Journal of Optimization Theory and Applications (2020) 186:759–778
https://doi.org/10.1007/s10957-020-01734-6

Demiclosedness Principles for Generalized Nonexpansive Mappings

Sedi Bartz[1] · **Rubén Campoy**[1] · **Hung M. Phan**[1]

Received: 29 April 2020 / Accepted: 28 July 2020 / Published online: 12 August 2020

Abstract
Demiclosedness principles are powerful tools in the study of convergence of iterative methods. For instance, a multi-operator demiclosedness principle for firmly nonexpansive mappings is useful in obtaining simple and transparent arguments for the weak convergence of the shadow sequence generated by the Douglas–Rachford algorithm. We provide extensions of this principle, which are compatible with the framework of more general families of mappings such as cocoercive and conically averaged mappings. As an application, we derive the weak convergence of the shadow sequence generated by the adaptive Douglas–Rachford algorithm.

Keywords Demiclosedness principle · Cocoercive mapping · Conically averaged mapping · Weak convergence · Douglas–Rachford algorithm · Adaptive Douglas–Rachford algorithm

Mathematics Subject Classification 47H05 · 47J25 · 49M27

1 Introduction

Demiclosedness principles play an important role in convergence analysis of fixed point algorithms. The concept of *demiclosedness* sheds light on topological properties

---

Communicated by Michel Théra.

✉ Rubén Campoy
ruben_campoygarcia@uml.edu

Sedi Bartz
sedi_bartz@uml.edu

Hung M. Phan
hung_phan@uml.edu

1 Department of Mathematical Sciences, Kennedy College of Sciences, University of Massachusetts Lowell, Lowell, MA, USA

of mappings, in particular, in the case where a weak topology is considered. More precisely, given a weakly sequentially closed subset D of a Hilbert space $\mathcal{H}$, the mapping $T : D \to \mathcal{H}$ is said to be *demiclosed* at $x \in D$, if for every sequence (x_k) in D such that (x_k) converges weakly to x and $T(x_k)$ converges strongly, say, to u, it follows that $T(x) = u$. By its definition, demiclosedness holds trivially whenever T is weakly sequentially continuous; however, it does not hold in general. Let Id denote the identity mapping on $\mathcal{H}$. A fundamental result in the theory of nonexpansive mappings is Browder's celebrated *demiclosedness principle* [1], which asserts that, if T is nonexpansive, then the mapping $\mathrm{Id} - T$ is demiclosed at every point in D. Browder's result holds in more general settings and, by now, has become a key tool in the study of asymptotic and ergodic properties of nonexpansive mappings; see [2–5], for example.

In [6], Browder's demiclosedness principle was extended and a version for finitely many firmly nonexpansive mappings was provided. As an application, a simple proof of the weak convergence of the *Douglas–Rachford (DR) algorithm* [7,8] was also provided in [6]: The DR algorithm belongs to the class of splitting methods for the problem of finding a *zero of the sum* of two maximally monotone operators $A, B : \mathcal{H} \rightrightarrows \mathcal{H}$; see (25). The DR algorithm generates a sequence by an iterative application of the DR operator (see (26), (27) and the comment thereafter), which can be expressed in terms of the resolvents (see Definition 2.3) of A and B. If the solution set is nonempty, then the DR sequence converges weakly to a fixed point such that the resolvent of A maps it to a zero of $A+B$. Thus, we see that, in fact, we are interested in the image of the DR sequence under the resolvent of A. This image is often referred to as the *shadow sequence*. The resolvent of a maximally monotone operator is continuous (in fact, firmly nonexpansive), but not weakly continuous, in general. Hence, the convergence of the shadow sequence cannot be derived directly from the convergence of the DR sequence, unless the latter converges in norm. However, in general, norm convergence does not hold: In [9], an example of a DR iteration which does not converge in norm was explicitly constructed. Regardless of this fact, the weak convergence of the shadow sequence was established by Svaiter in [10]. A simpler and more accessible proof of the weak convergence of the shadow sequence was later given in [6] by employing a multi-operator demiclosedness principle. A demiclosedness principle for circumcenter mappings, a class of operators that is generally not continuous, was recently developed in [11].

In this paper, we present an extended demiclosedness principle for more general families of operators, which are not necessarily firmly nonexpansive, provided that they satisfy a (firm) nonexpansiveness balance condition. We are motivated by the *adaptive Douglas–Rachford* (*aDR*) *algorithm* which was recently studied in [12] in order to find a zero of the sum of a weakly monotone operator and a strongly monotone operator. Furthermore, the framework of [12] has been recently extended in [13] in order to hold for monotonicity and comonotonicity settings as well. In both studies [12,13], the convergence of the shadow sequence generated by the aDR is guaranteed only under the assumption that the sum of the operators is strongly monotone. Moreover, the corresponding resolvents in the aDR are not necessarily firmly nonexpansive, and consequently, the demiclosedness principles of [6] cannot be directly applied in this framework. Our current approach is compatible with the framework of the aDR.

Consequently, we employ our generalized demiclosedness principles in order to obtain weak convergence of the shadow sequence of the aDR in most cases. To this end, we employ and extend techniques and results from [6].

The remainder of this paper is organized as follows: In Sect. 2, we review preliminaries and basic results. New demiclosedness principles are provided in Sect. 3. In Sect. 4, we employ the demiclosedness principles from Sect. 3 in order to obtain the weak convergence of the shadow sequence of the adaptive Douglas–Rachford algorithm. Finally, in Sect. 5, we conclude our discussion.

2 Preliminaries

Throughout this paper, $\mathcal{H}$ is a real Hilbert space equipped with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. The weak convergence and strong convergence are denoted by $\rightharpoonup$ and $\to$, respectively. We set $\mathbb{R}_+ := \{r \in \mathbb{R} : r \geq 0\}$ and $\mathbb{R}_{++} := \{r \in \mathbb{R} : r > 0\}$. Given a set-valued operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$, the *graph*, the *domain*, the set of *fixed points* and the set of *zeros* of A, are denoted, respectively, by $\operatorname{gra} A$, $\operatorname{dom} A$, $\operatorname{Fix} A$ and $\operatorname{zer} A$; i.e.,

$$\operatorname{gra} A := \big\{(x,u) \in \mathcal{H} \times \mathcal{H} : u \in A(x)\big\}, \qquad \operatorname{dom} A := \big\{x \in \mathcal{H} : A(x) \neq \varnothing\big\},$$
$$\operatorname{Fix} A := \big\{x \in \mathcal{H} : x \in A(x)\big\} \qquad \text{and} \quad \operatorname{zer} A := \big\{x \in \mathcal{H} : 0 \in A(x)\big\}.$$

The *identity* mapping is denoted by Id and the *inverse* of A is denoted by A^{-1}, i.e., $\operatorname{gra} A^{-1} := \{(u,x) \in \mathcal{H} \times \mathcal{H} : u \in A(x)\}$.

Definition 2.1 Let $D \subseteq \mathcal{H}$ be nonempty, let $T : D \to \mathcal{H}$ be a mapping, set $\tau > 0$ and $\theta > 0$. The mapping T is said to be

(i) *nonexpansive*, if

$$\big\|T(x) - T(y)\big\| \leq \|x - y\|, \quad \forall x, y \in D;$$

(ii) *firmly nonexpansive*, if

$$\big\|T(x) - T(y)\big\|^2 + \big\|(\operatorname{Id} - T)(x) - (\operatorname{Id} - T)(y)\big\|^2 \leq \|x - y\|^2, \quad \forall x, y \in D,$$

equivalently,

$$\big\langle x - y, T(x) - T(y)\big\rangle \geq \big\|T(x) - T(y)\big\|^2, \quad \forall x, y \in D;$$

(iii) τ-*cocoercive*, if τT is firmly nonexpansive, i.e.,

$$\big\langle x - y, T(x) - T(y)\big\rangle \geq \tau \big\|T(x) - T(y)\big\|^2, \quad \forall x, y \in D;$$

(iv) *conically* θ-*averaged*, if there exists a nonexpansive operator $R : D \to \mathcal{H}$ such that

$$T = (1 - \theta) \operatorname{Id} + \theta R.$$

Conically θ-averaged mappings, introduced in [14] and originally named *conically nonexpansive mappings*, are natural extensions of the classical θ-averaged mappings; more precisely, a conically θ-averaged mapping is θ-averaged whenever $\theta \in]0, 1[$. Additional properties and further discussions of conically averaged mappings, such as the following result, are available in [13,15].

Fact 2.1 *Let $D \subseteq \mathcal{H}$ be nonempty, let $T : D \to \mathcal{H}$ and let $\theta, \sigma > 0$. Then the following assertions are equivalent:*

(i) *T is conically θ-averaged;*
(ii) *$(1-\sigma)\,\mathrm{Id} + \sigma T$ is conically $\sigma\theta$-averaged;*
(iii) *For all $x, y \in D$,*

$$\left\|T(x) - T(y)\right\|^2 \le \|x-y\|^2 - \frac{1-\theta}{\theta}\left\|(\mathrm{Id} - T)(x) - (\mathrm{Id} - T)(y)\right\|^2.$$

Proof See [13, Proposition 2.2]. □

Lemma 2.1 *Let $D \subseteq \mathcal{H}$ be nonempty, let $T : D \to \mathcal{H}$ and let $\tau, \theta > 0$.*

(i) *If T is τ-cocoercive, then it is τ'-cocoercive for any $\tau' \in]0, \tau]$.*
(ii) *If T is conically θ-averaged, then it is conically θ'-averaged for any $\theta' \in [\theta, \infty[$.*

Proof (i): Follows immediately from the definition of cocoercivity. (ii): Follows from the equivalence (i) $\iff$ (iii) in Fact 2.1. □

Remark 2.1 We note that cocoercivity and conical averagedness generalize the notion of firm nonexpansiveness as follows:

(i) By Definition 2.1(ii) and (iii), the mapping T is firmly nonexpansive if and only if T is 1-cocoercive. Consequently, by employing Lemma 2.1(i), we see that whenever $\tau \ge 1$, a τ-cocoercive mapping is firmly nonexpansive.
(ii) Similarly, the mapping T is firmly nonexpansive if and only if it is conically $\frac{1}{2}$-averaged. Consequently, by employing Lemma 2.1(ii), we see that whenever $\theta \le \frac{1}{2}$, a conically θ-averaged mapping is firmly nonexpansive.

In our study, we will employ the following generalized notions of monotonicity.

Definition 2.2 Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ and let $\alpha \in \mathbb{R}$. Then A is said to be

(i) *α-monotone*, if

$$\langle x-y, u-v\rangle \ge \alpha\|x-y\|^2, \quad \forall (x,u), (y,v) \in \operatorname{gra} A;$$

(ii) *α-comonotone*, if A^{-1} is α-monotone, i.e.,

$$\langle x-y, u-v\rangle \ge \alpha\|u-v\|^2, \quad \forall (x,u), (y,v) \in \operatorname{gra} A.$$

The α-monotone operator A is said to be *maximally α-monotone* (resp. *maximally α-comonotone*), if there is no α-monotone (resp. α-comonotone) operator $B : \mathcal{H} \rightrightarrows \mathcal{H}$ such that $\operatorname{gra} A$ is properly contained in $\operatorname{gra} B$.

Remark 2.2 Common notions of monotonicity are related to the notions in Definition 2.2 as follows:

- In the case where $\alpha = 0$, 0-monotonicity and 0-comonotonicity simply mean *monotonicity* (see, for example, [16, Definition 20.1]).
- In the case where $\alpha > 0$, α-monotonicity is also known as α-*strong monotonicity* (see, for example, [16, Definition 22.1(iv)]). Similarly, α-comonotonicity is α-cocoercivity in Definition 2.1(iii).
- In the case where $\alpha < 0$, α-monotonicity and α-comonotonicity are also known as α-*hypomonotonicity* and α-*cohypomonotonicity*, respectively (see, for example, [17, Definition 2.2]). In addition, α-monotonicity is referred to as α-*weak monotonicity* in [12].

We continue our preliminary discussion by recalling the definition of the resolvent.

Definition 2.3 Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$. The *resolvent* of A is the operator defined by

$$J_A := (\mathrm{Id} + A)^{-1}.$$

The *relaxed resolvent* of A with parameter $\lambda > 0$ is defined by

$$J_A^{\lambda} := (1 - \lambda)\,\mathrm{Id} + \lambda J_A.$$

When $\lambda = 2$, we set $R_A := J_A^2 = 2J_A - \mathrm{Id}$, also known as the *reflected resolvent* of A.

We conclude our preliminary discussion by relating monotonicity and comonotonicity properties with corresponding properties for resolvents by recalling the following facts.

Fact 2.2 (*resolvents of monotone operators*) *Let* $A : \mathcal{H} \rightrightarrows \mathcal{H}$ *be* α*-monotone, where* $\alpha \in \mathbb{R}$*. If* $\gamma > 0$ *is such that* $1 + \gamma\alpha > 0$*, then*

(i) $J_{\gamma A}$ *is single-valued and* $(1 + \gamma\alpha)$*-cocoercive;*
(ii) $\operatorname{dom} J_{\gamma A} = \mathcal{H}$ *if and only if* A *is maximally* α*-monotone.*

Proof See [12, Lemma 3.3(ii) and Proposition 3.4]. □

Fact 2.3 (*resolvents of comonotone operators*) *Let* $A : \mathcal{H} \rightrightarrows \mathcal{H}$ *be* α*-comonotone, where* $\alpha \in \mathbb{R}$*. If* $\gamma > 0$ *is such that* $\gamma + \alpha > 0$*, then*

(i) $J_{\gamma A}$ *is at most single-valued and conically* $\frac{\gamma}{2(\gamma+\alpha)}$*-averaged;*
(ii) $\operatorname{dom} J_{\gamma A} = \mathcal{H}$ *if and only if* A *is maximally* α*-comonotone.*

Proof See [13, Propositions 3.7 and 3.8(i)]. □

3 New Demiclosedness Principles for Generalized Nonexpansive Mappings

In this section, we extend the multi-operator demiclosedness principles for firmly nonexpansive mappings from [6] to more general families of operators. To this end, we employ and extend techniques and results from [6] in a weighted inner product space. We begin our discussion by recalling the following fact.

Fact 3.1 *Let $F : \mathcal{H} \to \mathcal{H}$ be a firmly nonexpansive mapping, let $(x_k)_{k=0}^{\infty}$ be a sequence in $\mathcal{H}$, and let $C, D \subseteq \mathcal{H}$ be closed affine subspaces such that $C - C = (D - D)^{\perp}$. Suppose that*

$$x_k \rightharpoonup x, \tag{1a}$$
$$F(x_k) \rightharpoonup y, \tag{1b}$$
$$F(x_k) - P_C(F(x_k)) \to 0, \tag{1c}$$
$$(x_k - F(x_k)) - P_D(x_k - F(x_k)) \to 0. \tag{1d}$$

Then $y \in C$, $x \in y + D$, and $y = F(x)$.

Proof See [6, Corollary 2.7]. □

3.1 Demiclosedness Principles for Cocoercive Operators

Let $n \geq 2$ be an integer and set $\boldsymbol{\omega} := (\omega_1, \omega_2, \ldots, \omega_n) \in \mathbb{R}^n_{++}$. We equip the space

$$\boldsymbol{\mathcal{H}} := \mathcal{H}^n = \underbrace{\mathcal{H} \times \mathcal{H} \times \cdots \times \mathcal{H}}_{n}, \tag{2a}$$

with the weighted inner product $\langle \cdot, \cdot \rangle_{\boldsymbol{\omega}}$ defined by

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\boldsymbol{\omega}} := \sum_{i=1}^{n} \omega_i \langle x_i, y_i \rangle, \quad \forall \boldsymbol{x} = (x_1, x_2, \ldots, x_n), \boldsymbol{y} = (y_1, y_2, \ldots, y_n) \in \boldsymbol{\mathcal{H}}. \tag{2b}$$

Thus, $\boldsymbol{\mathcal{H}}$ is a Hilbert space with the induced norm $\|\boldsymbol{x}\|_{\boldsymbol{\omega}} = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle_{\boldsymbol{\omega}}}$.

Let $\boldsymbol{\tau} := (\tau_1, \tau_2, \ldots, \tau_n) \in \mathbb{R}^n \smallsetminus \big\{(0, \ldots, 0)\big\}$. Let $\boldsymbol{C} \subset \boldsymbol{\mathcal{H}}$ be the subspace defined by

$$\boldsymbol{C} := \big\{(\tau_1 x, \tau_2 x, \ldots, \tau_n x) : x \in \mathcal{H}\big\}.$$

In the following lemma, we provide a formula for the projector $P_{\boldsymbol{C}}$, which will be useful later.

Lemma 3.1 *Let $\boldsymbol{x} := (x_1, x_2, \ldots, x_n) \in \boldsymbol{\mathcal{H}}$. Then the projection of $\boldsymbol{x}$ onto $\boldsymbol{C}$ is*

$$P_{\boldsymbol{C}}(\boldsymbol{x}) = (\tau_1 \overline{u}, \tau_2 \overline{u}, \ldots, \tau_n \overline{u}), \quad \text{where } \overline{u} := \frac{\sum_{i=1}^{n} \omega_i \tau_i x_i}{\sum_{i=1}^{n} \omega_i \tau_i^2}. \tag{3}$$

Proof Fix $\boldsymbol{x} := (x_1, x_2, \ldots, x_n) \in \boldsymbol{\mathcal{H}}$. Since $\sum_{i=1}^n \omega_i \tau_i^2 > 0$, $\overline{u}$ is well defined by (3). Set $\boldsymbol{y} = (\tau_1\overline{u}, \tau_2\overline{u}, \ldots, \tau_n\overline{u})$. We prove that $(\boldsymbol{x}-\boldsymbol{y}) \perp \boldsymbol{z}$ for each $\boldsymbol{z} \in \boldsymbol{C}$; consequently, $P_{\boldsymbol{C}}(\boldsymbol{x}) = \boldsymbol{y}$. Indeed, let $\boldsymbol{z} = (\tau_1 v, \tau_2 v, \ldots, \tau_n v) \in \boldsymbol{C}$. Then

$$\begin{aligned}\langle \boldsymbol{z}, \boldsymbol{x}-\boldsymbol{y}\rangle_{\boldsymbol{\omega}} &= \sum_{i=1}^n \omega_i \langle \tau_i v, x_i - \tau_i \overline{u}\rangle \\ &= \left\langle v, \sum_{i=1}^n \omega_i \tau_i (x_i - \tau_i\overline{u})\right\rangle \\ &= \left\langle v, \sum_{i=1}^n \omega_i \tau_i x_i - \left(\sum_{i=1}^n \omega_i \tau_i^2\right)\overline{u}\right\rangle = \langle v, 0\rangle = 0.\end{aligned}$$

□

Theorem 3.1 (demiclosedness principle for cocoercive operators) *Let* $(\rho_1, \rho_2, \ldots, \rho_n), (\tau_1, \tau_2, \ldots, \tau_n) \in \mathbb{R}^n_{++}$. *For each* $i \in \{1, 2, \ldots, n\}$, *let* $F_i : \mathcal{H} \to \mathcal{H}$ *be* τ_i*-cocoercive and let* $(x_{i,k})_{k=0}^\infty$ *be a sequence in* $\mathcal{H}$. *Suppose that*

$$\forall i \in \{1, 2, \ldots, n\}, \qquad x_{i,k} \rightharpoonup x_i, \tag{4a}$$

$$\forall i \in \{1, 2, \ldots, n\}, \qquad F_i(x_{i,k}) \rightharpoonup y, \tag{4b}$$

$$\sum_{i=1}^n \rho_i (x_{i,k} - \tau_i F_i(x_{i,k})) \to -\left(\sum_{i=1}^n \rho_i \tau_i\right) y + \sum_{i=1}^n \rho_i x_i, \tag{4c}$$

$$\forall i, j \in \{1, 2, \ldots, n\}, \qquad F_i(x_{i,k}) - F_j(x_{j,k}) \to 0. \tag{4d}$$

Then $F_1(x_1) = F_2(x_2) = \cdots = F_n(x_n) = y$.

Proof For each $i \in \{1, 2, \ldots, n\}$, set

$$\omega_i := \frac{\rho_i}{\tau_i}$$

and equip $\boldsymbol{\mathcal{H}}$ with the inner product $\langle\cdot,\cdot\rangle_{\boldsymbol{\omega}}$ as in (2). Let $\boldsymbol{F} : \boldsymbol{\mathcal{H}} \to \boldsymbol{\mathcal{H}}$ be the mapping defined by

$$\boldsymbol{F}(\boldsymbol{z}) = (\tau_1 F_1(z_1), \tau_2 F_2(z_2), \ldots, \tau_n F_n(z_n)), \quad \boldsymbol{z} = (z_1, z_2, \ldots, z_n) \in \boldsymbol{\mathcal{H}}.$$

Then for every $\boldsymbol{u} = (u_1, u_2, \ldots, u_n), \boldsymbol{v} = (v_1, v_2, \ldots, v_n) \in \boldsymbol{\mathcal{H}}$, since F_i is τ_i cocoercive, it follows that

$$\begin{aligned}\langle \boldsymbol{u}-\boldsymbol{v}, \boldsymbol{F}(\boldsymbol{u}) - \boldsymbol{F}(\boldsymbol{v})\rangle_{\boldsymbol{\omega}} &= \sum_{i=1}^n \omega_i \langle u_i - v_i, \tau_i F_i(u_i) - \tau_i F_i(v_i)\rangle \\ &\geq \sum_{i=1}^n \omega_i \tau_i^2 \|F_i(u_i) - F_i(v_i)\|^2 = \|\boldsymbol{F}(\boldsymbol{u}) - \boldsymbol{F}(\boldsymbol{v})\|^2_{\boldsymbol{\omega}},\end{aligned}$$

which implies that $\boldsymbol{F}$ is firmly nonexpansive. Set $\boldsymbol{x} := (x_1, x_2, \ldots, x_n)$, $\boldsymbol{y} := (\tau_1 y, \tau_2 y, \ldots, \tau_n y)$ and, for each $k = 0, 1, 2, \ldots$, set $\boldsymbol{x}_k := (x_{1,k}, x_{2,k}, \ldots, x_{n,k})$. Then

$$\boldsymbol{x}_k \rightharpoonup \boldsymbol{x} \quad \text{and} \quad \boldsymbol{F}(\boldsymbol{x}_k) \rightharpoonup \boldsymbol{y}. \tag{5}$$

Let $\boldsymbol{C}$ and $\boldsymbol{D}$ be the affine subspaces of $\boldsymbol{\mathcal{H}}$ defined by

$$\boldsymbol{C} := \{(\tau_1 x, \tau_2 x, \ldots, \tau_n x) : x \in \mathcal{H}\} \quad \text{and} \quad \boldsymbol{D} := \boldsymbol{x} - \boldsymbol{y} + \boldsymbol{C}^{\perp}.$$

Then $\boldsymbol{C} - \boldsymbol{C} = (\boldsymbol{D} - \boldsymbol{D})^{\perp}$. Consequently, by employing Lemma 3.1, we arrive at

$$P_{\boldsymbol{C}}(\boldsymbol{F}(\boldsymbol{x}_k)) = (\tau_1 \overline{v}_k, \ldots, \tau_n \overline{v}_k), \quad \text{where } \overline{v}_k := \frac{\sum_{i=1}^n \omega_i \tau_i^2 F_i(x_{i,k})}{\sum_{i=1}^n \omega_i \tau_i^2} = \frac{\sum_{i=1}^n \rho_i \tau_i F_i(x_{i,k})}{\sum_{i=1}^n \rho_i \tau_i}.$$

Since $\overline{v}_k$ is a weighted average of the $F_i(x_{i,k})$'s, (4d) implies that

$$\forall i \in \{1, 2, \ldots, n\}, \quad F_i(x_{i,k}) - \overline{v}_k \to 0 \quad \text{as} \quad k \to \infty;$$

consequently, we conclude that

$$\boldsymbol{F}(\boldsymbol{x}_k) - P_{\boldsymbol{C}}(\boldsymbol{F}(\boldsymbol{x}_k)) \to 0. \tag{6}$$

We now employ the projections

$$P_{\boldsymbol{C}}(\boldsymbol{x}_k - \boldsymbol{F}(\boldsymbol{x}_k)) = (\tau_1 \overline{u}_k, \ldots, \tau_n \overline{u}_k), \quad \text{where } \overline{u}_k = \frac{\sum_{i=1}^n \rho_i \left(x_{i,k} - \tau_i F_i(x_{i,k})\right)}{\sum_{i=1}^n \rho_i \tau_i},$$

and

$$P_{\boldsymbol{C}}(\boldsymbol{x} - \boldsymbol{y}) = (\tau_1 \overline{u}, \ldots, \tau_n \overline{u}) - \boldsymbol{y}, \quad \text{where } \overline{u} = \frac{\sum_{i=1}^n \rho_i x_i}{\sum_i^n \rho_i \tau_i}.$$

By invoking (4c), we see that $\overline{u}_k \to -y + \overline{u}$, which, in turn, implies that

$$P_{\boldsymbol{C}}(\boldsymbol{x}_k - \boldsymbol{F}(\boldsymbol{x}_k)) \to P_{\boldsymbol{C}}(\boldsymbol{x} - \boldsymbol{y}).$$

Consequently,

$$\begin{aligned} \boldsymbol{x}_k - \boldsymbol{F}(\boldsymbol{x}_k) - &P_{\boldsymbol{D}}(\boldsymbol{x}_k - \boldsymbol{F}(\boldsymbol{x}_k)) \\ &= \boldsymbol{x}_k - \boldsymbol{F}(\boldsymbol{x}_k) - P_{\boldsymbol{x}-\boldsymbol{y}+\boldsymbol{C}^{\perp}}(\boldsymbol{x}_k - \boldsymbol{F}(\boldsymbol{x}_k)) \\ &= \boldsymbol{x}_k - \boldsymbol{F}(\boldsymbol{x}_k) - \left(\boldsymbol{x} - \boldsymbol{y} + P_{\boldsymbol{C}^{\perp}}(\boldsymbol{x}_k - \boldsymbol{F}(\boldsymbol{x}_k) - (\boldsymbol{x} - \boldsymbol{y}))\right) \\ &= (\operatorname{Id} - P_{\boldsymbol{C}^{\perp}})(\boldsymbol{x}_k - \boldsymbol{F}(\boldsymbol{x}_k)) - (\operatorname{Id} - P_{\boldsymbol{C}^{\perp}})(\boldsymbol{x} - \boldsymbol{y}) \\ &= P_{\boldsymbol{C}}(\boldsymbol{x}_k - \boldsymbol{F}(\boldsymbol{x}_k)) - P_{\boldsymbol{C}}(\boldsymbol{x} - \boldsymbol{y}) \to 0. \end{aligned} \tag{7}$$

Finally, since (5), (6) and (7) satisfy (1), we may employ Fact 3.1 in order to obtain $\boldsymbol{y} = \boldsymbol{F}(\boldsymbol{x})$, that is,

$$F_i(x_i) = y, \quad \forall i \in \{1, 2, \ldots, n\},$$

which concludes the proof. □

As a consequence of Theorem 3.1, we obtain the demiclosedness principle for firmly nonexpansive operators [6, Theorem 2.10].

Corollary 3.1 (demiclosedness principle for firmly nonexpansive operators) *For each $i \in \{1, 2, \ldots, n\}$ let $F_i : \mathcal{H} \to \mathcal{H}$ be firmly nonexpansive and let $(x_{i,k})_{k=0}^{\infty}$ be a sequence in $\mathcal{H}$. Suppose further that*

$$\forall i \in \{1, 2, \ldots, n\}, \qquad x_{i,k} \rightharpoonup x_i, \tag{8a}$$

$$\forall i \in \{1, 2, \ldots, n\}, \qquad F_i(x_{i,k}) \rightharpoonup y, \tag{8b}$$

$$\sum_{i=1}^{n} (x_{i,k} - F_i(x_{i,k})) \to -ny + \sum_{i=1}^{n} x_i, \tag{8c}$$

$$\forall i, j \in \{1, 2, \ldots, n\}, \qquad F_i(x_{i,k}) - F_j(x_{j,k}) \to 0. \tag{8d}$$

Then $F_1(x_1) = F_2(x_2) = \cdots = F_n(x_n) = y$.

Proof The proof follows by observing that firmly nonexpansive operators are cocoercive with constant $\tau = 1$ and by setting $\tau_1 = \tau_2 = \cdots = \tau_n = 1$ and $\rho_1 = \rho_2 = \cdots = \rho_n = 1$ in Theorem 3.1. □

Remark 3.1 By letting $n = 1$ in Corollary 3.1, we obtain a special case of Fact 3.1 where $C = \mathcal{H}$ and $D = \{x - y\}$, which, in turn, is equivalent to Browder's original demiclosedness principle [1] (alternatively, see [16, Theorem 4.27]).

Remark 3.2 (Theorem 3.1 vs. Corollary 3.1) A demiclosedness principle for cocoercive mappings can be derived directly from Corollary 3.1 when we apply the latter to the firmly nonexpansive mappings $\tau_1 F_1, \tau_2 F_2, \ldots, \tau_n F_n$. However, this does not yield Theorem 3.1: In this case, the cocoercivity constants will appear in (8b) and (8d); however, they are neither a part of (4b) nor (4d).

Following Remark 3.2, in Theorem 3.1, the cocoercivity constants τ_i are not a part of the conditions in (4) except for the condition (4c). In the following result, we do not incorporate cocoercivity constants in any of the convergence conditions; however, in exchange, we do impose a certain balance condition on these constants.

Theorem 3.2 (closedness principle for balanced cocoercive operators) *For each $i \in \{1, 2, \ldots, n\}$, let $F_i : \mathcal{H} \to \mathcal{H}$ be a τ_i-cocoercive mapping where $\tau_i > 0$ and let $(x_{i,k})_{k=0}^{\infty}$ be a sequence in $\mathcal{H}$. Suppose that there exists $(\rho_1, \rho_2, \ldots, \rho_n) \in \mathbb{R}_{++}^n$ such that the weighted average*

$$\frac{\sum_{i=1}^{n} \rho_i \tau_i}{\sum_{i=1}^{n} \rho_i} \geq 1, \tag{9}$$

and suppose further that

$$\forall i \in \{1, 2, \ldots, n\}, \qquad x_{i,k} \rightharpoonup x_i, \tag{10a}$$

$$\forall i \in \{1, 2, \ldots, n\}, \qquad F_i(x_{i,k}) \rightharpoonup y, \tag{10b}$$

$$\sum_{i=1}^{n} \rho_i (x_{i,k} - F_i(x_{i,k})) \to -\Big(\sum_{i=1}^{n} \rho_i\Big) y + \sum_{i=1}^{n} \rho_i x_i, \tag{10c}$$

$$\forall i, j \in \{1, 2, \ldots, n\}, \qquad F_i(x_{i,k}) - F_j(x_{j,k}) \to 0. \tag{10d}$$

Then $F_1(x_1) = F_2(x_2) = \cdots = F_n(x_n) = y$.

Proof In view of (9), we may choose $(\tau_1', \tau_2', \ldots, \tau_n')$, $\tau_i' \in]0, \tau_i]$ such that

$$\frac{\sum_{i=1}^{n} \rho_i \tau_i'}{\sum_{i=1}^{n} \rho_i} = 1.$$

By invoking Lemma 2.1(i), we see that for each $i \in \{1, 2, \ldots, n\}$, F_i is τ_i'-cocoercive. Consequently, we may assume without loss of generality that there is equality in (9), which we rewrite in the form

$$\rho_1(\tau_1 - 1) + \rho_2(\tau_2 - 1) + \cdots + \rho_n(\tau_n - 1) = 0. \tag{11}$$

Conditions (10a), (10b) and (10d) are the same as (4a), (4b) and (4d), respectively. Thus, in order to employ Theorem 3.1 3.1 and complete the proof, it suffices to prove that (4c) holds. Indeed, by combining (10c), (10d) and (11), we arrive at

$$\begin{aligned}\sum_{i=1}^{n} \rho_i \big(x_{i,k} - \tau_i F_i(x_{i,k})\big) &= \sum_{i=1}^{n} \rho_i \big(x_{i,k} - F_i(x_{i,k})\big) + \sum_{i=1}^{n} \rho_i (1 - \tau_i) F_i(x_{i,k}) \\ &= \sum_{i=1}^{n} \rho_i \big(x_{i,k} - F_i(x_{i,k})\big) + \sum_{i=2}^{n} \rho_i (1 - \tau_i) \big(F_i(x_{i,k}) - F_1(x_{1,k})\big) \\ &\to -\Big(\sum_{i=1}^{n} \rho_i\Big) y + \sum_{i=1}^{n} \rho_i x_i = -\Big(\sum_{i=1}^{n} \rho_i \tau_i\Big) y + \sum_{i=1}^{n} \rho_i x_i,\end{aligned}$$

which is (4c). □

Remark 3.3 (on the balance condition (9)) We note that the conditions in (10) for cocoercive mappings are a weighted version of the conditions in (8) for firmly nonexpansive mappings. However, the cocoercivity constants are required to be balanced as in (9); that is, the weighted average of the cocoercivity constants has to be at least 1, which is always true for firmly nonexpansive mappings (see Remark 2.1(i)).

3.2 Demiclosedness Principles for Conically Averaged Operators

In this section, we provide a demiclosedness principle for finitely many conically averaged operators. This is yet another generalization of the demiclosedness principle for firmly nonexpansive operators (Corollary 3.1), which we employ in our proof.

Theorem 3.3 (demiclosedness principle for conically averaged operators) *For each $i \in \{1, 2, \ldots, n\}$, let $T_i : \mathcal{H} \to \mathcal{H}$ be conically θ_i-averaged where $\theta_i > 0$, and let $\left(x_{i,k}\right)_{k=0}^{\infty}$ be a sequence in $\mathcal{H}$. Suppose that*

$$\forall i \in \{1, \ldots, n\}, \quad x_{i,k} \rightharpoonup x_i, \tag{12a}$$

$$\forall i \in \{1, \ldots, n\}, \quad T_i(x_{i,k}) \rightharpoonup 2\theta_i y + (1 - 2\theta_i)x_i, \tag{12b}$$

$$\sum_{i=1}^{n} \frac{x_{i,k} - T_i(x_{i,k})}{2\theta_i} \to -ny + \sum_{i=1}^{n} x_i, \tag{12c}$$

$$\forall i, j \in \{1, 2, \ldots, n\}, \ (x_{i,k} - x_{j,k}) - \left(\frac{x_{i,k} - T_i(x_{i,k})}{2\theta_i} - \frac{x_{j,k} - T_j(x_{j,k})}{2\theta_j} \right) \to 0. \tag{12d}$$

Then $T_i(x_i) = 2\theta_i y + (1 - 2\theta_i)x_i$ for all $i \in \{1, \ldots, n\}$.

Proof For each $i \in \{1, \ldots, n\}$, set

$$F_i := \left(1 - \frac{1}{2\theta_i}\right) \mathrm{Id} + \frac{1}{2\theta_i} T_i = \mathrm{Id} - \left(\frac{\mathrm{Id} - T_i}{2\theta_i}\right).$$

Then Fact 2.1(ii) and Remark 2.1(ii) imply that F_i is firmly nonexpansive. By employing (12a) and (12b), we see that

$$F_i(x_{i,k}) = \left(1 - \frac{1}{2\theta_i}\right) x_{i,k} + \frac{1}{2\theta_i} T_i(x_{i,k}) \rightharpoonup \left(1 - \frac{1}{2\theta_i}\right) x_i + y + \frac{1 - 2\theta_i}{2\theta_i} x_i = y. \tag{13}$$

Next, by invoking (12c), we obtain

$$\sum_{i=1}^{n} (x_{i,k} - F_i(x_{i,k})) = \sum_{i=1}^{n} \frac{x_{i,k} - T_i(x_{i,k})}{2\theta_i} \to -ny + \sum_{i=1}^{n} x_i. \tag{14}$$

Finally, by employing (12d), we see that for all $i, j \in \{1, \ldots, n\}$,

$$\begin{aligned} F_i(x_{i,k}) - F_j(x_{j,k}) &= \left(1 - \frac{1}{2\theta_i}\right) x_{i,k} - \left(1 - \frac{1}{2\theta_j}\right) x_{j,k} + \frac{T_i(x_{i,k})}{2\theta_i} - \frac{T_j(x_{j,k})}{2\theta_j} \\ &= (x_{i,k} - x_{j,k}) + \frac{T_i(x_{i,k}) - x_{i,k}}{2\theta_i} - \frac{T_j(x_{j,k}) - x_{j,k}}{2\theta_j} \to 0. \end{aligned} \tag{15}$$

Consequently, in view of (12a), (13), (14) and (15), we apply Corollary 3.1 to obtain

$$y = F_i(x_i) = \left(1 - \frac{1}{2\theta_i}\right)x_i + \frac{1}{2\theta_i}T_i(x_i), \quad \forall i \in \{1, \ldots, n\},$$

which concludes the proof. □

Remark 3.4 (Theorem 3.3 vs. Corollary 3.1) Since firmly nonexpansive operators are conically θ-averaged with $\theta = \frac{1}{2}$, it is clear that the assertion of Theorem 3.3 is more general than the one of Corollary 3.1. However, in view of the proof of Theorem 3.3, we conclude that the two assertions are equivalent.

We proceed in a similar manner to our discussion of demiclosedness principles for cocoercive operators; namely, we would like to have convergence conditions as in (12) of Theorem 3.3 that do not incorporate the conical average constants θ_i. Indeed, in the following result, we provide such conditions while, yet again, imposing a balance condition on the θ_i's. We focus our attention on a result concerning two mappings, which we will use in applications.

Theorem 3.4 (demiclosedness principle for two balanced averaged operators) *Let $T_1, T_2 : \mathcal{H} \to \mathcal{H}$ be θ_1- and θ_2-averaged mappings where $\theta_1, \theta_2 \in]0, 1[$, respectively, and suppose that there exist scalars $\rho_1, \rho_2 > 0$ such that*

$$\theta_1 \leq \frac{\rho_2}{\rho_1 + \rho_2} \quad \textit{and} \quad \theta_2 \leq \frac{\rho_1}{\rho_1 + \rho_2}. \tag{16}$$

Let $(x_{1,k})_{k=0}^{\infty}$ and $(x_{2,k})_{k=0}^{\infty}$ be sequences in $\mathcal{H}$ such that

$$x_{1,k} \rightharpoonup x_1 \quad \textit{and} \quad x_{2,k} \rightharpoonup x_2, \tag{17a}$$

$$T_1(x_{1,k}) \rightharpoonup y \quad \textit{and} \quad T_2(x_{2,k}) \rightharpoonup y, \tag{17b}$$

$$\rho_1(x_{1,k} - T_1(x_{1,k})) + \rho_2(x_{2,k} - T_2(x_{2,k})) \to 0, \tag{17c}$$

$$T_1(x_{1,k}) - T_2(x_{2,k}) \to 0. \tag{17d}$$

Then $T_1(x_1) = T_2(x_2) = y$.

Proof As in the proof of Theorem 3.2, by employing Lemma 2.1(ii), we may assume without the loss of generality that there is equality in (16), which we rewrite in the form

$$\left(\tfrac{1}{2} - \theta_1\right) + \left(\tfrac{1}{2} - \theta_2\right) = 0, \tag{18a}$$

$$\text{and} \quad \rho_1\theta_1 = \rho_2\theta_2. \tag{18b}$$

By combining (17a), (17b) and (17c), we see that $\rho_1(x_1 - y) + \rho_2(x_2 - y) = 0$, equivalently,

$$y = \frac{\rho_1 x_1 + \rho_2 x_2}{\rho_1 + \rho_2} = \theta_2 x_1 + \theta_1 x_2. \tag{19}$$

We set $\overline{y} := \frac{1}{2}(x_1 + x_2)$. By invoking (18a), it follows that

$$\begin{aligned} 2\theta_1\overline{y} + (1-2\theta_1)x_1 &= (1-\theta_1)x_1 + \theta_1 x_2 = \theta_2 x_1 + \theta_1 x_2, \\ 2\theta_2\overline{y} + (1-2\theta_2)x_2 &= \theta_2 x_1 + (1-\theta_2)x_2 = \theta_2 x_1 + \theta_1 x_2. \end{aligned} \tag{20}$$

Consequently, (17b) and (19) imply that

$$\begin{aligned} T_1(x_{1,k}) &\rightharpoonup 2\theta_1\overline{y} + (1-2\theta_1)x_1, \\ T_2(x_{2,k}) &\rightharpoonup 2\theta_2\overline{y} + (1-2\theta_2)x_2. \end{aligned} \tag{21}$$

Now, by (18b), (17c) and the definition of $\overline{y}$, it follows that

$$\begin{aligned} &\frac{x_{1,k} - T_1(x_{1,k})}{2\theta_1} + \frac{x_{2,k} - T_2(x_{2,k})}{2\theta_2} \\ &\quad = \frac{1}{2\rho_2\theta_2}\left(\rho_1\left(x_{1,k} - T_1(x_{1,k})\right) + \rho_2\left(x_{2,k} - T_2(x_{2,k})\right)\right) \to 0 = -2\overline{y} + x_1 + x_2. \end{aligned} \tag{22}$$

In addition, from (18a) it follows that

$$\begin{aligned} &(x_{1,k} - x_{2,k}) - \left(\frac{x_{1,k} - T_1(x_{1,k})}{2\theta_1} - \frac{x_{2,k} - T_2(x_{2,k})}{2\theta_2}\right) = \\ &\quad = \frac{T_1(x_{1,k}) - (1-2\theta_1)x_{1,k}}{2\theta_1} + \frac{(1-2\theta_2)x_{2,k} - T_2(x_{2,k})}{2\theta_2} \\ &\quad = \frac{(1-2\theta_1)}{2\theta_1}\left(T_1(x_{1,k}) - x_{1,k}\right) + T_1(x_{1,k}) - T_2(x_{2,k}) + \frac{(1-2\theta_2)}{2\theta_2}\left(x_{2,k} - T_2(x_{2,k})\right) \\ &\quad = (1-2\theta_2)\left(\frac{x_{1,k} - T_1(x_{1,k})}{2\theta_1} + \frac{x_{2,k} - T_2(x_{2,k})}{2\theta_2}\right) + \left(T_1(x_{1,k}) - T_2(x_{2,k})\right). \end{aligned} \tag{23}$$

By combining (23) with (22) and (17d), we obtain

$$(x_{1,k} - x_{2,k}) - \left(\frac{x_{1,k} - T_1(x_{1,k})}{2\theta_1} - \frac{x_{2,k} - T_2(x_{2,k})}{2\theta_2}\right) \to 0. \tag{24}$$

Finally, in view of (17a), (21), (22) and (24), we employ Theorem 3.3 in order to obtain

$$T_1(x_1) = 2\theta_1\overline{y} + (1-2\theta_1)x_1 \quad \text{and} \quad T_2(x_2) = 2\theta_2\overline{y} + (1-2\theta_2)x_2.$$

By recalling (19) and (20), we arrive at $T_1(x_1) = T_2(x_2) = y$. □

Remark 3.5 (on the balance condition) In Theorem 3.4, we did not provide an explicit balance condition for the averagedness constants as we did in (9). However, we did impose a stronger condition (16), which indeed implies

$$\frac{\rho_1\theta_1 + \rho_2\theta_2}{\rho_1 + \rho_2} \le \frac{2\rho_1\rho_2}{(\rho_1 + \rho_2)^2} \le \frac{1}{2};$$

that is, the weighted average of the θ_i's is at most $\frac{1}{2}$. This is always true for firmly nonexpansive mappings (see Remark 2.1(ii)).

4 Applications to the Adaptive Douglas–Rachford Algorithm

We recall that the problem of finding a zero of the sum of two operators $A, B : \mathcal{H} \rightrightarrows \mathcal{H}$ is

$$\text{find } x \in \mathcal{H} \text{ such that } 0 \in Ax + Bx. \tag{25}$$

In this section, we apply our generalized demiclosedness principles and derive the *weak convergence* of the shadow sequence of the *adaptive Douglas–Rachford (aDR)* algorithm, originally introduced in [12], in order to solve (25) for a weakly and a strongly monotone operators. The analysis is then extended in [13], which includes weakly and strongly comonotone operators as well.

Given $(\gamma, \delta, \lambda, \mu, \kappa) \in \mathbb{R}^5_{++}$, the *aDR operator* is defined by

$$T_{\text{aDR}} := (1-\kappa)\,\text{Id} + \kappa R_2 R_1, \tag{26}$$

where

$$J_1 := J_{\gamma A} = (\text{Id} + \gamma A)^{-1}, \quad R_1 := J^{\lambda}_{\gamma A} = (1-\lambda)\,\text{Id} + \lambda J_1,$$
$$J_2 := J_{\delta B} = (\text{Id} + \delta B)^{-1}, \quad R_2 := J^{\mu}_{\delta A} = (1-\mu)\,\text{Id} + \mu J_2.$$

Set an initial point $x_0 \in \mathcal{H}$. The aDR algorithm generates a sequence $(x_k)_{k=0}^{\infty}$ by the recurrence

$$x_{k+1} \in T_{\text{aDR}}(x_k), \quad k = 0, 1, 2, \ldots. \tag{27}$$

We observe that (26) coincides with the classical Douglas–Rachford operator in the case where $\lambda = \mu := 2$, $\gamma = \delta$, and $\kappa = 1/2$. Similarly to the classical DR algorithm, the fixed points of T_{aDR} are not explicit solutions of (25). Nonetheless, under the assumptions (see [13, Section 5])

$$(\lambda - 1)(\mu - 1) = 1 \quad \text{and} \quad \delta = (\lambda - 1)\gamma, \tag{28a}$$

equivalently,

$$\lambda = 1 + \frac{\delta}{\gamma} \quad \text{and} \quad \mu = 1 + \frac{\gamma}{\delta}, \tag{28b}$$

the fixed points are useful in order to obtain a solution as we show next.

Fact 4.1 (*aDR and solutions to the inclusion problem*) *Suppose that* $(\gamma, \delta) \in \mathbb{R}^2_{++}$, *that* λ, μ *are defined by* (28), *that* $\kappa > 0$, *and that* J_1 *is single-valued. Then*

(i) $\text{Id} - T_{\text{aDR}} = \kappa\mu(J_1 - J_2 R_1)$;
(ii) $J_1(\text{Fix}\, T_{\text{aDR}}) = \text{zer}(A + B)$.

Proof See [12, Lemma 4.1]. □

Suppose that $(x_k)_{k=0}^{\infty}$ is generated by the aDR algorithm and converges weakly to the limit point $x^\star \in \operatorname{Fix} T_{\mathrm{aDR}}$. Then Fact 4.1(ii) asserts that the *shadow* limit point $J_1(x^\star)$ is a solution of (25). Our aim is to prove that, under certain assumptions, the shadow sequence $(J_1(x_k))_{k=0}^{\infty}$ converges weakly to the shadow limit $J_1(x^\star)$.

In our analysis, we will employ the convergence results [13, Theorems 5.4 and 5.7]: Under the assumptions therein, the aDR operator (27) is shown to be averaged and, hence, single-valued. Consequently, in these cases, we will employ equality in (27).

4.1 Adaptive DR Algorithm for Monotone Operators

We begin our discussion with the case where the operators are maximally α-monotone and maximally β-monotone. We will prove the weak convergence of the aDR algorithm shadow sequence by means of a generalized demiclosedness principle. To this end, we recall the following fact regarding the convergence of the aDR algorithm.

Fact 4.2 (*aDR for monotone operators*) *Let $\alpha, \beta \in \mathbb{R}$ such that $\alpha + \beta \geq 0$ and let $A, B : \mathcal{H} \rightrightarrows \mathcal{H}$ be maximally α-monotone and maximally β-monotone, respectively, with* $\operatorname{zer}(A+B) \neq \varnothing$. *Let $(\gamma, \delta, \lambda, \mu) \in \mathbb{R}^4_{++}$ satisfy* (28) *and either*

$$\delta(1 + 2\gamma\alpha) = \gamma, \qquad \text{if } \alpha + \beta = 0, \tag{29a}$$

$$\text{or } (\gamma + \delta)^2 < 4\gamma\delta(1 + \gamma\alpha)(1 + \delta\beta), \qquad \text{if } \alpha + \beta > 0. \tag{29b}$$

Set $\kappa \in\,]0, \overline{\kappa}[$ where

$$\overline{\kappa} := \begin{cases} 1, & \text{if } \alpha + \beta = 0; \\ \dfrac{4\gamma\delta(1 + \gamma\alpha)(1 + \delta\beta) - (\gamma + \delta)^2}{2\gamma\delta(\gamma + \delta)(\alpha + \beta)}, & \text{if } \alpha + \beta > 0. \end{cases} \tag{30}$$

Set a starting point $x_0 \in \mathcal{H}$ and $(x_k)_{k=0}^{\infty}$ by $x_{k+1} = T_{\mathrm{aDR}}(x_k),\ k = 0, 1, 2, \ldots$. Then

(i) $x_k - x_{k+1} \to 0$;
(ii) $x_k \rightharpoonup x^\star \in \operatorname{Fix} T_{\mathrm{aDR}}$ *with* $J_1(x^\star) \in \operatorname{zer}(A + B)$.

Proof Combine [13, Theorem 5.7] with [13, Corollary 2.10] and Fact 4.1(ii). □

Theorem 4.1 (weak convergence of the shadow sequence under monotonicity) *Suppose that A and B are maximally α-monotone and maximally β-monotone, respectively, where $\alpha + \beta \geq 0$ and* $\operatorname{zer}(A + B) \neq \varnothing$. *Let $(\gamma, \delta, \lambda, \mu) \in \mathbb{R}^4_{++}$ satisfy* (28) *and* (29). *Let $\kappa \in\,]0, \overline{\kappa}[$ where $\overline{\kappa}$ is defined by* (30). *Set a starting point $x_0 \in \mathcal{H}$ and $(x_k)_{k=0}^{\infty}$ by $x_{k+1} = T_{\mathrm{aDR}}(x_k),\ k = 0, 1, 2, \ldots$. Then the shadow sequence $(J_1(x_k))_{k=0}^{\infty}$ converges weakly*

$$J_1(x_k) \rightharpoonup J_1(x^\star) \in \operatorname{zer}(A + B).$$

Proof Fact 4.2(ii) asserts that

$$x_k \rightharpoonup x^\star \in \operatorname{Fix} T_{\mathrm{aDR}} \text{ with } J_1(x^\star) \in \operatorname{zer}(A + B). \tag{31}$$

Consequently, $(x_k)_{k=0}^{\infty}$ is bounded. By combining (28) and (29), it follows that $1 + \gamma\alpha > 0$ and $1 + \delta\beta > 0$ (see [13, Theorem 5.7]). Thus, we employ Fact 2.2 which asserts that J_1 and J_2 are τ_1-cocoercive and τ_2-cocoercive, respectively, with full domain, where

$$\tau_1 := 1 + \gamma\alpha \quad \text{and} \quad \tau_2 := 1 + \delta\beta.$$

Due to the cocoerciveness of J_1, the shadow sequence $(J_1(x_k))_{k=0}^{\infty}$ is bounded and has a weak converging subsequence, say,

$$J_1\left(x_{k_j}\right) \rightharpoonup y^\star. \tag{32}$$

Set $z_k := R_1(x_k)$, for each $k = 0, 1, 2, \ldots$. Then

$$z_{k_j} \rightharpoonup (1-\lambda)x^\star + \lambda y^\star =: z^\star. \tag{33}$$

Moreover, Fact 4.2(i) and Fact 4.1(i) imply that

$$J_1(x_k) - J_2(z_k) \to 0. \tag{34}$$

which, when combined with (32), implies that

$$J_2(z_{k_j}) \rightharpoonup y^\star. \tag{35}$$

Thus, on the one hand, $J_1(x_{k_j}) - J_2(z_{k_j}) \to 0$ while, on the other hand,

$$\begin{aligned} J_1(x_{k_j}) - J_2(z_{k_j}) &= J_1(x_{k_j}) - R_1(x_{k_j}) + R_1(x_{k_j}) - J_2(z_{k_j}) \\ &= (\lambda - 1)\left(x_{k_j} - J_1(x_{k_j})\right) + \left(z_{k_j} - J_2(z_{k_j})\right) \\ &\rightharpoonup (\lambda - 1)(x^\star - y^\star) + (z^\star - y^\star) = -\lambda y^\star + (\lambda - 1)x^\star + z^\star. \end{aligned} \tag{36}$$

By combining (31)–(36), we see that the sequences $(x_{k_j})_{j=0}^{\infty}$ and $(z_{k_j})_{j=0}^{\infty}$ satisfy the conditions in (10) by setting

$$\rho_1 := \lambda - 1 > 0 \quad \text{and} \quad \rho_2 := 1 > 0. \tag{37}$$

With this choice of the parameters ρ_i's, we observe that the balance condition (9) is satisfied as well. Indeed, (28) implies that

$$\rho_1(\tau_1 - 1) + \rho_2(\tau_2 - 1) = (\lambda - 1)\gamma\alpha + \delta\beta = \delta(\alpha + \beta) \geq 0.$$

Consequently, we apply Theorem 3.2 in order to obtain $y^\star = J_1(x^\star)$. □

Remark 4.1 We observe that Theorem 4.1 guarantees the weak convergence of the shadow sequence whenever the original sequence converges weakly. In particular, this is guaranteed under the conditions on the parameters in Fact 4.2. However, this is a meaningful contribution only in the case where $\alpha + \beta = 0$: In the case where $\alpha + \beta > 0$, it is known that the shadow sequence converges not only weakly but, in fact, strongly (see [12, Theorem 4.5] and [13, Remark 5.8]).

4.2 Adaptive DR Algorithm for Comonotone Operators

We now address the weak convergence of the shadow sequence in the case where the operators are comonotone. To this end, we recall the following result regarding the convergence of the aDR algorithm.

Fact 4.3 (*aDR for comonotone operators*) *Let $\alpha, \beta \in \mathbb{R}$ such that $\alpha + \beta \geq 0$ and let $A, B : \mathcal{H} \rightrightarrows \mathcal{H}$ be maximally α-comonotone and maximally β-comonotone, respectively, such that* $\operatorname{zer}(A + B) \neq \varnothing$. *Let $(\gamma, \delta, \lambda, \mu) \in \mathbb{R}^4_{++}$ satisfy* (28) *and either*

$$\delta = \gamma + 2\alpha, \qquad \textit{if } \alpha + \beta = 0, \tag{38a}$$

$$\textit{or} \quad (\gamma + \delta)^2 < 4(\gamma + \alpha)(\delta + \beta), \qquad \textit{if } \alpha + \beta > 0. \tag{38b}$$

Set $\kappa \in \,]0, \overline{\kappa}[$ where

$$\overline{\kappa} := \begin{cases} 1, & \textit{if } \alpha + \beta = 0; \\ \dfrac{4(\gamma + \alpha)(\delta + \beta) - (\gamma + \delta)^2}{2(\gamma + \delta)(\alpha + \beta)}, & \textit{if } \alpha + \beta > 0. \end{cases} \tag{39}$$

Set a starting point $x_0 \in \mathcal{H}$ and $(x_k)_{k=0}^{\infty}$ by $x_{k+1} = T_{\mathrm{aDR}}(x_k),\ k = 0, 1, 2, \ldots$. Then

(i) $x_k - x_{k+1} \to 0$;
(ii) $x_k \rightharpoonup x^\star \in \operatorname{Fix} T_{\mathrm{aDR}}$ *with* $J_1(x^\star) \in \operatorname{zer}(A + B)$.

Proof Combine [13, Theorem 5.4] with [13, Corollary 2.10] and Fact 4.1(ii). □

Theorem 4.2 (weak convergence of the shadow sequence under comonotonicity) *Suppose that A and B are maximally α-comonotone and maximally β-comonotone such that $\alpha + \beta \geq 0$ and* $\operatorname{zer}(A + B) \neq \varnothing$. *Let $(\gamma, \delta, \lambda, \mu) \in \mathbb{R}^4_{++}$ satisfy* (28) *and suppose that*

$$(\gamma + \delta) \leq \min\{2(\gamma + \alpha), 2(\delta + \beta)\}. \tag{40}$$

Let $\kappa \in \,]0, \overline{\kappa}[$ where $\overline{\kappa}$ is defined by (39). *Set a starting point $x_0 \in \mathcal{H}$ and $(x_k)_{k=0}^{\infty}$ by $x_{k+1} = T_{\mathrm{aDR}}(x_k),\ k = 0, 1, 2, \ldots$. Then the shadow sequence $(J_1(x_k))_{k=0}^{\infty}$ converges weakly*

$$J_1(x_k) \rightharpoonup J_1(x^\star) \in \operatorname{zer}(A + B).$$

Proof We claim that (40) implies (38). Indeed, if $\alpha + \beta = 0$, we obtain

$$\begin{aligned}(\gamma+\delta) \leq \min\{2(\gamma+\alpha), 2(\delta-\alpha)\} \iff & (\gamma+\delta) \leq 2(\gamma+\alpha) \quad \text{and} \\ & (\gamma+\delta) \leq 2(\delta-\alpha) \\ \iff & \delta = \gamma + 2\alpha,\end{aligned}$$

which is (38a). On the other hand, observe that (40) trivially implies that

$$(\gamma+\delta)^2 \leq (\min\{2(\gamma+\alpha), 2(\delta+\beta)\})^2 \leq 4(\gamma+\alpha)(\delta+\beta).$$

Suppose that $(\gamma+\delta)^2 = (\min\{2(\gamma+\alpha), 2(\delta+\beta)\})^2 = 4(\gamma+\alpha)(\delta+\beta)$. Then

$$(\gamma+\delta) = 2(\gamma+\alpha) = 2(\beta+\delta),$$

which implies that $\alpha+\beta = 0$. Thus, (38b) holds whenever $\alpha+\beta > 0$ which concludes the proof of our claim. Consequently, we employ Fact 4.3 in order to obtain

$$x_k \rightharpoonup x^\star \in \operatorname{Fix} T_{\text{aDR}} \text{ with } J_1(x^\star) \in \operatorname{zer}(A+B).$$

We conclude that $(x_k)_{k=0}^\infty$ is bounded. Now, (28) and (38) imply that $\gamma+\alpha > 0$ and $\delta+\beta > 0$ (see [13, Theorem 5.4]). Thus, by Fact 2.3, J_1 and J_2 are conically θ_1-averaged and θ_2-averaged, respectively, with full domain, where

$$\theta_1 := \frac{\gamma}{2(\gamma+\alpha)} \quad \text{and} \quad \theta_2 := \frac{\delta}{2(\delta+\beta)}. \tag{41}$$

Since J_1 is conically averaged, the shadow sequence $(J_1(x_k))_{k=0}^\infty$ is bounded and has a weakly convergent subsequence, say, $J_1(x_{k_j}) \rightharpoonup y^\star$. By the same arguments as in the proof of Theorem 4.1, while employing Theorem 3.4 instead of Theorem 3.2, we arrive at

$$y^\star = J_1(x^\star),$$

which concludes the proof. To this end, it remains to verify that (16) holds, and then, Theorem 3.4 is applicable. Indeed, by (41), (37) and (28),

$$\begin{aligned}\theta_1 \leq \frac{\rho_2}{\rho_1+\rho_2} \quad \text{and} \quad \theta_2 \leq \frac{\rho_1}{\rho_1+\rho_2} \iff & \frac{\gamma}{2(\gamma+\alpha)} \leq \frac{1}{\lambda} \quad \text{and} \quad \frac{\delta}{2(\delta+\beta)} \leq \frac{\lambda-1}{\lambda} \\ \iff & \frac{\gamma}{2(\gamma+\alpha)} \leq \frac{1}{\lambda} \quad \text{and} \quad \frac{\gamma}{2(\delta+\beta)} \leq \frac{1}{\lambda} \\ \iff & \gamma\lambda \leq \min\{2(\gamma+\alpha), 2(\delta+\beta)\} \\ \iff & (\gamma+\delta) \leq \min\{2(\gamma+\alpha), 2(\delta+\beta)\},\end{aligned}$$

which is (40). □

Remark 4.2 We note that, in contrast to Theorem 4.1, Theorem 4.2 guarantees the weak convergence of the shadow sequence under more restrictive conditions on the parameters than the conditions in Fact 4.3, namely in the case where $\alpha + \beta > 0$. Nevertheless, Theorem 4.2 covers new ground since the convergence of the shadow sequence in the aDR for comonotone operators has not been previously addressed. Although outside the scope of this work, we believe that the convergence of the shadow sequence may be strong whenever $\alpha + \beta > 0$, as in the case of monotone operators.

5 Conclusions

In this paper, we extend the multi-operator demiclosedness principle [6] to more general classes of operators such as cocoercive and conically averaged operators. The new findings are natural and consistent with existing theory and are later justified by applications in which we show the weak convergence of the shadow sequence of the adaptive Douglas–Rachford algorithm. It remains of interest to find new connections between the demiclosedness principle and other classes of algorithms and optimization problems.

Acknowledgements We thank the referees for their valuable insights and corrections. Sedi Bartz was partially supported by a UMass Lowell faculty startup grant. Rubén Campoy was partially supported by a postdoctoral fellowship of UMass Lowell. Hung M. Phan was partially supported by Autodesk, Inc. via a gift made to the Department of Mathematical Sciences, UMass Lowell.

References

1. Browder, F.E.: Semicontractive and semiaccretive nonlinear mappings in Banach spaces. Bull. Am. Math. Soc. **74**, 660–665 (1968)
2. Garcia-Falset, J., Sims, B., Smyth, M.A.: The demiclosedness principle for mappings of asymptotically nonexpansive type. Houst. J. Math. **22**, 101–108 (1996)
3. Li, G., Kim, J.K.: Demiclosedness principle and asymptotic behavior for nonexpansive mappings in metric spaces. Appl. Math. Lett. **14**(5), 645–649 (2001)
4. Osilike, M.O., Udomene, A.: Demiclosedness principle and convergence theorems for strictly pseudocontractive mappings of Browder–Petryshyn type. J. Math. Anal. Appl. **256**(2), 431–445 (2001)
5. Zhou, H.: Demiclosedness principle with applications for asymptotically pseudo-contractions in Hilbert spaces. Nonlinear Anal. Theory Methods Appl. **70**(9), 3140–3145 (2009)
6. Bauschke, H.H.: New demiclosedness principles for (firmly) nonexpansive operators. In: Bailey, D., et al. (eds.) Computational and Analytical Mathematics. Springer Proceedings in Mathematics & Statistics, vol. 50, pp. 19–28. Springer, New York (2013)
7. Douglas, J., Rachford, H.H.: On the numerical solution of heat conduction problems in two and three space variables. Trans. Am. Math. Soc. **82**, 421–439 (1956)
8. Lions, P.L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. **16**(6), 964–979 (1979)
9. Bùi, M.N., Combettes, P.L.: The Douglas–Rachford Algorithm converges only weakly. SIAM J. Control Optim. **58**(2), 1118–1120 (2020)
10. Svaiter, B.F.: On weak convergence of the Douglas–Rachford method. SIAM J. Control Optim. **49**(1), 280–287 (2011)
11. Bauschke, H.H., Ouyang, H., Wang, X.: On circumcenter mappings induced by nonexpansive operators (2018). arXiv:1811.11420
12. Dao, M.N., Phan, H.M.: Adaptive Douglas–Rachford splitting algorithm for the sum of two operators. SIAM J. Optim. **29**(4), 2697–2724 (2019)

13. Bartz, S., Dao, M.N., Phan, H.M.: Conical averagedness and convergence analysis of fixed point algorithms (2019). arXiv:1910.14185
14. Bauschke, H.H., Moursi, W.M., Wang, X.: Generalized monotone operators and their averaged resolvents. Math. Program. (2020). https://doi.org/10.1007/s10107-020-01500-6
15. Giselsson, P., Moursi, W. M.: On compositions of special cases of Lipschitz continuous operators (2019). arXiv:1912.13165
16. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces, 2nd edn. Springer, Berlin (2017)
17. Combettes, P.L., Pennanen, T.: Proximal methods for cohypomonotone operators. SIAM J. Control Optim. **43**(2), 731–742 (2004)