, ρ-monotonicity is known as $|\rho|$-hypomonotonicity which has been studied in [6, 28]. In this case, the ρ-monotonicity is also known as $|\rho|$-cohypomonotonicity (see [10]).

(iii) When $\rho > 0$, ρ-monotonicity of B reduces to ρ-strong monotonicity of B, while ρ -comonotonicity of B reduces to ρ-cocoercivity of B.

Unlike classical monotonicity, ρ-comonotonicity of B is not equivalent to ρ- comonotonicity of B^{-1}.

Lemma 2.3. [1] *Let $B : \mathscr{H} \rightrightarrows \mathscr{H}$ and let $\rho \in \mathbb{R}$. Then the following are equivalent:*

(i) *B is ρ-comonotone.*

(ii) *$B^{-1} - \rho$ Id is monotone.*

(iii) *B^{-1} is ρ-monotone.*

Example 2.4. Let $\mathscr{H} = \mathbb{R}$, and consider the operator $B : \mathbb{R} \rightrightarrows \mathbb{R}$ defined by $B(x) = \beta x$, where $\beta > 0$. We verify the equivalence for $\rho \in \mathbb{R}$:

(i) B is ρ-comonotone: For $u, u' \in \mathbb{R}$, let $v = \beta u$ and $v' = \beta u'$. Then:

$$\langle v-v', u-u'\rangle = \beta(u-u')^2 \quad \text{and} \quad \rho\|v-v'\|^2 = \rho\beta^2(u-u')^2.$$

The ρ-comonotonicity inequality $\langle v-v', u-u'\rangle \geq \rho\|v-v'\|^2$ becomes

$$\beta(u-u')^2 \geq \rho\beta^2(u-u')^2 \implies \rho \leq \frac{1}{\beta}.$$

(ii) $B^{-1} - \rho$Id is monotone: The inverse operator is $B^{-1}(y) = \frac{y}{\beta}$. Then $B^{-1} - \rho$Id maps y to

$$\frac{y}{\beta} - \rho y = y\left(\frac{1}{\beta} - \rho\right).$$

This is monotone if and only if $\frac{1}{\beta} - \rho \geq 0$, i.e., $\rho \leq \frac{1}{\beta}$.

(iii) B^{-1} is ρ-monotone: For $y, y' \in \mathbb{R}$, $B^{-1}(y) = \frac{y}{\beta}$. The ρ-monotonicity condition requires:

$$\left\langle \frac{y}{\beta} - \frac{y'}{\beta}, y-y' \right\rangle \geq \rho\|y-y'\|^2.$$

Simplifying, we get:

$$\frac{1}{\beta}\|y-y'\|^2 \geq \rho\|y-y'\|^2 \implies \rho \leq \frac{1}{\beta}.$$

All three are equivalent if $\rho \leq \frac{1}{\beta}$. For instance, if $\beta = 2$, then $\rho = \frac{1}{2}$ satisfies all conditions.

Lemma 2.5. [1] *Let $B : \mathscr{H} \rightrightarrows \mathscr{H}$ and let $\rho \in \mathbb{R}$. Then the following are equivalent:*

(i) *B is maximally ρ-comonotone.*

(ii) *$B^{-1} - \rho$ Id is maximally monotone.*

(iii) B^{-1} *is maximally* ρ*-monotone.*

Remark 2.6. It is well known that when $\rho < 0$, the (maximal) monotonicity of $B^{-1} - \rho\,\mathrm{Id}$ is equivalent to the (maximal) monotonicity of the Yosida approximation $(B^{-1} - \rho Id)^{-1}$. See [6, Proposition 6.9.3]

For a ρ-comonotone operator, the resolvent plays an important role in approximating zero points.

Lemma 2.7. [1] *Let* $B : \mathscr{H} \rightrightarrows \mathscr{H}$ *be maximally* ρ*-comonotone with* $\rho \in \mathbb{R}$ *and let* $\lambda > 0$*. If* $\rho > -\lambda$*, then* λB *is maximally* $\frac{\rho}{\lambda}$*−comonotone with* $\frac{\rho}{\lambda} > -1$*, also* J_λ^B *is single-valued and* $D(J_\lambda^B) = ran(I + \lambda B) = \mathscr{H}$*.*

Proposition 2.8. *For* ρ*-comonotone* B *with* $\rho > -\lambda$ *:*

(i) J_λ^B *is* $(\lambda/(2(\rho+\lambda)))$*-conically nonexpansive.*
(ii) $I - J_\lambda^B$ *is* $\frac{(\rho+\lambda)}{(2\lambda)}$*-cocoercive.*

Let $T : \mathscr{H} \to \mathscr{H}$ be an operator. A point $x \in \mathscr{H}$ is called the fixed point of T if $Tx = x$. We denote the set of fixed points of T by $F(T)$.

Definition 2.9. The operator $T : \mathscr{H} \to \mathscr{H}$ is called L-Lipschitz continuous if $L > 0$ and

$$\|Tx - Ty\| \leq L\|x - y\|, \ \forall\, x, y \in \mathscr{H}.$$

If $L = 1$ T is called a nonexpansive mapping.

Definition 2.10. A mapping $T : \mathscr{H} \to \mathscr{H}$ is said to be demi-closed at 0 if, for any sequence $\{x_n\}$ in $\mathscr{H}$, the conditions $x_n \rightharpoonup z$ and $T(x_n) \to 0$, imply $Tz = 0$.

Definition 2.11. [1] Let $\mathscr{C}$ be a nonempty subset of $\mathscr{H}$, let $T : \mathscr{C} \to \mathscr{H}$, and let $\alpha \in (0, \infty)$. Then, T is α-conically nonexpansive on $\mathscr{C}$ if there exists a nonexpansive operator $G : \mathscr{C} \to \mathscr{H}$ such that $T = (1-\alpha)I + \alpha G$. An α-conically nonexpansive operator is α-averaged when $\alpha \in (0, 1)$ and nonexpansive when $\alpha = 1$.

Lemma 2.12. *Let* $\alpha \in (0, \infty)$ *and let* $T : \mathscr{H} \to \mathscr{H}$ *be* α*-conically nonexpansive. Then the following holds:*

(i) *The operator* $I - T$ *is demiclosed at 0.*
(ii) *The set of fixed points,* $F(T)$*, is closed and convex.*
(iii) *The operator* T *is Lipschitz continuous.*
(iv) *For any* $\bar{x} \in F(T)$ *and* $x \in \mathscr{H}$*, the following holds*

$$\langle x - \bar{x}, x - Tx \rangle \geq \frac{1}{2\alpha}\|x - Tx\|^2$$

Lemma 2.13. [1] *Let* $T : \mathscr{H} \to \mathscr{H}$ *be an operator.*

(i) T *is nonexpansive if and only if it is the resolvent of a maximally* $(-\frac{1}{2})$*-comonotone operator* $B : \mathscr{H} \rightrightarrows \mathscr{H}$*.*
(ii) *Let* $\alpha \in (0, \infty)$*. Then* T *is* α*-conically nonexpansive if and only if it is the resolvent of a maximally* ρ*-comonotone operator* $B : \mathscr{H} \rightrightarrows \mathscr{H}$*, where* $\rho = \frac{1}{2\alpha} - 1 > -\frac{1}{2}$*, i.e* $\alpha = \frac{1}{2(\rho+1)}$*.*
(iii) *Let* $\alpha \in (0, 1)$*. Then* T *is* $\alpha-$*averaged if and only if it is the resolvent of a maximally* $\rho-$ *comonotone operator* $B : \mathscr{H} \rightrightarrows \mathscr{H}$*, where* $\rho = \frac{1}{2\alpha} - 1 > -\frac{1}{2}$*, i.e* $\alpha = \frac{1}{2(\rho+1)}$*.*

Lemma 2.14. [15] *Let* $B : \mathscr{H} \rightrightarrows \mathscr{H}$ *be* ρ*-comonotone with* $\rho \in \mathbb{R}$ *and let* $\lambda, \mu > 0$*. If* $\rho > -\lambda, -\mu$*, then there exists a constant* $L > 0$*, such that*

$$\|x - J_\mu^B x\| \leq \Big(L + 1 + \frac{L\mu}{\lambda}\Big)\|x - J_\mu^B x\|, \ \forall\, x \in ran(I + \lambda B) \cap ran(I + \mu B)$$

Lemma 2.15. [1] *Let $B : \mathscr{H} \rightrightarrows \mathscr{H}$ be a maximally ρ-comonotone with $\rho > -\frac{1}{2}$. Then, $B^{-1}(0)$ is closed and convex.*

Lemma 2.16. [15] *Let $B : \mathscr{H} \rightrightarrows \mathscr{H}$ be a ρ-comonotone with $\rho > -1$. Then, the set $B^{-1}(0) = F(J_1^B)$, and consequently, $B^{-1}(0)$ is closed and convex.*

Proposition 2.17. [1] *Let $B : \mathscr{H} \rightrightarrows \mathscr{H}$ be such that $D(B) \neq \emptyset$. Let $\rho \in (-1, \infty)$, set $\mathscr{C} = ran(I + B)$, $T = J_\lambda^B$, i.e $B = T^{-1} - I$, and set $\alpha = \frac{1}{2(\rho+1)}$. Then the following are equivalent*

(i) *B is ρ-comonotone if and only if T is $\frac{1}{2(\rho+1)}$-conically nonexpansive.*
(ii) *B is maximally ρ-comonotone if and only if T is α-conically nonexpansive and $\mathscr{C} = \mathscr{H}$.*
(iii) *B is $(-\frac{1}{2})$-comonotone if and only if T is nonexpansive.*
(iv) *B is maximally $(-\frac{1}{2})$-comonotone if and only if T is nonexpansive and $\mathscr{C} = \mathscr{H}$.*
(v) *B is ρ-comonotone and $\rho > -\frac{1}{2}$ if and only if T is α-averaged.*
(vi) *B is maximally ρ-comonotone and $\rho > -\frac{1}{2}$ if and only if T is α-averaged and $\mathscr{C} = \mathscr{H}$.*

Lemma 2.18. *Let $\mathscr{H}$ be a real Hilbert space. Then, the following results hold for all $x, y \in \mathscr{H}$ and $\beta \in \mathbb{R}$*

(i) $\|x + y\|^2 \leq \|x\|^2 + 2\langle y, x + y\rangle$;
(ii) $\|x + y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2$;
(iii) $\|\beta x + (1-\beta)y\|^2 = \beta\|x\|^2 + (1-\beta)\|y\|^2 - \beta(1-\beta)\|x - y\|^2$.

Lemma 2.19. [13, Lemma 3.2] *Let $\{a_n\}$ and $\{b_n\}$ be sequences of positive numbers. Assume that the sequence $\{a_n\}$ is nonsummable, the sequence $\{b_n\}$ is decreasing and*

$$\sum_{i=0}^{\infty} a_i b_i < \infty.$$

Then,

$$b_n = o\Big(1/\sum_{i=0}^{n} b_i\Big)$$

where the o notation means that $t_n = o(1/s_n)$ if and only if $\lim_{n\to\infty} t_n s_n = 0$.

Remark 2.20. To analyze the proximal point algorithm, Dong [13] used Lemma 2.19.

Lemma 2.21. *(Opial) Let $\mathscr{C}$ be a nonempty subset of $\mathscr{H}$, and $\{x_n\}$ is a sequence in $\mathscr{H}$ such that the following conditions holds:*

(i) *for every $x \in \mathscr{C}$, $\lim_{n\to\infty} \|x_n - x\|$ exists;*
(ii) *every weak sequential limit point of $\{x_n\}$, as $n \to \infty$ belongs to $\mathscr{C}$.*

Then $\{x_n\}$ converges weakly as $n \to \infty$ to a point $\mathscr{C}$.

3. Main Results

In this section, we present our main result.

Theorem 3.1. *Let $B : \mathscr{H} \rightrightarrows \mathscr{H}$ be a maximally ρ-comonotone operator. Choose $\lambda > 0$ such that $\rho \in (-\lambda, 0]$. Given $x_0 \in \mathscr{H}$, generate*

$$x_{n+1} = (1-\beta_n)x_n + \beta_n J_\lambda^B(x_n), \tag{3.1}$$

where $0 < \liminf_{n\to\infty} \beta_n \leq \limsup_{n\to\infty} \beta_n < \frac{2(\rho+\lambda)}{\lambda}$. Then, $\{x_n\}$ converges weakly to an element of $Zer(B)$.

Proof. Let $x^* \in \text{Zer}(B)$. Then $x^* = J_\lambda^B x^*$. Since $\rho > -\lambda$, the resolvent J_λ^B is single-valued (see Lemma 2.7). Now, let $T = J_\lambda^B$. Then, $Tx_n = J_\lambda^B x_n$ is α-conically nonexpansive $\forall n \in \mathbb{N}$, where where $\alpha = \frac{\lambda}{2(\rho+\lambda)}$ (please, see Proposition 2.17 (i)). By Lemma 2.18 and Lemma 2.12 (iv), we have

$$\begin{aligned} \|J_\lambda^B x_n - x^*\|^2 &= \|x_n - x^* - (x_n - J_\lambda^B x_n)\|^2 \\ &= \|x_n - x^*\|^2 - 2\langle x_n - x^*, x_n - J_\lambda^B x_n\rangle + \|x_n - J_\lambda^B x_n\|^2 \\ &\leq \|x_n - x^*\|^2 - \frac{1}{\alpha}\|x_n - J_\lambda^B x_n\|^2 + \|x_n - J_\lambda^B x_n\|^2 \\ &= \|x_n - x^*\|^2 - \Big(\frac{1}{\alpha} - 1\Big)\|x_n - J_\lambda^B x_n\|^2 \\ &= \|x_n - x^*\|^2 - \frac{(2\rho+\lambda)}{\lambda}\|x_n - J_\lambda^B x_n\|^2. \end{aligned} \tag{3.2}$$

Consequently, from (3.8) and (3.2), we have

$$\begin{aligned} \|x_{n+1} - x^*\|^2 &= \|(1-\beta_n)(x_n - x^*) + \beta_n(J_\lambda^B x_n - x^*)\|^2 \\ &= (1-\beta_n)\|x_n - x^*\|^2 + \beta_n\|J_\lambda^B x_n - x^*\|^2 - \beta_n(1-\beta_n)\|x_n - J_\lambda^B x_n\|^2 \\ &\leq (1-\beta_n)\|x_n - x^*\|^2 + \beta_n\|x_n - x^*\|^2 - \beta_n\Big(\frac{2\rho+\lambda}{\lambda}\Big)\|x_n - J_\lambda^B x_n\|^2 \\ &\quad -\beta_n(1-\beta_n)\|x_n - J_\lambda^B x_n\|^2 \\ &= \|x_n - x^*\|^2 - \beta_n\Big(\frac{2\rho+\lambda}{\lambda} + 1 - \beta_n\Big)\|x_n - J_\lambda^B x_n\|^2. \end{aligned} \tag{3.3}$$

Since $0 < \liminf_{n\to\infty} \beta_n \leq \limsup_{n\to\infty} \beta_n < \frac{2(\rho+\lambda)}{\lambda}$, we obtain from (3.3) that

$$\|x_{n+1} - x^*\| \leq \|x_n - x^*\|, \; \forall\, x^* \in \text{Zer}(B),$$

and $\lim_{n\to\infty} \|x_n - x^*\|$ exists. Therefore, $\{x_n\}$ is bounded. Also

$$\lim_{n\to\infty} \|x_n - J_\lambda^B x_n\| = 0.$$

Since $\{x_n\}$ is bounded, there exists $\{x_{n_k}\} \subset \{x_n\}$ such that $x_{n_k} \rightharpoonup p \in \mathscr{H}$. By the fact that

$$\lim_{n\to\infty} \|x_n - Tx_n\| = \lim_{k\to\infty} \|x_{n_k} - Tx_{n_k}\| = 0$$

and demiclosedness property of $I - T$ at zero, we have that $p = Tp, T = J_\lambda^B$. Hence, $p \in \text{Zer}(B)$. By Lemma 2.21, we have that $\{x_n\}$ converges weakly to a point in $\text{Zer}(B)$. □

In the case when $\rho = 0$, we have the following result for the relaxed proximal point algorithm for a maximally monotone operator.

Corollary 3.2. *Suppose $B : \mathscr{H} \rightrightarrows \mathscr{H}$ is a maximally monotone operator. Choose $\lambda > 0$, $x_0 \in \mathscr{H}$, and generate*

$$x_{n+1} = (1-\beta_n)x_n + \beta_n J_\lambda^B(x_n), \tag{3.4}$$

where $0 < \liminf_{n\to\infty} \beta_n \leq \limsup_{n\to\infty} \beta_n < 2$. Then, $\{x_n\}$ converges weakly to an element of Zer(B).

In the case when $\rho < 0$, we have the following result for the relaxed proximal point algorithm for a maximally comonotone operator.

Corollary 3.3. *Suppose $B : \mathscr{H} \rightrightarrows \mathscr{H}$ is a maximally ρ-comonotone operator. Choose $\lambda > 0$ such that $\rho \in (-\lambda, 0)$. Choose $\lambda > 0$ such that $\rho + \lambda > 0$. Given $x_0 \in \mathscr{H}$, generate*

$$x_{n+1} = (1-\beta_n)x_n + \beta_n J_\lambda^B(x_n), \tag{3.5}$$

where $0 < \liminf_{n\to\infty} \beta_n \le \limsup_{n\to\infty} \beta_n < \frac{2(\rho+\lambda)}{\lambda}$. *Then,* $\{x_n\}$ *converges weakly to an element of Zer*(B).

We are now in the position to present the convergence rate of the proposed method.

Theorem 3.4. *Let* $B : \mathscr{H} \rightrightarrows \mathscr{H}$ *be a maximally* ρ*-comonotone operator. Choose* $\lambda > 0$ *such that* $\rho \in (-\lambda, 0]$. *Given* $x_0 \in \mathscr{H}$, *generate*

$$x_{n+1} = (1-\beta_n)x_n + \beta_n J_\lambda^B(x_n), \tag{3.6}$$

where $0 < \liminf_{n\to\infty} \beta_n \le \limsup_{n\to\infty} \beta_n < \frac{2(\rho+\lambda)}{\lambda}$ *such that* $\gamma_n := \sum_{j=0}^{n} \beta_j\Big(\frac{2\rho+\lambda}{\lambda} + 1 - \beta_j\Big)$ *for* $n \in \mathbb{N}$ *and* $\lim_{n\to\infty} \gamma_n = \infty$. *Then, the convergence rate estimate*

$$\|x_n - J_\lambda^B x_n\| = o\Big(1/\sqrt{\gamma_n}\Big)$$

holds, that is $\lim_{n\to\infty} \sqrt{\gamma_n}\|x_n - J_\lambda^B x_n\| = 0$.

Proof. From (3.3), we have

$$\|x_{n+1} - x^*\|^2 \quad \le \quad \|x_n - x^*\|^2 - \beta_n\Big(\frac{2\rho+\lambda}{\lambda} + 1 - \beta_n\Big)\|x_n - J_\lambda^B x_n\|^2$$

which implies that

$$\beta_n\Big(\frac{2\rho+\lambda}{\lambda} + 1 - \beta_n\Big)\|x_n - J_\lambda^B x_n\|^2 \quad \le \quad \|x_n - x^*\|^2 - \|x_{n+1} - x^*\|^2.$$

Summing from $j = 0$ to l we have

$$\begin{aligned}\sum_{j=0}^{l} \beta_j\Big(\frac{2\rho+\lambda}{\lambda} + 1 - \beta_j\Big)\|x_j - J_\lambda^B x_j\|^2 \quad &\le \quad \|x_j - x^*\|^2 - \|x_{j+1} - x^*\|^2 \\ &\le \quad \|x_0 - x^*\|.\end{aligned}$$

Letting $l \to \infty$, we have

$$\sum_{j=0}^{\infty} \beta_j\Big(\frac{2\rho+\lambda}{\lambda} + 1 - \beta_j\Big)\|x_j - J_\lambda^B x_j\|^2 < \infty,$$

which implies that

$$\|x_j - J_\lambda^B x_j\|^2 = o\Big(1/\sum_{j=0}^{\infty} \beta_j\Big(\frac{2\rho+\lambda}{\lambda} + 1 - \beta_j\Big)\Big).$$

Since $\lim_{n\to\infty} \gamma_n = \infty$, we obtain from Lemma 2.19 that

$$\|x_j - J_\lambda^B x_j\|^2 = o(1/\gamma_n).$$

Hence, we have that

$$\|x_j - J_\lambda^B x_j\| = o(1/\sqrt{\gamma_n}).$$

□

Remark 3.5.

(i) Theorem 3.4 improves the known big $\mathcal{O}$ in [11, Proposition 11] to little o without any other restrictions.

(ii) Let $\delta > 0$. The condition $\Big\{\beta_n\Big(\frac{2\rho+\lambda}{\lambda} + 1 - \beta_n\Big)\Big\} \subset (\delta, \infty)$ implies that $\lim_{n\to\infty} \gamma_n = \infty$ but the reverse does not hold.

(iii) Under the assumptions that $\delta > 0$ and

$$\left\{\beta_n\Big(\frac{2\rho+\lambda}{\lambda}+1-\beta_n\Big)\right\} \subset (\delta,\infty),$$

$$\sqrt{\delta(n+1)}\|\|x_n - J_\lambda^B x_n\|\| \leq \sqrt{\gamma_n}\|x_n - J_\lambda^B x_n\|.$$

Hence, the $o(1/\sqrt{n+1})$ rate in [12, Theorem 1] follows from Theorem 3.4.

In the case when $\rho = 0$, we have the following result for the relaxed proximal point algorithm for a maximally monotone operator.

Corollary 3.6. *Let $B : \mathscr{H} \rightrightarrows \mathscr{H}$ be a maximally monotone operator. Choose $\lambda > 0, x_0 \in \mathscr{H}$, and generate*

$$x_{n+1} = (1-\beta_n)x_n + \beta_n J_\lambda^B(x_n), \tag{3.7}$$

where $0 < \liminf_{n\to\infty} \beta_n \leq \limsup_{n\to\infty} \beta_n < 2$ such that $\gamma_n := \sum_{j=0}^{n} \beta_j\Big(2-\beta_j\Big)$ for $n \in \mathbb{N}$ and $\lim_{n\to\infty}\gamma_n = \infty$. Then, the convergence rate estimate

$$\|x_n - J_\lambda^B x_n\| = o\Big(1/\sqrt{\gamma_n}\Big)$$

holds, that is $\lim_{n\to\infty} \sqrt{\gamma_n}\|x_n - J_\lambda^B x_n\| = 0$.

In the case when $\rho < 0$, we have the following result for the relaxed proximal point algorithm for a maximally comonotone operator.

Corollary 3.7. *Let $B : \mathscr{H} \rightrightarrows \mathscr{H}$ be a maximally ρ-comonotone operator. Choose $\lambda > 0$ such that $\rho \in (-\lambda, 0)$. Choose $\lambda > 0$ such that $\rho + \lambda > 0$. Given $x_0 \in \mathscr{H}$, generate*

$$x_{n+1} = (1-\beta_n)x_n + \beta_n J_\lambda^B(x_n), \tag{3.8}$$

where $0 < \liminf_{n\to\infty} \beta_n \leq \limsup_{n\to\infty} \beta_n < \frac{2(\rho+\lambda)}{\lambda}$ such that $\gamma_n := \sum_{j=0}^{n} \beta_j\Big(\frac{2\rho+\lambda}{\lambda}+1-\beta_j\Big)$ for $n \in \mathbb{N}$ and $\lim_{n\to\infty}\gamma_n = \infty$. Then, the convergence rate estimate

$$\|x_n - J_\lambda^B x_n\| = o\Big(1/\sqrt{\gamma_n}\Big)$$

holds, that is $\lim_{n\to\infty} \sqrt{\gamma_n}\|x_n - J_\lambda^B x_n\| = 0$.

4. Conclusion

In this paper, we studied the generalized inclusion problem in the framework of real Hilbert spaces. We investigated the convergence of the proximal point algorithm for solving inclusion problems involving maximally comonotone operators. Using the relationship between comonotonicity and the averaged nature of resolvent operators, we obtained the weak convergence of our method. By relaxing monotonicity to comonotonicity, we expand the reach of the proximal point algorithm to problems in nonconvex optimization, game theory, and variational inequalities, where traditional approaches might stall. In the future, we may explore adaptive stepsize, accelerated variants, or applications to structured nonmonotone problems in machine learning and imaging sciences.

Statements and Declarations

The authors declare that they have no conflict of interest, and the manuscript has no associated data.

References

[1] H. H. Bauschke, W. A. Moursi, and X. Wang. Generalized monotone operators and their averaged resolvents. *Mathematical Programming Series B*, 189:55-74, 2020.

[2] H. H. Bauschke and P. L. Combettes. Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Springer, New york, 2nd edition, 2017.

[3] H. Brézis and P. L. Lions. Produits infinis de resolvantes. *Israel Journal of Mathematics*, 29:329–345, 1978.

[4] K. Bhatia and K. Sridharan. Online learning with dynamics: A minimax perspective. *Advances in Neural Information Processing Systems*, 33:15020–15030, 2020..

[5] L. Bottou, F. E. Curtis, and J. Nocedal. Optimization methods for large-scale machine learning. *SIAM Review*, 60(2):223–311, 2018.

[6] R. S. Burachik and A. N. Iusem. Set-Valued Mappings and Enlargements of Monotone Operators. Springer, Berlin, 2008.

[7] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging Vision*, 40(1):120—145, 2011.

[8] A. Chambolle and T. Pock. On the ergodic convergence rates of a first-order primal-dual algorithm. *Mathematical Programming* 159(1):253—287, 2016.

[9] P. L. Combettes. Monotone operator theory in convex optimization. *Mathematical Programming*, 170(1): 177–206, 2018.

[10] P. L. Combettes and T. Pennanen. Proximal methods for cohypomonotone operators. *SIAM Journal of Control and Optimization*, 43:731–742, 2004.

[11] R. Cominetti, J. A. Soto, and J. Vaisman. On the rate of convergence of Krasnoselśkiĩ–Mann iterations and their connection with sums of Bernoullis. *Israel Journal of Mathematics,* 199:757—772, 2014.

[12] D. Davis and W. Yin. Convergence rate analysis of several splitting schemes. In R. Glowinski, S. Osher, W. Yin, editors, *Splitting Methods in Communication and Imaging, Science and Engineering*, pages 115-163, Springer, New York.

[13] Y. Dong. Comments on "the proximal point algorithm revisited". *Journal of Optimization Theory and Application*, 116:343—349, 2015.

[14] J. Douglas and H. H. Rachford. On the numerical solution of heat conduction problems in two and three space variables. *Transactions of the American Mathematical Society*, 82(2):421—439, 1956.

[15] M. Eslamian. Split common null point problem for generalized monotone operators. *Optimization,* 1—17, 2024.

[16] D. Gabay and B. Mercier. A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Computers and Mathematics with Application*, 2(1):17—40, 1976.

[17] R. Glowinski and A. Marroco. Sur lápproximation, par éléments finis dórdre un, et la résolution, par pénalisation-dualité d'une classe de problémes de Dirichlet non linéaires. *Esaim-Mathematical Modelling and Numerical Analysis,* 9(R2):41—76, 1975.

[18] A. N. Iusem, T. Pennanen, and B. F. Svaiter. Inexact variants of the proximal point algorithm without monotonicity. *SIAM Journal of Optimization*, 13(4):1080–1097, 2003.

[19] A. Jabbar, X. Li, and B. Omar. A survey on generative adversarial networks: Variants, applications and training. *ACM Computing Surveys (CSUR)*, 54(8):1–49, 2021.

[20] G. Lan. First-order and Stochastic Optimization Methods for Machine Learning. Springer, Cham, 2020.

[21] F. Lin, X. Fang, and Z. Gao. Distributionally robust optimization: A review on theory and applications. *Numerical Algebra, Control and Optimization*, 12(1):159–212, 2022.

[22] P. L. Lions and B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal of Numerical Analysis*, 16(6):964—979, 1979.

[23] Y. Malitsky. Golden ratio algorithms for variational inequalities. *Mathematical Programming*, 184:383–410, 2020.

[24] B. Martinet. Regularisation d'inéquations variationelles par approximations successives. *Revue Française d'Automatique, Informatique, Recherche Opérationnelle*, 4:154–158, 1970.

[25] T. Pennanen. Local convergence of the proximal point algorithm and multiplier methods without monotonicity. *Mathematics of Operation Research*, 27:170–191, 2002.

[26] R. T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM Journal of Control and Optimization*, 14:877–898, 1976.

[27] R. T. Rockafellar. Augmented Lagrangians and applications of the proximal point algorithm in convex programming. *Mathematics of Operation Research*, 1(2):97—116, 1976.

[28] R. T. Rockafellar and R. J. -B. Wets. Variational Analysis. Springer, Berlin, 2009.

[29] M. V. Solodov and B. F. Svaiter. A new projection method for variational inequality problems. *SIAM Journal of Control and Optimization*, 37(3):765–776, 1999.

[30] P. Tseng. A modified forward-backward splitting method for maximal monotone mappings. *SIAM Journal of Control and Optimization*, 38(2):431–446, 2000.