On the Proximal Point Algorithm and its Halpern-type variant for generalized monotone operators in Hilbert space

Ulrich Kohlenbach
Department of Mathematics
Technische Universität Darmstadt
Schlossgartenstraße 7, 64289 Darmstadt, Germany
kohlenbach@mathematik.tu-darmstadt.de

April 3, 2021

Abstract

In a recent paper, Bauschke et al. study ρ-comonotonicity as a generalized notion of monotonicity of set-valued operators A in Hilbert space and characterize this condition on A in terms of the averagedness of its resolvent J_A. In this note we show that this result makes it possible to adapt many proofs of properties of the proximal point algorithm PPA and its strongly convergent Halpern-type variant HPPA to this more general class of operators. This also applies to quantitative results on the rates of convergence or metastability (in the sense of T. Tao). E.g. using this approach we get a simple proof for the convergence of the PPA in the boundedly compact case for ρ-comonotone operators and obtain an effective rate of metastability. If A has a modulus of regularity w.r.t. $zer\,A$ we also get a rate of convergence to some zero of A even without any compactness assumption. We also study a Halpern-type variant HPPA of the PPA for ρ-comonotone operators, prove its strong convergence (without any compactness or regularity assumption) and give a rate of metastability.

Keywords: Generalized monotone operators, proximal point algorithm, Halpern-type proximal point algorithm, rates of convergence, metastability, proof mining.
Mathematics Subject Classification (2010): 47H05, 47J25, 03F10

1 Introduction

A central theme in convex optimization is the computation of zeros $z \in zer\,A := A^{-1}(0)$ of (maximally) monotone set-valued operators $A \subseteq H \times H$ in Hilbert space H. This stems from the fact that for A being the subdifferential ∂f of a proper, convex and lower semi-continuous function $f : H \to (-\infty, \infty]$, $zer\,A$ coincides with the set of minimizers of f.

An important algorithm for the approximation of zeros of A is the Proximal Point Algorithm PPA ([17, 20])

$$x_{n+1} := J_{\gamma_n A} x_n, \quad (\gamma_n) \subset (0, \infty),$$

where $J_{\gamma_n} := (I + \gamma_n A)^{-1} : R(I + \gamma_n A) \to D(A)$ is the single valued resolvent of $\gamma_n A$ and A is assumed to satisfy some range condition such as $\overline{D(A)} \subseteq R(I + \lambda A)$ for all $\lambda > 0$ so that the iteration is defined for $x_0 \in \overline{D(A)} \subseteq R(I + \gamma_0 A)$. Here $D(A)$ and $R(A)$ denote the domain and range of A respectively as defined for set-valued mappings (see e.g. [4]).

The range condition trivially holds for maximally monotone operators A such as ∂f since then $R(I+\lambda A)=H$.

The crucial relation between A and $J_{\lambda A}$ is that the set of zeros of A coincides with the fixed point set of $J_{\lambda A}$ (which, therefore, in particular does not depend on the choice of $\lambda>0$). If A is monotone, then $J_{\lambda A}$ is firmly nonexpansive so that many results from metric fixed point theory apply (see e.g. [4] for all this).

In order to be able to treat functions f which are not necessarily convex, one needs to weaken the requirement of A to be monotone from

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

to e.g. stipulating

$$(++)\ \forall (x,u),(y,v)\in A\,(\langle x-y,u-v\rangle \geq \rho\|u-v\|^2),$$

where now ρ may also be negative (see e.g. [7, 8]).

In the recent paper [5], this condition - called ρ-comonotonicity - is thoroughly investigated and related to properties of J_A. One key result is that J_A is an averaged mapping whenever (++) holds with $\rho>-\frac{1}{2}$. The averaged mappings form a larger class of mappings than the firmly nonexpansive ones but still have nice properties, e.g. they are strongly nonexpansive.

In the recent papers [12, 13], we studied from a quantitative point of view the PPA as well as a strongly convergent so-called Halpern-type variant HPPA (in Banach spaces) making use essentially only of the fact that all firmly nonexpansive mappings have a common so-called modulus for being strongly nonexpansive (see [10]). This also holds true for the class of averaged mappings if we have some control on the averaging constant (see [21]). Putting all this together, it is rather straightforward to see that the main results on the PPA and HPPA established in [12, 13] generalize (in the case of Hilbert spaces) to ρ-comonotone operators which is the content of this short note. While the PPA has been considered for ρ-comonotone operators before (even for sequences of operators, error terms and relaxations, see [7]) our note shows that by the connection between the comonotonicity of A and the averagedness of J_A as established in [5], many proofs for properties of the PPA and the HPPA for monotone operators can be easily adapted to cover the ρ-comonotone case. We also provide new quantitative results on the convergence. For the HPPA, to the best of our knowledge, our note provides the first results in the absence of monotonicity.

2 Preparatory results

Throughout this paper H is a real Hilbert space and $A\subseteq H\times H$ a set-valued operator with the usual definitions of $D(A)$ and $zer\,A$. $\overline{D(A)}$ denotes the topological closure of $D(A)$. We always assume that $D(A)\neq\emptyset$.

Definition 2.1 ([5]). *Let $\rho\in\mathbb{R}$. A is called ρ-comonotone if*

$$\forall (x,u),(y,v)\in A\,(\langle x-y,u-v\rangle \geq \rho\|u-v\|^2).$$

In the case where $\rho<0$ which we are interested in, ρ-comonotonicity has been studied before in [7] under the name of $|\rho|$-cohypomonotonicity in the context of proximal methods as discussed in the introduction (see also Remark 3.4 below).

Let $J_A:=(I+A)^{-1}$ be the resolvent of A.

Proposition 2.2. *Let $\rho \in \mathbb{R}, \lambda > 0$ and A be ρ-comonotone. Then $D(J_{\lambda A}) = R(I + \lambda A)$, $x \in J_{\lambda A}x \leftrightarrow x \in zer\, A$ and, if $\rho > -1$, J_A is at most single-valued and $zer\, A = Fix(J_A)$.*

Proof: [4, Proposition 23.2] and [5, Proposition 2.13]. □

Lemma 2.3. *If A is ρ-comonotone for $\rho \in \mathbb{R}$, then for $\lambda > 0$ we have that λA is ρ/λ-comonotone.*

Proof: If $u \in \lambda Ax, v \in \lambda Ay$, then $\frac{u}{\lambda} \in Ax, \frac{v}{\lambda} \in Ay$ and so

$$\langle x - y, u - v\rangle = \lambda \left\langle x - y, \frac{u}{\lambda} - \frac{v}{\lambda}\right\rangle \geq \lambda \cdot \rho \left\|\frac{u}{\lambda} - \frac{v}{\lambda}\right\|^2 = \frac{\rho}{\lambda}\|u - v\|^2.$$

□

The following proposition, which is well-known for monotone operators, extends to ρ-comonotone operators:

Proposition 2.4. *Let $A \subseteq H \times H$ be ρ-comonotone with $\rho \in \mathbb{R}$. Let $\lambda, \mu > 0$.*

1. *If $\rho \geq -\frac{\lambda}{2}$, then $J_{\lambda A}$ is nonexpansive.*
2. *$J_{\lambda A}$ satisfies the resolvent equation in the following form: if $\rho > -\lambda, -\mu$, then*

$$J_{\lambda A}x = J_{\mu A}\left(\frac{\mu}{\lambda}x + \left(1 - \frac{\mu}{\lambda}\right) J_{\lambda A}x\right), \quad x \in D(J_{\lambda A}).$$

3. *If $\rho \geq -\frac{\lambda}{2}, -\frac{\mu}{2}$ then*

$$\|x - J_{\mu A}x\| \leq \left(2 + \frac{\mu}{\lambda}\right) \|x - J_{\lambda A}x\|$$

for all $x \in R(I + \lambda A) \cap R(I + \mu A)$.

Proof: 1) By the assumptions and lemma 2.3, λA is $-\frac{1}{2}$-comonotone and so - by [5, Proposition 3.11(iii)] - $J_{\lambda A}$ is nonexpansive.
2) follows as in [3][p.105] using Proposition 2.2 which is applicable since - by Lemma 2.3 - $J_{\lambda A}, J_{\mu A}$ are > -1-comonotone.
3) Using 1) and 2) we get

$$\begin{aligned}
&\|x - J_{\mu A}x\| \leq \|x - J_{\lambda A}x\| + \|J_{\lambda A}x - J_{\mu A}x\| = \\
&\|x - J_{\lambda A}x\| + \left\|J_{\mu A}\left(\tfrac{\mu}{\lambda}x + (1 - \tfrac{\mu}{\lambda})J_{\lambda A}x\right) - J_{\mu A}x\right\| \leq \\
&\|x - J_{\lambda A}x\| + \left\|\tfrac{\mu}{\lambda}x + (1 - \tfrac{\mu}{\lambda})J_{\lambda A}x - x\right\| = \\
&\|x - J_{\lambda A}x\| + \left|1 - \tfrac{\mu}{\lambda}\right| \|x - J_{\lambda A}x\| \leq \left(2 + \tfrac{\mu}{\lambda}\right) \|x - J_{\lambda A}x\|.
\end{aligned}$$

□

Definition 2.5 ([6])**.** *Let $C \subseteq H$ be a nonempty subset of H and $T : C \to H$ be a mapping.*

1. *T is called α-averaged with $\alpha \in (0, 1)$ if $T = (1 - \alpha)I + \alpha S$, where $S : C \to H$ is nonexpansive.*
2. *T is called strongly nonexpansive (SNE) if T is nonexpansive and for all sequences $(x_n), (y_n)$ in H the following implication is true:*

   if $((x_n - y_n)$ bounded $\wedge \|x_n - y_n\| - \|Tx_n - Ty_n\| \to 0)$, then $(x_n - y_n) - (Tx_n - Ty_n) \to 0$.

Lemma 2.6. [10, Lemma 2.2] $T : C \to H$ *is strongly nonexpansive iff* T *has as an* SNE-*modulus* $\omega : (0,\infty)^2 \to (0,\infty)$, *i.e.*

$$\forall b, \varepsilon > 0\, \forall x, y \in C\ (\|x-y\| \le b \wedge \|x-y\| - \|Tx - Ty\| < \omega(b,\varepsilon) \to \|(x-y) - (Tx - Ty)\| < \varepsilon)\,.$$

The proof of [21, Proposition 2.7] establishes:

Proposition 2.7 ([21])**.** *Let* $C \subseteq H$ *be some subset of* H *and* $T : C \to H$ *be an* α*-averaged mapping for some* $\alpha \in (0,1)$. *Then* T *is strongly nonexpansive with* SNE-*modulus*

$$\omega_\alpha(b,\varepsilon) := \frac{1-\alpha}{4b\alpha} \cdot \varepsilon^2.$$

Proposition 2.8. *Let* $(\gamma_n) \subset (0,\infty), \gamma > 0$ *be such that* $\gamma_n \ge \gamma > 0$ *for all* $n \in \mathbb{N}$. *Let* $\rho \in (-\frac{\gamma}{2}, 0]$ *and* $A \subseteq H \times H$ *be* ρ*-comonotone.*
Then for each $n \in \mathbb{N}$, $J_{\gamma_n A} : R(I + \gamma_n A) \to D(A)$ *is strongly nonexpansive with common* SNE-*modulus* ω_α, *where* $\alpha := \frac{1}{2((\rho/\gamma)+1)} \in (0,1)$.
In particular, if $D(A) \subseteq C \subseteq \bigcap_{n=0}^{\infty} R(I + \gamma_n A)$, *then* $(J_{\gamma_n A})$ *(restricted to* C*) is a strongly nonexpansive sequence of mappings* $C \to C$ *in the sense of the papers* [1, 2].

Proof: By the assumptions and Lemma 2.3, $\gamma_n A$ is (ρ/γ_n)-comonotone and so, since

$$\frac{\rho}{\gamma_n} \ge \frac{\rho}{\gamma} > -\frac{1}{2},$$

it a fortiori is η-comonotone with $\eta := \frac{\rho}{\gamma} > -\frac{1}{2}$. Hence by [5, Proposition 3.11(v)] applied to $\gamma_n A$, the resolvent $J_{\gamma_n A} : R(I + \gamma_n A) \to D(A)$ is α-averaged. The claim now follows from Proposition 2.7. □

3 The Proximal Point Algorithm PPA for comonotone operators

Let $A \subseteq H \times H$ be ρ-comonotone, $(\gamma_n) \subset (0,\infty)$ and assume that $D(A) \subseteq \bigcap_{n=0}^{\infty} R(I + \gamma_n A)$. We assume that $zer\,A \neq \emptyset$. The Proximal Point Algorithm PPA for A and (γ_n) is defined by ($n \in \mathbb{N} = \{0,1,2,\ldots\}$)

$$x_{n+1} := J_{\gamma_n A} x_n, \quad x_0 \in R(I + \gamma_0 A).$$

Throughout this section we also assume that $\gamma_n \ge \gamma > 0$ for all $n \in \mathbb{N}$ and that $\rho \in (-\frac{\gamma}{2}, 0]$.

Proposition 3.1. *1.*

$$\lim_{n\to\infty} \|x_n - J_{\gamma_0 A} x_n\| = \lim_{n\to\infty} \|x_n - x_{n+1}\| = 0.$$

Moreover, with $\alpha := \frac{1}{2((\rho/\gamma)+1)} \in (0,1)$, ω_α *as in Proposition 2.7 and* $b \ge \|x_0 - p\|$ *for some* $p \in zer\,A$,

$$\Delta(\varepsilon, L, b) := \lceil b/\omega_\alpha(b,\varepsilon) \rceil + L + 1$$

is a modulus of lim inf *(in the sense of* [14]*) i.e.*

$$\forall L \in \mathbb{N},\, \varepsilon > 0\, \exists n\ (L \le n \le \Delta(\varepsilon, L, b) \text{ and } \|x_n - x_{n+1}\| < \varepsilon)\,.$$

2. Define

$$u_n := \frac{x_n - x_{n+1}}{\gamma_n}.$$

Then $u_n \in Ax_{n+1}$, $\lim_{n\to\infty} u_n = 0$ *and*

$$\exists n \le \rho(\varepsilon, b, \gamma) := \Delta(\varepsilon\cdot\gamma, 0, b)\ (\|u_n\| < \varepsilon).$$

Proof: 1) Let $p \in zer\,A$. Then by Propositions 2.2 and 2.4 (using that $\gamma_n A$ is $> -\frac{1}{2} > -1$ comonotone)

$$\|x_{n+1} - p\| \le \|x_n - p\| \le b, \quad n \in \mathbb{N},$$

and so (x_n) is Fejér monotone w.r.t. $zer\,A = Fix(J_{\gamma_n A})$ and $(\|x_n - p\|)$ is convergent. Thus

$$|\|J_{\gamma_n A}x_n - J_{\gamma_n A}p\| - \|x_n - p\|| = |\|x_{n+1} - p\| - \|x_n - p\|| \to 0.$$

Hence by Proposition 2.8

$$\|x_{n+1} - x_n\| = \|J_{\gamma_n A}x_n - x_n\| \to 0.$$

By Proposition 2.4(3) (which is applicable in the nontrivial case where $n \ge 1$ due to $x_n \in D(A)$ and the range condition)

$$\|x_n - J_{\gamma_0 A}x_n\| \le \left(2 + \frac{\gamma_0}{\gamma}\right)\|x_n - J_{\gamma_n A}x_n\|$$

and so also $\lim_{n\to\infty} \|x_n - J_{\gamma_0 A}x_n\| = 0$.

The lim inf-bound is proved as in [13, Proposition 2.1] using Proposition 2.8. We include the proof here for completeness: Let $L \in \mathbb{N}$ and $\delta > 0$. Then there exists an $n \in \mathbb{N}$ with $L \le n \le L + \lceil b/\delta\rceil + 1$ such that

$$\|x_n - p\| - \|J_{\gamma_n A}x_n - J_{\gamma_n A}p\| = \|x_n - p\| - \|x_{n+1} - p\| < \delta$$

since, otherwise,

$$b \ge \|x_L - p\| \ge \|x_L - p\| - \|x_{L+\lceil b/\delta\rceil+1} - p\| \ge (\lceil b/\delta\rceil + 1)\cdot\delta > b.$$

Now fix $\delta := \omega_\alpha(b, \varepsilon)$. Then Proposition 2.8 implies the existence of an n with $L \le n \le \Delta(\varepsilon, L, b)$ such that

$$\|x_n - x_{n+1}\| = \|(x_n - p) - (J_{\gamma_n A}x_n - J_{\gamma_n A}p)\| < \varepsilon.$$

2) is immediate from 1). □

The PPA for maximally monotone operators, while being weakly convergent, fails to be strongly convergent as shown in [9]. In the boundedly compact (i.e. finite dimensional) case there is in general no computable rate of convergence unless some strong metric regularity assumption is made (see [19] and [11]). However, in the boundedly compact case, one can get effective rates Ψ of metastability in the sense of T. Tao [24, 25] for the Cauchy property of (x_n), i.e.

$$\forall \varepsilon > 0\, \forall g : \mathbb{N} \to \mathbb{N}\, \exists n \le \Psi(\varepsilon, g)\, \forall i, j \in [n, n + g(n)]\ (\|x_i - x_j\| < \varepsilon).$$

Note that, noneffectively, this property implies the Cauchy property of (x_n) and hence the existence of a limit x but does not allow one to convert Ψ into an effective rate of convergence. One can additionally ensure that for $i \in [n, n + g(n)]$, x_i is an approximate zero of A which guarantees that x is a zero of A.

We now extend our rate of metastability for the PPA from [13] to the ρ-comononotone case:

Theorem 3.2. *Let A be as above and assume additionally that $\overline{D(A)} \subseteq \bigcap_{n=0}^{\infty} R(I+\gamma_n A)$ is boundedly compact and $x_0 \in \overline{D(A)}$. Then (x_n) strongly converges to a zero of A. Moreover, the rate of metastability Ψ from* [13, Theorem 2.12] *also holds in our current situation with Δ being replaced by our definition in Proposition 3.1(1), i.e.*

$$(*)\ \forall k \in \mathbb{N}\, \forall g \in \mathbb{N}^{\mathbb{N}}\, \exists n \leq \Psi(k,g,\beta)\, \forall i,j \in [n, n+g(n)] \left(\|x_i - x_j\| \leq \frac{1}{k+1} \text{ and } x_i \in \tilde{F}_k \right),$$

where

$$\tilde{F}_k := \bigcap_{i \leq k} \left\{ x \in \overline{D(A)} \,:\, \|x - J_{\gamma_i A} x\| \leq \frac{1}{k+1} \right\}$$

and β is a modulus of total boundedness (in the sense of [13, Theorem 2.12]*) for $\overline{D(A)} \cap \overline{B}(0,M)$, where $\overline{B}(0,M) := \{x \in H : \|x\| \leq M\}$, with $M \geq b + \|p\|$ and $b \geq \|x_0 - p\|$ for some $p \in zer\, A$.*
If $C \subseteq H$ is closed and convex with $\overline{D(A)} \subseteq C \subseteq \bigcap_{n=0}^{\infty} R(I+\gamma_n A)$, then without compactness assumption, (x_n) converges weakly to a zero of A.

Proof: The proof of [13, Theorem 2.12] for the rate of metastability of (x_n) can be taken without any changes observing that [14, Lemma 8.1] holds with the same proof in our context and that Φ can be shown to be an approximate F-bound as in [13, Proposition 2.11] using Propositions 2.4(3) and 3.1(1) instead of [13, Prop.2.3(ii),Prop.2.1].
Since (x_n) is metastable (the first part of $(*)$), it is a Cauchy sequence and hence convergent with $x := \lim_n x_n \in \overline{D(A)}$. By the extra clause '$x_i \in \tilde{F}_k$' in $(*)$, which strengthens the usual formulation of a rate of metastability, we can conclude that $x \in zer\, A$. Indeed, choosing in $(*)$ for given $N \in \mathbb{N}$ the function $g(n) := N$ we get an $n_N \geq N$ with $\|x_{n_N} - J_{\gamma_0 A} x_{n_N}\| \leq \frac{1}{k+1}$. Using the nonexpansivity of $J_{\gamma_0 A}$ this implies that $x \in Fix(J_{\gamma_0 A}) = zer\, A$.
For the weak convergence in the noncompact case we reason as follows: let w be a weak sequential cluster point of (x_n). Then there is a subsequence (x_{n_k}) which weakly converges to w. By Proposition 3.1(1) (x_{n_k}) is an approximate fixed point sequence of $J_{\gamma_0 A}$. Hence by Browder's demiclosedness principle ([4, Corollary 4.28]) applied to $J_{\gamma_0 A}$ and C it follows that $w \in Fix(J_{\gamma_0 A})$. Hence we can - using again the fact that (x_n) is Fejér monotone w.r.t. $Fix(J_{\gamma_0 A})$ - conclude that (x_n) weakly converges to $w \in Fix(J_{\gamma_0 A}) = zer\, A$ by [4, Theorem 5.5]. ☐

Remark 3.3. *The range condition in Theorem 3.2 is trivially satisfied if A is maximally ρ-comonotone (in the sense of* [5, Definition 2.4.(iv)]*) since then by Lemma 2.3 λA is maximally (ρ/λ)-comonotone with $\rho/\lambda > -1$ for $\lambda \geq \gamma$ so that by* [5, Corollary 2.12] *$R(I + \lambda A) = H$.*

Remark 3.4. *Note that the conditions on ρ, γ_n made in* [7] *on their general* PPA *in the case of a single operator A and without relaxation (i.e. $\lambda_n := 1$) imply our condition that $\rho > -\frac{\inf\{\gamma_n : n \in \mathbb{N}\}}{2}$: observing that ρ in* [7] *corresponds to our $-\rho$, the conditions (iii),(iv) in* [7, Theorem 3.1] *state the existence of an $\varepsilon \in (0,1)$ s.t.*

$$\frac{1}{1+\rho/\gamma_n} \leq 2 - \varepsilon,\ \gamma_n > -\rho, \quad n \in \mathbb{N}.$$

An easy calculation shows that this implies that

$$\inf\{\gamma_n : n \in \mathbb{N}\} \geq -\rho \frac{2-\varepsilon}{1-\varepsilon} > -2\rho.$$

Also the converse holds: let $\delta > 0$ be such that $\gamma > -2\rho + \delta$. Then the condition

$$\frac{1}{1 + \frac{\rho}{\gamma_n}} < 2 - \varepsilon$$

is satisfied with $\varepsilon := 2 - \frac{2\gamma}{\gamma+\delta}$.
Error terms u_n subject to the condition that $\sum \|u_n\| < \infty$ (implied by condition (vi) in [7, Theorem 3.1]*) can be incorporated even in the quantitative part of our theorem (similar to* [15, Theorem 4.5]*). Our approach makes the relevance of the averagedness of $J_{\gamma_n A}$ explicit which only implicitly occurs in the proof of* [7, Theorem 3.1].

Definition 3.5 ([16])**.** *Let A be as at the beginning of this section with $p \in zer\,A$ and define $F(x) := \mathrm{dist}(0_X, A(x))$ (with $F(x) := \infty$ for $x \notin D(A)$). A function $\phi : (0,\infty) \to (0,\infty)$ is called a 'modulus of regularity for A w.r.t. $zer\,A$ and $\overline{B}(p,r)$ with $r > 0$' if for all $\varepsilon > 0$ and $x \in \overline{B}(p,r) := \{y \in H : \|y - p\| \leq r\}$ one has*

$$F(x) < \phi(\varepsilon) \to \mathrm{dist}(x, zer\,F) < \varepsilon.$$

As [13, Lemma 2.6] (but reasoning in the proof of $zer\,F \subseteq zer\,A$ with - say - $\gamma_0 A$ and $J_{\gamma_0 A}$ instead of A, J_A) one shows that

Lemma 3.6. *With F as defined in the previous definition, $zer\,F = zer\,A$ and so (x_n) as defined by the* PPA *for A is Fejér monotone w.r.t. $zer\,F = zer\,A$, i.e.*

$$\forall p \in zer\,F\, \forall n \in \mathbb{N}\, (\|x_{n+1} - p\| \leq \|x_n - p\|).$$

As in the case of [13, Theorem 2.8] one now gets

Theorem 3.7. *Let A and (γ_n) be as above and assume that $\overline{D(A)} \subseteq \bigcap_{n=0}^{\infty} R(I + \gamma_n A)$. Let $p \in zer\,A$ and $b \geq \|x_0 - p\|$. If A has a modulus ϕ of regularity w.r.t $zer\,A$ and $\overline{B}(p,b)$, then (x_n) converges to a zero $z := \lim x_n$ of A with rate of convergence $\rho(\phi(\varepsilon/2), b, \gamma) + 1$, where ρ is as in Proposition 3.1(2).*

Proof: The proof is largely identical to that of [13, Theorem 2.8]. We only have to observe that in that latter proof it suffices to have the existence of an $n \leq \rho(\varepsilon, b, \gamma)\, (|F(x_{n+1})| \leq \|u_n\| \leq \varepsilon)$ (rather than that this holds for all $n \geq \rho(\varepsilon, b, \gamma)$) and that this follows from Proposition 3.1(2). □

4 The Halpern-type Proximal Point Algorithm HPPA for comonotone operators

Whereas the PPA even for monotone operators A in general is not strongly convergent ([9]) a Halpern-type variant strongly converges also for ρ-comonotone operators as we show in this section. Again we assume that $(\gamma_n) \subset (0,\infty)$ with $\gamma_n \geq \gamma > 0$ for all $n \in \mathbb{N}$ and that A is ρ-comonotone with $\rho \in (-\frac{\gamma}{2}, 0]$ with $zer\,A \neq \emptyset$. Let $C \subseteq H$ be a nonempty closed and convex subset such that $\overline{D(A)} \subseteq C \subseteq \bigcap_{n=0}^{\infty} R(I + \gamma_n A)$.

Definition 4.1 ([2, 18, 23]). *Let $S \subseteq H$ be some nonempty subset of H and $T : S \to H$ a mapping and (S_n) be a sequence of mappings $S_n : S \to H$. Let $F((S_n)) := \bigcap_{n\in\mathbb{N}} Fix(S_n)$ be the set of all common fixed points of S_n for all n. (S_n) is said to satisfy the* NST condition (I) with *T if $F((S_n)) \neq \emptyset$, $Fix(T) \subseteq F((S_n))$ and $x_n - Tx_n \to 0$ whenever (x_n) is a bounded sequence in S with $x_n - S_nx_n \to 0$.*

Proposition 4.2. *Let $T := J_{\gamma_0 A} : C \to C$ and $S_n := J_{\gamma_n A} : C \to C$. Then (S_n) (strictly speaking the sequence of the restrictions of S_n to C) satisfies the* NST *condition* (I) *with T.*

Proof: Clearly, $Fix(T) = Fix(S_n) = zer\, A \neq \emptyset$. Let (x_n) be a bounded sequence in C with $\lim_n \|x_n - S_nx_n\| = 0$. Then by Proposition 2.4(2) also $\lim_n \|x_n - Tx_n\| = 0$. □

Theorem 4.3. *Let $(\alpha_n) \subset (0,1]$ be such that $\lim_n \alpha_n = 0$ and $\sum_{n=0}^{\infty} \alpha_n = \infty$. For $u, x_0 \in C$ define the Halpern-type proximal point algorithm* (HPPA) *by*

$$x_{n+1} := \alpha_n u + (1-\alpha_n) J_{\gamma_n A} x_n \in C.$$

Then (x_n) strongly converges to the zero of A which is closest to u. Moreover, the rate of metastability from [12, Theorem 4.1] *also holds for our current situation if ω_η is replaced by ω_α from Proposition 2.7 above with $\alpha := \frac{1}{2((\rho/\gamma)+1)}$ and $\omega_J(b,\varepsilon) := \varepsilon$.*

Proof: The strong convergence follows from [2, Theorem 3.1] whose assumptions are satisfied by Propositions 2.4(1), 2.8 and 4.2 using also that H has the fixed point property for nonexpansive mappings. The strong convergence also follows using [12, Theorem 4.1] which, moreover, gives the rate of metastability stated in the theorem. For this we only have to observe that the proof of [12, Theorem 4.1] only uses properties of $J_{\gamma_n A}$ which by the results stated above also hold true for ρ-comonotone operators A where now we use ω_α and Proposition 2.8 instead of ω_η and [12, Lemma 2.4]. Finally, we note that we can take $\omega_J(b,\varepsilon) := \varepsilon$ as modulus of uniform continuity for the normalized duality map on $\overline{B}(0,b)$ since we are in a Hilbert space. □

Remark 4.4. *Remark 3.3 applies here as well: if A is maximally ρ-comonotone, then the range condition is satisfied for any closed and convex subset $C \subseteq H$ satisfying $\overline{D(A)} \subseteq C$.*

Acknowledgments: The author has been supported by the German Science Foundation (DFG Project KO 1737/6-1).

References

[1] Aoyama, K., Toyoda, M., Approximation of zeros of accretive operators in a Banach space. Israel J. Math. **220**, pp. 803-816 (2017).

[2] Aoyama, K., Toyoda, M., Approximation of common fixed points of strongly nonexpansive sequences in a Banach space. J. Fixed Point Theory and Appl. **21**, Article no. 35 (2019).

[3] Barbu, V., *Nonlinear semigroups and differential equations in Banach spaces*, Noordhoff International Publishing, Leyden, The Netherlands, 1976.

[4] Bauschke, H.H., Combettes, P.L., *Convex analysis and monotone operator theory in Hilbert spaces*, Springer, New York, 2nd edn., 2017.

[5] Bauschke, H.H., Moursi, W.A, Wang, X., Generalized monotone operators and their averaged resolvents. Math. Programming Ser. B, https://doi.org/10.1007/s10107-020-01500-6 (2020).

[6] Bruck, R.E., Reich, S., Nonexpansive projections and resolvents of accretive operators in Banach spaces. Houston J. Math. **3**, pp. 459-470 (1977).

[7] Combettes, P.L., Pennanen, T., Proximal methods for cohypomonotone operators. SIAM J. Control Optim. **43**, pp. 731-742 (2004).

[8] Diakonikolas, J., Daskalakis, C., Jordan, M.I., Efficient methods for structured nonconvexnon-concave Min-Max optimization. arXiv:2011.00364.

[9] Güler, O., On the convergence of the proximal point algorithm for convex minimization. SIAM J. Control Optim. **29**, pp. 403-419 (1991).

[10] Kohlenbach, U., On the quantitative asymptotic behavior of strongly nonexpansive mappings in Banach and geodesic spaces. Israel J. Math. **216**, pp. 215-246 (2016).

[11] Kohlenbach, U., On the reverse mathematics and Weihrauch complexity of moduli of regularity and uniqueness. Computability **8**, pp. 377-387 (2019).

[12] Kohlenbach, U., Quantitative analysis of a Halpern-type proximal point algorithm for accretive operators in Banach spaces. J. Nonlin. Convex Anal. **9**, pp. 2125-2138 (2020).

[13] Kohlenbach, U., Quantitative results on the proximal point algorithm in uniformly convex Banach spaces. J. Convex Anal. **28**, pp. 11-18 (2021).

[14] Kohlenbach, U., Leuştean, L., Nicolae, A., Quantitative results on Fejér monotone sequences. Comm. Contemp. Math. vol. 20 (2018), 42pp., DOI: 10.1142/S0219199717500158.

[15] Kohlenbach, U., López-Acedo, G., Nicolae, A., Quantitative asymptotic regularity results for the composition of two mappings. Optimization **66**, pp. 1291-1299 (2017).

[16] Kohlenbach, U., López-Acedo, G., Nicolae, A., Moduli of regularity and rates of convergence for Fejér monotone sequences. Israel J. Math. **232**, pp. 261-297 (2019).

[17] Martinet, B., Régularisation d'inéquations variationnelles par approximations successives. Rev. Française Informat. Recherche Opérationnelle **4**, pp. 154-158 (1970).

[18] Nakajo, K., Shimoji, K., Takahashi, W., Strong convergence to common fixed points of families of nonexpansive mappings in Banach spaces. J. Nonlinear Convex Anal. **8**, pp-11-34 (2007).

[19] Neumann, E., Computational problems in metric fixed point theory and their Weihrauch degrees. Log. Method. Comput. Sci. **11**, 44pp. (2015).

[20] Rockafellar, R.T. Monotone operators and the proximal point algorithm. SIAM J. Control Optim. **14**, pp. 877-898 (1976).

[21] Sipoş, A., Quantitative inconsistent feasibility for averaged mappings, arXiv:2001.01513, 2020.

[22] Takahashi, W., Nonlinear Functional Analysis. Yokohama Publishers, iv+276pp., 2000.

[23] Takahashi, W., Viscosity approximation methods for countable families of nonexpansive mappings in Banach spaces. Nonlinear Anal. **70**, pp. 719-734 (2009).

[24] Tao, T., Soft analysis, hard analysis, and the finite convergence principle. Essay posted May 23, 2007. Appeared in: 'T. Tao, Structure and Randomness: Pages from Year One of a Mathematical Blog. AMS, 298pp., 2008'.

[25] Tao, T., Norm convergence of multiple ergodic averages for commuting transformations. Ergod. Th. & Dyn. Sys. **28**, pp. 657-688 (2008).