

Federation University ResearchOnline

<https://researchonline.federation.edu.au>

Copyright Notice

This is the post-peer-review, pre-copyedit version of an article published in Journal of Global Optimization. The final authenticated version is available online at:

<https://doi.org/10.1007/s10898-021-01057-4>

Copyright © 2021, The Author(s)

Use of the Accepted Manuscript is subject to [AM terms of use](#), which permit users to view, print, copy, download and text and data-mine the content, for the purposes of academic research, subject always to the full conditions of use. Under no circumstances may the AM be shared or distributed under a Creative Commons, or other form of open access license, nor may it be reformatted or enhanced.

See this record in Federation ResearchOnline at:

<http://researchonline.federation.edu.au/vital/access/HandleResolver/1959.17/184764>

Conical averagedness and convergence analysis of fixed point algorithms

Sedi Bartz*, Minh N. Dao† and Hung M. Phan‡

November 27, 2020

Abstract

We study a conical extension of averaged nonexpansive operators and the role it plays in convergence analysis of fixed point algorithms. Various properties of conically averaged operators are systematically investigated, in particular, the stability under relaxations, convex combinations and compositions. We derive conical averagedness properties of resolvents of generalized monotone operators. These properties are then utilized in order to analyze the convergence of the proximal point algorithm, the forward-backward algorithm, and the adaptive Douglas–Rachford algorithm. Our study unifies, improves and casts new light on recent studies of these topics.

AMS Subject Classifications: Primary: 47H10, 49M27; Secondary: 65K05, 65K10.

Keywords: Adaptive Douglas–Rachford algorithm, cocoercivity, conically averaged operator, forward-backward algorithm, proximal point algorithm, strong monotonicity, weak monotonicity.

1. Introduction

Averaged nonexpansive operators, originally introduced in [1], are well known to be useful in convergence analysis of various fixed point algorithms, see [2, 4, 8, 10, 20, 26] and the references therein. In particular, iterative sequences generated by several fixed point algorithms can be expressed in terms of Krasnosel’skiĭ–Mann iterations [21, 23], the convergence of which relies upon an averagedness property. Although frequently understood in the single-valued setting, some averaged nonexpansive properties can also be explored in the set-valued framework. For instance, the notion of *union averaged nonexpansive operators* has been recently studied in [13] with applications to the local convergence of proximal algorithms; see also [30].

Each averaged operator is an under-relaxation of some nonexpansive operator. We demonstrate that over-relaxations of nonexpansive operators also arise naturally in several situations. In our study, we consider *conically averaged operators* and provide a framework that unifies both types of relaxations for nonexpansive operators. We then show that this class of operators plays a significant role in convergence analysis of several fixed point algorithms, in particular, the *proximal point algorithm*, the *forward–backward algorithm*, and the *adaptive Douglas–Rachford (DR) algorithm*.

*Department of Mathematical Sciences, Kennedy College of Sciences, University of Massachusetts Lowell, Lowell, MA 01854, USA. E-mail: sedi_bartz@uml.edu.

†School of Engineering, Information Technology and Physical Sciences, Federation University Australia, Ballarat, VIC 3353, Australia. E-mail: m.dao@federation.edu.au

‡Department of Mathematical Sciences, Kennedy College of Sciences, University of Massachusetts Lowell, Lowell, MA 01854, USA. E-mail: hung_phan@uml.edu.

The proximal point algorithm was introduced by Martinet [24] and further developed by Rockafellar [29] for finding a zero of a maximally monotone operator. This was followed by other early studies such as [6]. Since then, the proximal point algorithm turned into an indispensable tool of optimization, both in theory and in applications. In fact, several iterative optimization algorithms can be reformulated as special cases of the proximal point algorithm, see, for example, [15] and references therein. On the other hand, the forward-backward algorithm was first proposed by Lions and Mercier [22] and Passty [27] for finding a zero of the sum of two maximally monotone operators. This splitting idea can be traced back to the projected gradient method [18]. Another well-known splitting algorithm is the DR algorithm, initially studied by Douglas and Rachford [14] in the setting of linear operators and was later generalized for maximally monotone operators by Lions and Mercier, also in [22]. It is worth mentioning that both the forward-backward and the DR algorithms reduce to the proximal point algorithm in the case where one operator is zero. Recently, the so-called adaptive DR algorithm has been proposed in [11] to deal with scenarios that lack classical monotonicity.

The main objectives of our study are as follows.

- (i) We systematically study the conical averagedness property and prove that it is stable under relaxations, convex combinations and compositions. In order to establish these properties, we extend known results from [2, Section 4.5] and [10, Section 2] for nonexpansive and averaged operators. In particular, under suitable conditions, in Theorem 2.7 we show that compositions of scaled conically averaged operators are also conically averaged. Furthermore, we show that relaxed resolvents of generalized monotone operators either belong to or directly linked to the class of conically averaged operators.
- (ii) Under generalized monotonicity assumptions, we provide convergence analysis by means of conical averagedness, in particular, in Theorem 4.2 for the relaxed proximal point algorithm, in Theorem 4.3 for the relaxed forward-backward algorithm, in Theorems 5.4 and 5.7 for the adaptive DR algorithm. We emphasize that these splitting algorithms incorporate several operators, some of which may not be averaged due to the lack of classical monotonicity. Nevertheless, we prove that their compositions are, in fact, conically averaged. Consequently, we derive convergence and rate of asymptotic regularity. This approach sheds light on the structures of these algorithms and yields simple and transparent convergence proofs. An application to the strongly-weakly convex optimization problem is included in Theorem 5.12. Our analysis improves several contemporary results on this topic, e.g., [2, Proposition 26.1], [11, Theorems 4.5 and 5.4], and [19, Theorems 4.4 and 4.6].

The remainder of this paper is organized as follows. In Section 2 we present conically averaged operators as well as several interesting properties which are useful for our analysis. In Section 3 we study relaxed resolvents of generalized monotone operators in relation with their conical averagedness properties. In Sections 4 and 5 we provide our main results regarding conical averagedness of operators associated with the previously mentioned algorithms, which, in turn, lead to their convergence and rate of asymptotic regularity.

We will employ standard notations that generally follow [2]. Throughout, X is a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. The set of nonnegative integers is denoted by $\mathbb{N}$, the set of real numbers by $\mathbb{R}$, the set of nonnegative real numbers by $\mathbb{R}_+ := \{x \in \mathbb{R} \mid x \geq 0\}$, and the set of the positive real numbers by $\mathbb{R}_{++} := \{x \in \mathbb{R} \mid x > 0\}$. We use the notation $A: X \rightrightarrows X$ to indicate that A is a set-valued operator on X and the notation $A: X \rightarrow X$ to indicate that A is a single-valued operator on X . We denote the *domain* of the mapping $A: X \rightrightarrows X$ by

$\text{dom } A := \{x \in X \mid Ax \neq \emptyset\}$, its set of *zeros* by $\text{zer } A := \{x \in X \mid 0 \in Ax\}$, and its set of *fixed points* by $\text{Fix } A := \{x \in X \mid x \in Ax\}$. Id denotes the *identity mapping*.

We will often use the following identity. For every $\sigma, \tau \in \mathbb{R}$ and $s, t \in X$,

$$\|\sigma s + \tau t\|^2 = \sigma(\sigma + \tau)\|s\|^2 + \tau(\sigma + \tau)\|t\|^2 - \sigma\tau\|s - t\|^2; \quad (1)$$

moreover, if $\sigma + \tau \neq 0$, then

$$\sigma\|s\|^2 + \tau\|t\|^2 = \frac{1}{\sigma + \tau}\|\sigma s + \tau t\|^2 + \frac{\sigma\tau}{\sigma + \tau}\|s - t\|^2. \quad (2)$$

2. Conically averaged operators

We recall that the mapping $T: X \rightarrow X$ is *nonexpansive* if it is Lipschitz continuous with constant 1 on its domain, i.e.,

$$\forall x, y \in \text{dom } T, \quad \|Tx - Ty\| \leq \|x - y\|. \quad (3)$$

T is said to be *θ -averaged* if $\theta \in]0, 1[$ and $T = (1 - \theta)\text{Id} + \theta N$ for some nonexpansive operator $N: X \rightarrow X$, see, e.g., [2, Definition 4.33]. We now extend this concept to allow for $\theta \in \mathbb{R}_{++}$.

Definition 2.1 (conically averaged operator). We say that an operator $T: X \rightarrow X$ is *conically averaged* with constant $\theta \in \mathbb{R}_{++}$, or *conically θ -averaged*, if there exists a nonexpansive operator $N: X \rightarrow X$ such that

$$T = (1 - \theta)\text{Id} + \theta N. \quad (4)$$

Let T be conically θ -averaged. Then T is nonexpansive when $\theta = 1$, and T is θ -averaged when $\theta \in]0, 1[$. As one would expect, conically averaged operators also possess properties similar to averaged operators. Indeed, we now present several properties which generalize similar properties from [2, 10], see also [3] for a related development where conically averaged operators were referred to as conically nonexpansive operators.

Proposition 2.2. *Let $T: X \rightarrow X$, $\theta \in \mathbb{R}_{++}$, and $\lambda \in \mathbb{R}_{++}$. Then the following assertions are equivalent:*

- (i) T is conically θ -averaged.
- (ii) $(1 - \lambda)\text{Id} + \lambda T$ is conically $\lambda\theta$ -averaged.
- (iii) For all $x, y \in \text{dom } T$,

$$\|Tx - Ty\|^2 \leq \|x - y\|^2 - \frac{1 - \theta}{\theta}\|(\text{Id} - T)x - (\text{Id} - T)y\|^2. \quad (5)$$

- (iv) For all $x, y \in \text{dom } T$,

$$\|Tx - Ty\|^2 + (1 - 2\theta)\|x - y\|^2 \leq 2(1 - \theta)\langle x - y, Tx - Ty \rangle. \quad (6)$$

Proof. Set $N := (1 - 1/\theta)\text{Id} + (1/\theta)T$. Then $T = (1 - \theta)\text{Id} + \theta N$ and $(1 - \lambda)\text{Id} + \lambda T = (1 - \lambda\theta)\text{Id} + \lambda\theta N$. By definition,

$$T \text{ is conically } \theta\text{-averaged} \iff N \text{ is nonexpansive} \quad (7a)$$

$$\iff (1 - \lambda)\text{Id} + \lambda T \text{ is conically } \lambda\theta\text{-averaged}, \quad (7b)$$

which implies the equivalence between (i) and (ii).

Next, we note that $\text{Id} - N = (\text{Id} - T)/\theta$ and $\text{dom } T = \text{dom } N =: D$. Now, by employing (1), for all $x, y \in D$

$$\|Tx - Ty\|^2 = \|(1 - \theta)(x - y) + \theta(Nx - Ny)\|^2 \quad (8a)$$

$$= (1 - \theta)\|x - y\|^2 + \theta\|Nx - Ny\|^2 - \theta(1 - \theta)\|(\text{Id} - N)x - (\text{Id} - N)y\|^2 \quad (8b)$$

$$= \|x - y\|^2 - \frac{1 - \theta}{\theta}\|(\text{Id} - T)x - (\text{Id} - T)y\|^2 + \theta(\|Nx - Ny\|^2 - \|x - y\|^2). \quad (8c)$$

Consequently, we see that

$$T \text{ is conically } \theta\text{-averaged} \quad (9a)$$

$$\iff N \text{ is nonexpansive} \quad (9b)$$

$$\iff \forall x, y \in D, \quad \|Nx - Ny\| \leq \|x - y\| \quad (9c)$$

$$\iff \forall x, y \in D, \quad \|Tx - Ty\|^2 \leq \|x - y\|^2 - \frac{1 - \theta}{\theta}\|(\text{Id} - T)x - (\text{Id} - T)y\|^2 \quad (9d)$$

$$\iff \forall x, y \in D, \quad \|Tx - Ty\|^2 + (1 - 2\theta)\|x - y\|^2 \leq 2(1 - \theta) \langle x - y, Tx - Ty \rangle, \quad (9e)$$

and we get the equivalence between (i), (iii), and (iv). The proof is complete. $\blacksquare$

In view of Proposition 2.2, we see that T is $1/2$ -averaged if and only if, for all $x, y \in \text{dom } T$

$$\|Tx - Ty\|^2 \leq \|x - y\|^2 - \|(\text{Id} - T)x - (\text{Id} - T)y\|^2, \quad (10)$$

in which case T is said to be *firmly nonexpansive* (see also [17, Proposition 11.2]).

Proposition 2.3. *Let $T: X \rightarrow X$ and $\lambda \in \mathbb{R}_{++}$. Then T is firmly nonexpansive if and only if $\text{Id} - \lambda T$ is conically $\lambda/2$ -averaged.*

Proof. Set $N := 2T - \text{Id}$. Then $T = (1/2)\text{Id} + (1/2)N$ is firmly nonexpansive $\iff N$ is nonexpansive $\iff \text{Id} - \lambda T = (1 - \lambda/2)\text{Id} + (\lambda/2)(-N)$ is conically $\lambda/2$ -averaged. $\blacksquare$

The following results reiterate and extend similar facts regarding nonexpansive and averaged operators from [2, Section 4.5] and [10, Section 2].

Proposition 2.4 (convex combination of conically averaged operators). *Let I be a finite index set. For each $i \in I$ let $T_i: X \rightarrow X$ be conically θ_i -averaged. Let $\{w_i\}_{i \in I} \subseteq \mathbb{R}_{++}$, $\sum_{i \in I} w_i = 1$. Then $\sum_{i \in I} w_i T_i$ is conically θ -averaged where $\theta := \sum_{i \in I} w_i \theta_i$.*

Proof. For each $i \in I$, there exists a nonexpansive operator N_i such that $T_i = (1 - \theta_i)\text{Id} + \theta_i N_i$. It follows that

$$\sum_{i \in I} w_i T_i = \sum_{i \in I} w_i (1 - \theta_i) \text{Id} + \sum_{i \in I} w_i \theta_i N_i = (1 - \theta) \text{Id} + \theta \sum_{i \in I} \frac{w_i \theta_i}{\theta} N_i. \quad (11)$$

Since, by [2, Proposition 4.9(i)], $\sum_{i \in I} \frac{w_i \theta_i}{\theta} N_i$ is nonexpansive, the proof is complete. $\blacksquare$

Proposition 2.5 (composition of two conically averaged operators). *Let $T_1: X \rightarrow X$ and $T_2: X \rightarrow X$ be conically θ_1 -averaged and conically θ_2 -averaged, respectively. Suppose that either $\theta_1 = \theta_2 = 1$ or $\theta_1 \theta_2 < 1$. Let $\omega \in \mathbb{R} \setminus \{0\}$ and set*

$$T := \left(\frac{1}{\omega} T_2\right)(\omega T_1) \quad \text{and} \quad \theta := \begin{cases} 1 & \text{if } \theta_1 = \theta_2 = 1, \\ \frac{\theta_1 + \theta_2 - 2\theta_1 \theta_2}{1 - \theta_1 \theta_2} & \text{if } \theta_1 \theta_2 < 1. \end{cases} \quad (12)$$

Then T is conically θ -averaged. Moreover,

(i) $\theta = 1$ if and only if $\theta_1 = 1$ or $\theta_2 = 1$.

(ii) $\theta < 1$ if and only if $\theta_1 < 1$ and $\theta_2 < 1$.

Proof. Let $x, y \in \text{dom } T_2(\omega T_1)$. By applying Proposition 2.2 to T_2 and then to T_1 , we obtain

$$\left\| \left(\frac{1}{\omega} T_2 \right) (\omega T_1) x - \left(\frac{1}{\omega} T_2 \right) (\omega T_1) y \right\|^2 = \frac{1}{\omega^2} \|T_2(\omega T_1 x) - T_2(\omega T_1 y)\|^2 \quad (13a)$$

$$\leq \frac{1}{\omega^2} \left(\|\omega T_1 x - \omega T_1 y\|^2 - \frac{1 - \theta_2}{\theta_2} \|(\text{Id} - T_2)(\omega T_1 x) - (\text{Id} - T_2)(\omega T_1 y)\|^2 \right) \quad (13b)$$

$$= \|T_1 x - T_1 y\|^2 - \frac{1 - \theta_2}{\theta_2} \left\| \left(T_1 - \left(\frac{1}{\omega} T_2 \right) (\omega T_1) \right) x - \left(T_1 - \left(\frac{1}{\omega} T_2 \right) (\omega T_1) \right) y \right\|^2 \quad (13c)$$

$$\leq \|x - y\|^2 - \frac{1 - \theta_1}{\theta_1} \|(\text{Id} - T_1)x - (\text{Id} - T_1)y\|^2 - \frac{1 - \theta_2}{\theta_2} \left\| \left(T_1 - \left(\frac{1}{\omega} T_2 \right) (\omega T_1) \right) x - \left(T_1 - \left(\frac{1}{\omega} T_2 \right) (\omega T_1) \right) y \right\|. \quad (13d)$$

Case 1: If $\theta_1 = \theta_2 = 1$, then (13) implies that $(\frac{1}{\omega} T_2)(\omega T_1)$ is nonexpansive, i.e., conically 1-averaged.

Case 2: If $\theta_1 \theta_2 < 1$, then

$$\frac{1}{\theta_1} + \frac{1}{\theta_2} \geq \frac{2}{\sqrt{\theta_1 \theta_2}} > 2, \quad (14)$$

which implies

$$\frac{1 - \theta_1}{\theta_1} + \frac{1 - \theta_2}{\theta_2} > 0. \quad (15)$$

By setting $s := (\text{Id} - T_1)x - (\text{Id} - T_1)y$ and $t := \left(T_1 - \left(\frac{1}{\omega} T_2 \right) (\omega T_1) \right) x - \left(T_1 - \left(\frac{1}{\omega} T_2 \right) (\omega T_1) \right) y$ and by employing (2) and (15) we see that

$$\frac{1 - \theta_1}{\theta_1} \|s\|^2 + \frac{1 - \theta_2}{\theta_2} \|t\|^2 = \frac{\frac{1 - \theta_1}{\theta_1} \frac{1 - \theta_2}{\theta_2}}{\frac{1 - \theta_1}{\theta_1} + \frac{1 - \theta_2}{\theta_2}} \|s + t\|^2 + \frac{1}{\frac{1 - \theta_1}{\theta_1} + \frac{1 - \theta_2}{\theta_2}} \left\| \frac{1 - \theta_1}{\theta_1} s - \frac{1 - \theta_2}{\theta_2} t \right\|^2 \quad (16a)$$

$$\geq \frac{\frac{1 - \theta_1}{\theta_1} \frac{1 - \theta_2}{\theta_2}}{\frac{1 - \theta_1}{\theta_1} + \frac{1 - \theta_2}{\theta_2}} \left\| \left(\text{Id} - \left(\frac{1}{\omega} T_2 \right) (\omega T_1) \right) x - \left(\text{Id} - \left(\frac{1}{\omega} T_2 \right) (\omega T_1) \right) y \right\|^2 \quad (16b)$$

$$= \frac{(1 - \theta_1)(1 - \theta_2)}{\theta_1 + \theta_2 - 2\theta_1\theta_2} \left\| \left(\text{Id} - \left(\frac{1}{\omega} T_2 \right) (\omega T_1) \right) x - \left(\text{Id} - \left(\frac{1}{\omega} T_2 \right) (\omega T_1) \right) y \right\|^2 \quad (16c)$$

$$= \frac{1 - \theta}{\theta} \left\| \left(\text{Id} - \left(\frac{1}{\omega} T_2 \right) (\omega T_1) \right) x - \left(\text{Id} - \left(\frac{1}{\omega} T_2 \right) (\omega T_1) \right) y \right\|^2, \quad (16d)$$

where

$$\theta := \frac{\theta_1 + \theta_2 - 2\theta_1\theta_2}{1 - \theta_1\theta_2} > 0. \quad (17)$$

By recalling (13) and Proposition 2.2(i)&(iii), we conclude that $(\frac{1}{\omega} T_2)(\omega T_1)$ is conically θ -averaged.

(i): From (17) it follows that

$$\theta = 1 \iff \theta_1 + \theta_2 - 2\theta_1\theta_2 = 1 - \theta_1\theta_2 \iff (1 - \theta_1)(1 - \theta_2) = 0. \quad (18)$$

Consequently, by recalling the definition of θ , we conclude that $\theta = 1$ if and only if $\theta_1 = 1$ or $\theta_2 = 1$.

(ii): It follows from our assumption and the definition of θ that

$$\theta < 1 \iff \theta_1\theta_2 < 1 \text{ and } \theta_1 + \theta_2 - 2\theta_1\theta_2 < 1 - \theta_1\theta_2 \quad (19a)$$

$$\iff \theta_1\theta_2 < 1 \text{ and } (1 - \theta_1)(1 - \theta_2) > 0 \quad (19b)$$

$$\iff \theta_1 < 1 \text{ and } \theta_2 < 1 \quad (19c)$$

which concludes the proof. $\blacksquare$

It will be of convenience in our applications to include an equivalent reformulation of Proposition 2.5. The following result generalizes [2, Proposition 4.44] and also recaptures [16, Proposition 3.12], in which the composition of an averaged operator and a so-called *negatively averaged* operator was considered.

Corollary 2.6. *Let $T_1, T_2: X \rightarrow X$, $\omega_1, \omega_2 \in \mathbb{R} \setminus \{0\}$, and $\theta_1, \theta_2 \in \mathbb{R}_{++}$. Suppose that $\omega_1 T_1$ and $\omega_2 T_2$ are conically θ_1 -averaged and conically θ_2 -averaged, respectively, where either $\theta_1 = \theta_2 = 1$ or $\theta_1\theta_2 < 1$. Then $\omega_1\omega_2 T_2 T_1$ and $\omega_1\omega_2 T_1 T_2$ are both conically θ -averaged, where*

$$\theta := \begin{cases} 1 & \text{if } \theta_1 = \theta_2 = 1, \\ \frac{\theta_1 + \theta_2 - 2\theta_1\theta_2}{1 - \theta_1\theta_2} & \text{if } \theta_1\theta_2 < 1. \end{cases} \quad (20)$$

Proof. We note that $\omega_1\omega_2 T_2 T_1 = \omega_1(\omega_2 T_2)(\frac{1}{\omega_1}(\omega_1 T_1))$ and $\omega_1\omega_2 T_1 T_2 = \omega_2(\omega_1 T_1)(\frac{1}{\omega_2}(\omega_2 T_2))$. Consequently, the proof follows from Proposition 2.5. $\blacksquare$

The following theorem extends [2, Proposition 4.9(i) and Proposition 4.46] beyond the composition of conically averaged operators by incorporating *scalar multiplications* of the operators.

Theorem 2.7 (composition of conically averaged operators). *Let $m \geq 2$ be an integer. For each $i \in I := \{1, \dots, m\}$ let $T_i: X \rightarrow X$ be conically θ_i -averaged. Let $\{\omega_i\}_{i \in I} \subseteq \mathbb{R}$, $\omega_1\omega_2 \cdots \omega_m = 1$. Set*

$$T := (\omega_m T_m)(\omega_{m-1} T_{m-1}) \cdots (\omega_1 T_1). \quad (21)$$

Then

(i) *If $\max_{i \in I} \theta_i \leq 1$, then T is nonexpansive.*

(ii) *If $\theta_i \neq 1$ for each $i \in I$ and*

$$\forall k \in \{2, \dots, m\}, \quad \theta_k < 1 + \frac{1}{\sum_{i=1}^{k-1} \frac{\theta_i}{1 - \theta_i}}, \quad (22)$$

then T is conically θ -averaged where

$$\theta := \frac{1}{1 + \frac{1}{\sum_{i \in I} \frac{\theta_i}{1 - \theta_i}}}. \quad (23)$$

(iii) *If $\max_{i \in I} \theta_i < 1$, then T is θ -averaged where $\theta < 1$ is given by (23).*

Proof. (i): The proof follows from Proposition 2.5 by induction on m .

(ii): We prove by induction on m . The case where $m = 2$ is a straightforward conclusion from Proposition 2.5. Now, suppose that the statement is true for $m - 1$. For each $i \in \{1, \dots, m\}$ let T_i be a conically θ_i -averaged mapping such that

$$\forall k \in \{2, \dots, m\}, \quad \theta_k < 1 + \frac{1}{\sum_{i=1}^{k-1} \frac{\theta_i}{1 - \theta_i}}. \quad (24)$$

Let $\{\omega_i\}_{i \in I} \subseteq \mathbb{R}$, $\omega_1 \omega_2 \cdots \omega_m = 1$. We prove that the operator

$$T = (\omega_m T_m)(\omega_{m-1} T_{m-1}) \cdots (\omega_1 T_1) \quad (25)$$

is conically θ -averaged with θ given by (23).

Since the statement is true for $m - 1$ conically averaged operators, we have

$$T^* := (\omega_m \omega_{m-1} T_{m-1})(\omega_{m-2} T_{m-2}) \cdots (\omega_1 T_1) \quad (26)$$

is θ^* -averaged where

$$\theta^* := \frac{1}{1 + \frac{1}{\sum_{i=1}^{m-1} \frac{\theta_i}{1-\theta_i}}}. \quad (27)$$

By letting $k = m$ in (24), we see that

$$\theta_m < \frac{1}{\theta^*}. \quad (28)$$

We now apply Proposition 2.5 to two operators T^* and T_m in order to conclude that $T = (\omega_m T_m)(\frac{1}{\omega_m} T^*)$ is conically θ_0 -averaged where

$$\theta_0 := \frac{\theta^* + \theta_m - 2\theta^* \theta_m}{1 - \theta^* \theta_m}. \quad (29)$$

It follows that

$$\frac{\theta_0}{1 - \theta_0} = \frac{\theta^*}{1 - \theta^*} + \frac{\theta_m}{1 - \theta_m} = \sum_{i=1}^{m-1} \frac{\theta_i}{1 - \theta_i} + \frac{\theta_m}{1 - \theta_m} = \sum_{i=1}^m \frac{\theta_i}{1 - \theta_i} \quad (30)$$

Consequently,

$$\theta_0 = \frac{1}{1 + \frac{1}{\sum_{i=1}^m \frac{\theta_i}{1-\theta_i}}} = \theta. \quad (31)$$

which concludes the proof in the case $k = m$.

(iii): For each $i \in I$, since $\theta_i < 1$, then $\frac{\theta_i}{1-\theta_i} > 0$. Consequently,

$$\forall k \in \{2, \dots, m\}, \quad \theta_k < 1 < 1 + \frac{1}{\sum_{i=1}^{k-1} \frac{\theta_i}{1-\theta_i}}, \quad (32)$$

in particular, we see that (22) is satisfied. The conclusion now follows by employing (ii). $\blacksquare$

Corollary 2.8. *Let $m \geq 2$ be an integer. For each $i \in I := \{1, \dots, m\}$ let $T_i: X \rightarrow X$ and $\omega_i \in \mathbb{R} \setminus \{0\}$ be such that $\omega_i T_i$ conically θ_i -averaged. Set*

$$T := (\omega_m \omega_{m-1} \cdots \omega_1)(T_m T_{m-1} \cdots T_1). \quad (33)$$

Then

(i) *If $\max_{i \in I} \theta_i \leq 1$, then T is nonexpansive.*

(ii) *If $\theta_i \neq 1$ for each $i \in I$ and*

$$\forall k \in \{2, \dots, m\}, \quad \theta_k < 1 + \frac{1}{\sum_{i=1}^{k-1} \frac{\theta_i}{1-\theta_i}}, \quad (34)$$

then T is conically θ -averaged where

$$\theta := \frac{1}{1 + \frac{1}{\sum_{i \in I} \frac{\theta_i}{1-\theta_i}}}. \quad (35)$$

(iii) If $\max_{i \in I} \theta_i < 1$, then T is θ -averaged where $\theta < 1$ is given by (35).

Proof. The proof follows by observing that

$$T = \left((\omega_1 \omega_2 \cdots \omega_{m-1})(\omega_m T_m) \right) \left(\frac{1}{\omega_{m-1}}(\omega_{m-1} T_{m-1}) \right) \cdots \left(\frac{1}{\omega_1}(\omega_1 T_1) \right) \quad (36)$$

and by applying Theorem 2.7 to operators $\omega_1 T_1, \omega_2 T_2, \dots, \omega_{m-1} T_{m-1}, \omega_m T_m$ and scalars $\frac{1}{\omega_1}, \frac{1}{\omega_2}, \dots, \frac{1}{\omega_{m-1}}, \omega_1 \cdots \omega_{m-1}$. ■

We conclude this section by employing Fejér monotonicity in order to prove the convergence of sequences generated by an averaged/conically averaged operator. Recall that a sequence $(x_n)_{n \in \mathbb{N}}$ is said to be *Fejér monotone* with respect to a nonempty subset C of X if

$$\forall c \in C, \forall n \in \mathbb{N}, \quad \|x_{n+1} - c\| \leq \|x_n - c\|. \quad (37)$$

The following result somewhat extends [2, Theorem 5.15], see also [5, 28]. We include a proof for convenience.

Proposition 2.9 (Krasnosel'skiĭ–Mann iterations). *Let T be a conically θ -averaged operator with full domain and $\text{Fix } T \neq \emptyset$. Let $x_0 \in X$. For each $n \in \mathbb{N}$ set*

$$x_{n+1} = (1 - \lambda_n)x_n + \lambda_n T x_n, \quad (38)$$

where $(\lambda_n)_{n \in \mathbb{N}}$ is a sequence in $[0, 1/\theta]$ such that $\sum_{n=0}^{+\infty} \lambda_n(1 - \theta\lambda_n) = +\infty$. Then

- (i) $(x_n)_{n \in \mathbb{N}}$ is Fejér monotone with respect to $\text{Fix } T$.
- (ii) $(x_n - T x_n)_{n \in \mathbb{N}}$ converges strongly to 0.
- (iii) $(x_n)_{n \in \mathbb{N}}$ converges weakly to a point in $\text{Fix } T$.
- (iv) If $\liminf_{n \rightarrow +\infty} \lambda_n(1 - \theta\lambda_n) > 0$, then $\|x_n - T x_n\| = o(1/\sqrt{n})$.

Proof. Since T is conically θ -averaged with full domain, then $T = (1 - \theta)\text{Id} + \theta N$ for some nonexpansive operator $N: X \rightarrow X$. Consequently, $\text{Id} - T = \theta(\text{Id} - N)$, $\text{Fix } T = \text{Fix } N$, and

$$\forall n \in \mathbb{N}, \quad x_{n+1} = (1 - \theta\lambda_n)x_n + \theta\lambda_n N x_n. \quad (39)$$

(i): For all $y \in \text{Fix } N$ and all $n \in \mathbb{N}$, by employing (1) and the nonexpansiveness of N we obtain

$$\|x_{n+1} - y\|^2 = \|(1 - \theta\lambda_n)(x_n - y) + \theta\lambda_n(N x_n - y)\|^2 \quad (40a)$$

$$= (1 - \theta\lambda_n)\|x_n - y\|^2 + \theta\lambda_n\|N x_n - N y\|^2 - \theta\lambda_n(1 - \theta\lambda_n)\|x_n - N x_n\|^2 \quad (40b)$$

$$\leq \|x_n - y\|^2 - \theta\lambda_n(1 - \theta\lambda_n)\|x_n - N x_n\|^2 \quad (40c)$$

Since $\lambda_n(1 - \theta\lambda_n) \geq 0$ for all n , $(x_n)_{n \in \mathbb{N}}$ is Fejér monotone with respect to $\text{Fix } N = \text{Fix } T$.

(ii): By telescoping (40) over $n \in \mathbb{N}$,

$$\theta \sum_{n=0}^{+\infty} \lambda_n(1 - \theta\lambda_n)\|x_n - N x_n\|^2 \leq \|x_0 - y\|^2 < +\infty. \quad (41)$$

Since $\sum_{n=0}^{+\infty} \lambda_n(1 - \theta\lambda_n) = +\infty$, it follows that

$$\liminf_{n \rightarrow +\infty} \|x_n - N x_n\| = 0. \quad (42)$$

Moreover, since $x_{n+1} = (1 - \theta\lambda_n)x_n + \theta\lambda_n Nx_n$ and since N is nonexpansive,

$$\|x_{n+1} - Nx_{n+1}\| = \|(1 - \theta\lambda_n)(x_n - Nx_n) + (Nx_n - Nx_{n+1})\| \quad (43a)$$

$$\leq (1 - \theta\lambda_n)\|x_n - Nx_n\| + \|Nx_n - Nx_{n+1}\| \quad (43b)$$

$$\leq (1 - \theta\lambda_n)\|x_n - Nx_n\| + \|x_n - x_{n+1}\| \quad (43c)$$

$$= \|x_n - Nx_n\|. \quad (43d)$$

We see that $(\|x_n - Nx_n\|)_{n \in \mathbb{N}}$ is decreasing and bounded below by 0, hence it converges. Consequently, (42) implies that

$$x_n - Tx_n = \theta(x_n - Nx_n) \rightarrow 0 \quad \text{as } n \rightarrow +\infty. \quad (44)$$

(iii): Let x^* be a weak cluster point of $(x_n)_{n \in \mathbb{N}}$, that is, there exists a subsequence $(x_{k_n})_{n \in \mathbb{N}}$ such that $x_{k_n} \rightharpoonup x^*$. By combining (44) and [2, Corollary 4.28], $x^* \in \text{Fix } N$. In turn, [2, Theorem 5.5] implies that $(x_n)_{n \in \mathbb{N}}$ converges weakly to a point in $\text{Fix } N = \text{Fix } T$.

(iv): It follows from $\liminf_{n \rightarrow +\infty} \lambda_n(1 - \theta\lambda_n) > 0$ and (41) that $\sum_{n=0}^{+\infty} \|x_n - Nx_n\|^2 < +\infty$, which, when combined with (43), implies that

$$\frac{n}{2} \|x_n - Nx_n\|^2 \leq \sum_{k=\lfloor n/2 \rfloor}^n \|x_k - Nx_k\|^2 \rightarrow 0 \quad \text{as } n \rightarrow +\infty, \quad (45)$$

where $\lfloor n/2 \rfloor$ is the largest integer majorized by $n/2$. Therefore, $\|x_n - Tx_n\| = \theta\|x_n - Nx_n\| = o(1/\sqrt{n})$ as $n \rightarrow +\infty$. ■

Corollary 2.10 (convergence of averaged operators). *Let T be a θ -averaged operator with full domain and $\text{Fix } T \neq \emptyset$. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence generated by T . Then $(x_n)_{n \in \mathbb{N}}$ converges weakly to a point in $\text{Fix } T$ and the rate of asymptotic regularity of T is $o(1/\sqrt{n})$, i.e., $\|x_n - Tx_n\| = o(1/\sqrt{n})$ as $n \rightarrow +\infty$.*

Proof. We employ Proposition 2.9 with $\theta < 1$ and $\lambda_n = 1$ for all $n \in \mathbb{N}$. ■

3. Generalized monotonicity

We recall the following standard notations. Let $A: X \rightrightarrows X$. The *graph* of A is the set $\text{gra } A := \{(x, u) \in X \times X \mid u \in Ax\}$ and the *inverse* of A , denoted by A^{-1} , is the operator with graph $\text{gra } A^{-1} := \{(u, x) \in X \times X \mid u \in Ax\}$. The *resolvent* of A is the mapping defined by

$$J_A := (\text{Id} + A)^{-1} \quad (46)$$

and the *relaxed resolvent* of A with parameter $\lambda \in \mathbb{R}_+$ is the mapping defined by

$$J_A^\lambda := (1 - \lambda)\text{Id} + \lambda J_A. \quad (47)$$

The mapping $R_A := 2J_A - \text{Id}$ is the *reflected resolvent* of A .

Let $\alpha \in \mathbb{R}$. We recall that A is α -*monotone* (see, for example, [11]) if

$$\forall (x, u), (y, v) \in \text{gra } A, \quad \langle x - y, u - v \rangle \geq \alpha \|x - y\|^2 \quad (48)$$

and α -*comonotone* (see, for example, [3]) if

$$\forall (x, u), (y, v) \in \text{gra } A, \quad \langle x - y, u - v \rangle \geq \alpha \|u - v\|^2. \quad (49)$$

We say that A is *maximally α -monotone* (respectively, *maximally α -comonotone*) if it is α -monotone (respectively, α -comonotone) and there is no α -monotone (respectively, α -comonotone) operator $B: X \rightrightarrows X$ such that $\text{gra } A$ is properly contained in $\text{gra } B$.

We note that both 0-monotonicity and 0-comonotonicity simply mean monotonicity. If $\alpha > 0$, then α -monotonicity is actually α -strong monotonicity in [2, Definition 22.1(iv)], while α -comonotonicity coincides with α -cocoercivity [2, Definition 4.10]. If $\alpha < 0$, then α -monotonicity and α -comonotonicity are respectively α -hypomonotonicity and α -cohypomonotonicity in [9, Definition 2.2]. Additionally, α -monotonicity with $\alpha < 0$ is also referred to as *weak monotonicity* in [11, Section 3]. We refer the reader to [2, 7] for more discussions on maximal monotonicity and some of its variants.

Remark 3.1. Several properties are immediate from the definitions.

- (i) A is α -comonotone if and only if A^{-1} is α -monotone.
- (ii) A is maximally α -comonotone if and only if A^{-1} is maximally α -monotone.
- (iii) If A is α -comonotone with $\alpha \geq 0$, then A is also monotone.
- (iv) If A is α -comonotone with $\alpha > 0$, then A is single-valued and $1/\alpha$ -Lipschitz continuous.

In the case where $\alpha \geq 0$, the following characterizations of maximal α -monotonicity and maximal α -comonotonicity hold.

Proposition 3.2 (maximal α -monotonicity and α -comonotonicity). *Let $A: X \rightrightarrows X$ and $\alpha \in \mathbb{R}_+$. Then*

- (i) A is maximally α -monotone if and only if A is α -monotone and maximally monotone.
- (ii) A is maximally α -comonotone if and only if A is α -comonotone and maximally monotone.

Proof. (i): See [11, Proposition 3.5(i)].

(ii): A is maximally α -comonotone $\iff A^{-1}$ is maximally α -monotone $\iff A^{-1}$ is α -monotone and maximally monotone (by (i)) $\iff A$ is α -comonotone and maximally monotone. ■

We now collect several useful properties of relaxed resolvents of α -monotone and α -comonotone operators. Some parts of the following results are available in [2, 3, 11, 12]. For convenience, these results are included here as well as their proofs. In particular, we will show that if an operator is either α -monotone or α -comonotone, then its relaxed resolvent is, to a certain extent, related to a conically averaged operator. These results play a crucial role in convergence analysis of several iterative algorithms. We begin our discussion with the following auxiliary properties.

Proposition 3.3. *Let $A: X \rightarrow X$ and $\alpha, \lambda \in \mathbb{R}_{++}$. Then the following assertions are equivalent:*

- (i) A is α -comonotone (i.e., α -cocoercive).
- (ii) αA is firmly nonexpansive.
- (iii) $\text{Id} - \lambda A$ is conically $\frac{\lambda}{2\alpha}$ -averaged.

Proof. The equivalence between (i) and (ii) follows, e.g., from [2, Remark 4.34(iv)]. The equivalence between (ii) and (iii) follows from Proposition 2.3 by observing that $\text{Id} - \lambda A = \text{Id} - \frac{\lambda}{\alpha}(\alpha A)$. ■

Proposition 3.4 (single-valuedness and full domain). *Let $A: X \rightrightarrows X$ be α -monotone and let $\gamma \in \mathbb{R}_{++}$ such that $1 + \gamma\alpha > 0$. Then*

- (i) $J_{\gamma A}$ is (at most) single-valued.
- (ii) $\text{dom } J_{\gamma A} = X$ if and only if A is maximally α -monotone.

Proof. See [11, Proposition 3.4]. ■

Proposition 3.5 (relaxed resolvents of α -monotone operators). *Let $A: X \rightrightarrows X$ be α -monotone and let $\gamma \in \mathbb{R}_{++}$ be such that $1 + \gamma\alpha > 0$. Set $R := (1 - \lambda)\text{Id} + \lambda J_{\gamma A}$ where $\lambda \in]1, +\infty[$. Then*

- (i) $J_{\gamma A}$ is $(1 + \gamma\alpha)$ -comonotone. Consequently, $(1 + \gamma\alpha)J_{\gamma A}$ is firmly nonexpansive.
- (ii) $\frac{1}{1-\lambda}R$ is conically $\frac{\lambda}{2(\lambda-1)(1+\gamma\alpha)}$ -averaged.

Proof. (i): By Proposition 3.4(i) $J_{\gamma A}$ is single-valued, and, by [11, Lemma 3.3], it is $(1 + \gamma\alpha)$ -cocoercive, i.e., $(1 + \gamma\alpha)$ -comonotone. Since $1 + \gamma\alpha > 0$, Proposition 3.3 implies that $(1 + \gamma\alpha)J_{\gamma A}$ is firmly nonexpansive.

(ii): Since $(1 + \gamma\alpha)J_{\gamma A}$ is firmly nonexpansive, Proposition 2.3 implies that

$$\frac{1}{1-\lambda}R = \text{Id} + \frac{\lambda}{1-\lambda}J_{\gamma A} = \text{Id} - \frac{\lambda}{(\lambda-1)(1+\gamma\alpha)}(1+\gamma\alpha)J_{\gamma A} \quad (50)$$

is conically $\frac{\lambda}{2(\lambda-1)(1+\gamma\alpha)}$ -averaged. ■

Lemma 3.6 (resolvents of α -comonotone operators). *Let $A: X \rightrightarrows X$ and $\gamma \in \mathbb{R}_{++}$. Then A is α -comonotone if and only if for all $(x, a), (y, b) \in \text{gra } J_{\gamma A}$,*

$$(\gamma + 2\alpha) \langle x - y, a - b \rangle \geq \alpha \|x - y\|^2 + (\gamma + \alpha) \|a - b\|^2. \quad (51)$$

Consequently, if A is α -comonotone and $J_{\gamma A}$ is single-valued, then, for all $x, y \in \text{dom } J_{\gamma A}$,

$$(\gamma + 2\alpha) \langle x - y, J_{\gamma A}x - J_{\gamma A}y \rangle \geq \alpha \|x - y\|^2 + (\gamma + \alpha) \|J_{\gamma A}x - J_{\gamma A}y\|^2. \quad (52)$$

Proof. Let $(a, u), (b, v) \in \text{gra } A$ and set $x := a + \gamma u$ and $y := b + \gamma v$. Then

$$\langle a - b, u - v \rangle \geq \alpha \|u - v\|^2 \quad (53a)$$

$$\iff \gamma \langle a - b, \gamma u - \gamma v \rangle \geq \alpha \|\gamma u - \gamma v\|^2 \quad (53b)$$

$$\iff \gamma \langle a - b, (x - y) - (a - b) \rangle \geq \alpha \|(x - y) - (a - b)\|^2 \quad (53c)$$

$$\iff \gamma \langle x - y, a - b \rangle - \gamma \|a - b\|^2 \geq \alpha \left(\|x - y\|^2 + \|a - b\|^2 - 2 \langle x - y, a - b \rangle \right) \quad (53d)$$

$$\iff (\gamma + 2\alpha) \langle x - y, a - b \rangle \geq \alpha \|x - y\|^2 + (\gamma + \alpha) \|a - b\|^2 \quad (53e)$$

which completes the proof. ■

Proposition 3.7 (single-valuedness and full domain). *Let $A: X \rightrightarrows X$ be α -comonotone and let $\gamma \in \mathbb{R}_{++}$ be such that $\gamma + \alpha > 0$. Then*

- (i) $J_{\gamma A}$ is (at most) single-valued.

(ii) $\text{dom } J_{\gamma A} = X$ if and only if A is maximally α -comonotone.

Proof. (i): This follows from (51) in Lemma 3.6 and the fact that $\gamma + \alpha > 0$.

(ii): Since A is α -comonotone, $A' := A^{-1} - \alpha \text{Id}$ is monotone. For $x \in X$,

$$a \in (\text{Id} - J_{\gamma A})x \iff x \in (x - a) + \gamma A(x - a) \quad (54a)$$

$$\iff x - a \in A^{-1} \left(\frac{a}{\gamma} \right) = (\alpha \text{Id} + A') \left(\frac{a}{\gamma} \right) \quad (54b)$$

$$\iff x \in (\gamma + \alpha) \left(\text{Id} + \frac{1}{\gamma + \alpha} A' \right) \left(\frac{a}{\gamma} \right) \quad (54c)$$

$$\iff a \in \gamma \left(\text{Id} + \frac{1}{\gamma + \alpha} A' \right)^{-1} \left(\frac{x}{\gamma + \alpha} \right). \quad (54d)$$

Therefore,

$$\text{Id} - J_{\gamma A} = \gamma J_{\frac{1}{\gamma + \alpha} A'} \left(\frac{1}{\gamma + \alpha} \text{Id} \right). \quad (55)$$

We conclude that $\text{dom } J_{\gamma A} = X$ if and only if $\text{dom } J_{\frac{1}{\gamma + \alpha} A'} = X$, which is equivalent to A' being maximally monotone (see [2, Theorem 21.1 and Proposition 20.22]). The proof follows by observing that the maximal monotonicity of A' is equivalent to the maximal α -comonotonicity of A . ■

The following result is an extension of [2, Proposition 23.14].

Proposition 3.8 (relaxed resolvents of α -comonotone operators). *Let $A: X \rightrightarrows X$ be α -comonotone and let $\gamma \in \mathbb{R}_{++}$ be such that $\gamma + \alpha > 0$. Set $R := (1 - \lambda) \text{Id} + \lambda J_{\gamma A}$ with $\lambda \in \mathbb{R}_{++}$. Then*

(i) $J_{\gamma A}$ is conically $\frac{\gamma}{2(\gamma + \alpha)}$ -averaged.

(ii) R is conically $\frac{\lambda\gamma}{2(\gamma + \alpha)}$ -averaged.

Proof. Proposition 3.7(i) implies that $J_{\gamma A}$ is single-valued. Consequently, Lemma 3.6 implies that, for all $x, y \in \text{dom } J_{\gamma A}$,

$$(\gamma + 2\alpha) \langle x - y, J_{\gamma A}x - J_{\gamma A}y \rangle \geq \alpha \|x - y\|^2 + (\gamma + \alpha) \|J_{\gamma A}x - J_{\gamma A}y\|^2 \quad (56)$$

which is equivalent to

$$2 \left(1 - \frac{\gamma}{2(\gamma + \alpha)} \right) \langle x - y, J_{\gamma A}x - J_{\gamma A}y \rangle \geq \left(1 - \frac{\gamma}{\gamma + \alpha} \right) \|x - y\|^2 + \|J_{\gamma A}x - J_{\gamma A}y\|^2. \quad (57)$$

Proposition 2.2(i)&(iv) now implies that $J_{\gamma A}$ is conically $\frac{\gamma}{2(\gamma + \alpha)}$ -averaged. In turn, Proposition 2.2(i)&(ii) implies that R is conically $\frac{\lambda\gamma}{2(\gamma + \alpha)}$ -averaged. ■

We will now address the connection between conical averagedness and several fixed point algorithms including the forward-backward algorithm and the adaptive Douglas–Rachford algorithm.

4. Relaxed forward-backward algorithm

Let $A: X \rightrightarrows X$, $B: X \rightarrow X$. We consider the problem

$$\text{find } x \in X \text{ such that } 0 \in Ax + Bx. \quad (58)$$

Let $\gamma \in \mathbb{R}_{++}$ and $\kappa \in \mathbb{R}_{++}$. Set $x_0 \in X$. The *relaxed forward-backward (rFB) algorithm* for problem (58) generates a sequence $(x_n)_{n \in \mathbb{N}}$ via

$$\forall n \in \mathbb{N}, \quad x_{n+1} \in T_{\text{FB}} x_n \quad \text{where} \quad T_{\text{FB}} := (1 - \kappa) \text{Id} + \kappa J_{\gamma A} (\text{Id} - \gamma B). \quad (59)$$

In the case where $\kappa = 1$, the rFB algorithm is the well-studied *forward-backward algorithm*, see, e.g., [2, Section 26.5].

When one considers an iterative fixed point algorithm in order to solve a problem, the relations between the fixed points of that algorithm and the solutions of the problem under consideration is crucial. The following well-known result, see, e.g., [2, Proposition 26.1(iv)(a)], asserts that the fixed points of the forward-backward algorithm are, in fact, solutions of problem (58). For completeness, we include a proof as well.

Lemma 4.1. *With the settings of (58)–(59),*

$$\text{Fix } T_{\text{FB}} = \text{Fix } (J_{\gamma A} (\text{Id} - \gamma B)) = \text{zer}(A + B). \quad (60)$$

Proof. We note that $\text{Id} - T_{\text{FB}} = \kappa (\text{Id} - J_{\gamma A} (\text{Id} - \gamma B))$. Consequently, $\text{Fix } T_{\text{FB}} = \text{Fix } (J_{\gamma A} (\text{Id} - \gamma B))$. It follows that $x \in \text{Fix } (J_{\gamma A} (\text{Id} - \gamma B)) \iff x \in J_{\gamma A} (x - \gamma Bx) \iff x - \gamma Bx \in x + \gamma Ax \iff 0 \in Ax + Bx$. ■

In the case where $B = 0$, problem (58) reduces to finding a zero of the operator $A: X \rightrightarrows X$, i.e.,

$$\text{find } x \in X \text{ such that } 0 \in Ax \quad (61)$$

and the corresponding rFB algorithm reduces to the *relaxed proximal point algorithm* of the form

$$\forall n \in \mathbb{N}, \quad x_{n+1} \in T_{\text{PP}} x_n \quad \text{where} \quad T_{\text{PP}} := (1 - \kappa) \text{Id} + \kappa J_{\gamma A}. \quad (62)$$

In this case, Lemma 4.1 implies that

$$\text{Fix } T_{\text{PP}} = \text{Fix } J_{\gamma A} = \text{zer } A. \quad (63)$$

The following results are the main results of this section. We provide the averagedness of T_{PP} and T_{FB} as well as the convergence of the corresponding algorithms in cases where A is not necessarily monotone. Classical results for monotone operators can be found, for example, in [2, Example 23.40 and Proposition 26.1(iv)(d)].

Theorem 4.2 (relaxed proximal point algorithm). *Suppose that A is maximally α -comonotone with $\alpha \in \mathbb{R}$ and that $\gamma > \max\{0, -\alpha\}$. Then T_{PP} is conically $\frac{\kappa}{\kappa^*}$ -averaged and has full domain, where $\kappa^* := \frac{2(\gamma + \alpha)}{\gamma}$. Moreover, if $\text{zer } A \neq \emptyset$ and $\kappa < \kappa^*$, then every sequence $(x_n)_{n \in \mathbb{N}}$ generated by T_{PP} converges weakly to a point in $\text{zer } A$ and the rate of asymptotic regularity of T_{PP} is $o(1/\sqrt{n})$.*

Proof. As $\gamma + \alpha > 0$, employing Proposition 3.7, we see that $J_{\gamma A}$ and hence T_{PP} are single-valued and have full domain. By Proposition 3.8(ii), T_{PP} is conically θ -averaged, where

$$\theta := \frac{\kappa \gamma}{2(\gamma + \alpha)} = \frac{\kappa}{\kappa^*}. \quad (64)$$

The proof follows from Corollary 2.10 and (63). ■

Theorem 4.3 (relaxed forward-backward algorithm). *Suppose that A is maximally α -comonotone with $\alpha \in \mathbb{R}$, that B is β -comonotone with $\beta \in \mathbb{R}_{++}$ (i.e., β -cocoercive), and that either*

- (i) $\alpha + \beta = 0$ and $\gamma = 2\beta$; or
- (ii) $\alpha + \beta > 0$ and $\max\{0, 2\beta - 2\sqrt{\beta(\alpha + \beta)}\} < \gamma < 2\beta + 2\sqrt{\beta(\alpha + \beta)}$.

Then T_{FB} is conically $\frac{\kappa}{\kappa^*}$ -averaged and has full domain, where

$$\kappa^* := \begin{cases} 1 & \text{if } \alpha + \beta = 0, \\ \frac{4(\gamma + \alpha)\beta - \gamma^2}{2\gamma(\alpha + \beta)} & \text{if } \alpha + \beta > 0. \end{cases} \quad (65)$$

Moreover, if $\text{zer}(A + B) \neq \emptyset$ and $\kappa < \kappa^*$, then every sequence $(x_n)_{n \in \mathbb{N}}$ generated by T_{FB} converges weakly to a point in $\text{zer}(A + B)$ and the rate of asymptotic regularity of T_{FB} is $o(1/\sqrt{n})$.

Proof. On the one hand, (i) implies that $\gamma + \alpha = 2\beta - \beta = \beta > 0$ and $\gamma = 2(\gamma + \alpha) = 2\beta$. On the other hand, (ii) is equivalent to $\gamma^2 < 4(\gamma + \alpha)\beta$, which implies that $\gamma + \alpha > 0$ (since $\beta > 0$) and $\kappa^* > 0$ (since $\gamma > 0$). Now, since A is maximally α -comonotone, Proposition 3.7 implies that $J_{\gamma A}$ is single-valued and has full domain, and so does T_{FB} . By Proposition 3.8(i), $J_{\gamma A}$ is conically θ_1 -averaged, where

$$\theta_1 := \frac{\gamma}{2(\gamma + \alpha)}. \quad (66)$$

Next, since B is β -comonotone, Proposition 3.3 implies that $\text{Id} - \gamma B$ is conically θ_2 -averaged, where

$$\theta_2 := \frac{\gamma}{2\beta}. \quad (67)$$

We observe that if (i) holds, then $\theta_1 = \theta_2 = 1$ and $\frac{1}{\kappa^*} = 1$; if (ii) holds, then $\theta_1 \theta_2 < 1$ and

$$\frac{1}{\kappa^*} = \frac{2\gamma(\alpha + \beta)}{4(\gamma + \alpha)\beta - \gamma^2} = \frac{\theta_1 + \theta_2 - 2\theta_1\theta_2}{1 - \theta_1\theta_2}. \quad (68)$$

In view of Proposition 2.5, $J_{\gamma A}(\text{Id} - \gamma B)$ is conically $\frac{1}{\kappa^*}$ -averaged, and by Proposition 2.2(i)&(ii), T_{FB} is conically $\frac{\kappa}{\kappa^*}$ -averaged. The proof then follows from Corollary 2.10 and Lemma 4.1. $\blacksquare$

Corollary 4.4. *Suppose that A is maximally monotone, that B is β -comonotone with $\beta \in \mathbb{R}_{++}$ (i.e., β -cocoercive), and that $\gamma \in]0, 4\beta[$. Then T_{FB} is conically $\frac{\kappa}{\kappa^*}$ -averaged and has full domain, where $\kappa^* := \frac{4\beta - \gamma}{2\beta}$. Moreover, if $\text{zer}(A + B) \neq \emptyset$ and $\kappa < \kappa^*$, then every sequence $(x_n)_{n \in \mathbb{N}}$ generated by T_{FB} converges weakly to a point in $\text{zer}(A + B)$ and the rate of asymptotic regularity of T_{FB} is $o(1/\sqrt{n})$.*

Proof. Since A is maximally 0-comonotone, we apply Theorem 4.3(ii) with $\alpha = 0$. $\blacksquare$

Remark 4.5 (range of parameter γ). We recall that classical convergence analysis for the forward-backward algorithm requires that $\gamma \in]0, 2\beta[$, see, for example, [2, Proposition 26.1(iv)(d) and Theorems 26.14(i)]. Corollary 4.4 improves upon that by only requiring $\gamma \in]0, 4\beta[$.

5. Adaptive Douglas–Rachford algorithm

We focus on problem (58) where $A: X \rightrightarrows X$ and $B: X \rightrightarrows X$. We set $(\gamma, \delta) \in \mathbb{R}_{++}^2$ and $(\lambda, \mu, \kappa) \in \mathbb{R}_{++}^3$. The *adaptive DR operator*, introduced and studied in [11], is defined by

$$T := T_{A,B} := (1 - \kappa) \text{Id} + \kappa R_2 R_1 \quad (69)$$

where

$$R_1 := (1 - \lambda) \text{Id} + \lambda J_{\gamma A} \quad \text{and} \quad R_2 := (1 - \mu) \text{Id} + \mu J_{\delta B}. \quad (70)$$

Set $x_0 \in X$. Then the *adaptive DR (aDR) algorithm* for problem (58) generates a sequence $(x_n)_{n \in \mathbb{N}}$, also called a *DR sequence*, by letting

$$\forall n \in \mathbb{N}, \quad x_{n+1} \in T x_n. \quad (71)$$

Naturally, we refer to the case where $\delta = \gamma > 0$ and $\lambda = \mu = 2$ as the *classical DR algorithm* (or simply DR).

Unlike in the case of the forward-backward counterpart, the fixed points of the adaptive DR algorithm, in general, do not directly solve (58). Nevertheless, by choosing compatible parameters, we show that the images of the fixed points under the resolvent are, in fact, solutions. To this end, similarly to [11, Section 4], we assume that

$$(\lambda - 1)(\mu - 1) = 1 \quad \text{and} \quad \delta = \gamma(\lambda - 1), \quad (72)$$

equivalently,

$$\lambda = 1 + \frac{\delta}{\gamma} \quad \text{and} \quad \mu = 1 + \frac{\gamma}{\delta} \quad (73)$$

which clearly holds for the classical DR algorithm. Our settings are justified by the following fact.

Fact 5.1 (fixed points of the aDR operator). *Suppose that (73) holds. Then $\text{Fix } T \neq \emptyset$ if and only if $\text{zer}(A + B) \neq \emptyset$. Moreover, if $J_{\gamma A}$ is single-valued, then*

$$J_{\gamma A}(\text{Fix } T) = \text{zer}(A + B). \quad (74)$$

Proof. See [11, Lemma 4.1(iii)]. ■

In view of Fact 5.1 we focus on the convergence of the adaptive DR algorithm to a fixed point of the operator T under condition (73). In turn, such convergence can be guaranteed by the averagedness as we show in the following result.

Proposition 5.2 (convergence of the aDR algorithm via averagedness). *Suppose that T is θ -averaged and has full domain, that $\text{zer}(A + B) \neq \emptyset$, and that (73) holds. Then the rate of asymptotic regularity of T is $o(1/\sqrt{n})$ and every sequence $(x_n)_{n \in \mathbb{N}}$ generated by T converges weakly to a point $\bar{x} \in \text{Fix } T$. Moreover, if $J_{\gamma A}$ is single-valued, then $J_{\gamma A} \bar{x} \in \text{zer}(A + B)$.*

Proof. Since $\text{zer}(A + B) \neq \emptyset$, Fact 5.1 implies that $\text{Fix } T \neq \emptyset$. Now, by Corollary 2.10, the rate of asymptotic regularity of T is $o(1/\sqrt{n})$ and every sequence $(x_n)_{n \in \mathbb{N}}$ generated by T converges weakly to a point $\bar{x} \in \text{Fix } T$. Moreover, if $J_{\gamma A}$ is single-valued, then Fact 5.1 implies that $J_{\gamma A} \bar{x} \in \text{zer}(A + B)$. ■

Motivated by these observations, we focus on the averagedness of the adaptive DR operator. To this end, we look for compatible parameters $\gamma, \delta, \lambda, \mu, \kappa$. In fact, it is enough to determine only $\gamma, \delta, \kappa > 0$ which, in turn, determine λ, μ via (73).

5.1. The case of α -comonotone and β -comonotone operators

We consider the adaptive DR operators for two comonotone operators. In particular, we derive convergence by employing conical averagedness. To this end, we will make use of the following lemma.

Lemma 5.3 (existence of parameters). *Let $\alpha, \beta \in \mathbb{R}$ be such that $\alpha + \beta \geq 0$, and let $\gamma, \delta \in \mathbb{R}_{++}$. Set*

$$\gamma_0 := \begin{cases} 0 & \text{if } \alpha \geq 0, \\ 2\beta - 2\sqrt{\beta(\alpha + \beta)} & \text{if } \alpha < 0. \end{cases} \quad (75)$$

and set $\Delta := (\gamma + \alpha)(\alpha + \beta)$. Then $\gamma_0 \geq \max\{0, -\alpha\}$ and the following assertions are equivalent:

- (i) $(\gamma + \delta)^2 \leq 4(\gamma + \alpha)(\delta + \beta)$.
- (ii) $\gamma > \gamma_0$ and $\gamma + 2\alpha - 2\sqrt{\Delta} \leq \delta \leq \gamma + 2\alpha + 2\sqrt{\Delta}$.

Consequently, if $\alpha + \beta = 0$, then (i) and (ii) are equivalent to $\gamma > \max\{0, -2\alpha\}$ and $\delta = \gamma + 2\gamma$; if $\alpha + \beta > 0$, then there always exist $\gamma, \delta \in \mathbb{R}_{++}$ such that

$$\gamma > \gamma_0 \quad \text{and} \quad \max\{0, \gamma + 2\alpha - 2\sqrt{\Delta}\} < \delta < \gamma + 2\alpha + 2\sqrt{\Delta}, \quad (76)$$

in which case, all inequalities in (i) and (ii) are strict. Moreover, if (i) or (ii) holds, then $\gamma + \alpha > 0$ and $\delta + \beta > 0$.

Proof. If $\alpha \geq 0$, then $\gamma_0 = 0 = \max\{0, -\alpha\}$. If $\alpha < 0$, then $\beta > \alpha + \beta \geq 0$ and

$$\gamma_0 = 2\beta - 2\sqrt{\beta(\alpha + \beta)} = (\sqrt{\beta} - \sqrt{\alpha + \beta})^2 - \alpha \geq -\alpha = \max\{0, -\alpha\}. \quad (77)$$

Next, it is clear that

$$(\gamma + \delta)^2 \leq 4(\gamma + \alpha)(\delta + \beta) \quad (78a)$$

$$\iff \delta^2 - 2(\gamma + 2\alpha)\delta + \gamma^2 - 4(\gamma + \alpha)\beta \leq 0 \quad (78b)$$

$$\iff \Delta = (\gamma + \alpha)(\alpha + \beta) \geq 0 \quad \text{and} \quad \gamma + 2\alpha - 2\sqrt{\Delta} \leq \delta \leq \gamma + 2\alpha + 2\sqrt{\Delta}. \quad (78c)$$

In addition, since $\gamma_0 \geq \{0, -\alpha\}$ and $\alpha + \beta \geq 0$, we have $\Delta \geq 0$ as soon as $\gamma > \gamma_0$. By combining this with (78), we see that (ii) implies (i).

We now assume (i). Then, by (78), $\Delta \geq 0$ and $\gamma + 2\alpha - 2\sqrt{\Delta} \leq \delta \leq \gamma + 2\alpha + 2\sqrt{\Delta}$. Since $\delta > 0$, it follows that $\gamma + 2\alpha + 2\sqrt{\Delta} > 0$. To show (ii), it suffices to show that $\gamma > \gamma_0$. We distinguish between the following cases.

Case 1: $\alpha + \beta = 0$. Then $\gamma_0 = \max\{0, -2\alpha\}$, $\Delta = 0$, and

$$\gamma + 2\alpha + 2\sqrt{\Delta} > 0 \iff \gamma + 2\alpha > 0, \quad (79)$$

which implies that $\gamma > \gamma_0 = \max\{0, -2\alpha\}$.

Case 2: $\alpha + \beta > 0$. Since $\Delta \geq 0$, then $\gamma + \alpha \geq 0$. If $\alpha \geq 0$, then $\gamma_0 = 0$, so $\gamma > \gamma_0$. Now, assume that $\alpha < 0$. Then $\beta > \alpha + \beta \geq 0$ and

$$\gamma + 2\alpha + 2\sqrt{\Delta} > 0 \iff (\sqrt{\gamma + \alpha} + \sqrt{\alpha + \beta})^2 > \beta \quad (80a)$$

$$\iff \sqrt{\gamma + \alpha} > \sqrt{\beta} - \sqrt{\alpha + \beta} \quad (80b)$$

$$\iff \gamma + \alpha > (\sqrt{\beta} - \sqrt{\alpha + \beta})^2 \quad (80c)$$

$$\iff \gamma > \gamma_0 = 2\beta - 2\sqrt{\beta(\alpha + \beta)}. \quad (80d)$$

This completes the proof of the equivalence between (i) and (ii). On the other hand, if we set $\gamma > \gamma_0$, then by (80), there exists δ such that $\max\{0, \gamma + 2\alpha - 2\sqrt{\Delta}\} < \delta < \gamma + 2\alpha + 2\sqrt{\Delta}$, which leads to strict inequalities in (i) and (ii).

Finally, since (i) and (ii) are equivalent, if one of them holds, then $\gamma > \gamma_0 \geq \max\{0, -\alpha\}$, so $\gamma + \alpha > 0$, which, when combined with (i), implies $\delta + \beta > 0$. $\blacksquare$

Theorem 5.4 (aDR for α -comonotone and β -comonotone operators). *Suppose that A and B are maximally α -comonotone and β -comonotone, respectively. Set λ, μ as in (73), and suppose that either*

(i) $\alpha + \beta = 0$, $\gamma > \max\{0, -2\alpha\}$, $\delta = \gamma + 2\alpha$, and $\kappa^* := 1$; or

(ii) $\alpha + \beta > 0$ and $\kappa^* := \frac{4(\gamma+\alpha)(\delta+\beta)-(\gamma+\delta)^2}{2(\gamma+\delta)(\alpha+\beta)} > 0$.

Then the adaptive DR operators $T_{A,B}$ and $T_{B,A}$ are conically $\frac{\kappa}{\kappa^}$ -averaged and have full domain. Moreover, if $\text{zer}(A+B) \neq \emptyset$ and $\kappa < \kappa^*$, then, for any $(T, C) \in \{(T_{A,B}, \gamma A), (T_{B,A}, \delta B)\}$, every sequence $(x_n)_{n \in \mathbb{N}}$ generated by T converges weakly to a point $\bar{x} \in \text{Fix } T$ with $J_C \bar{x} \in \text{zer}(A+B)$ and the rate of asymptotic regularity of T is $o(1/\sqrt{n})$.*

Proof. We first observe that in (i), the existence of $\gamma, \delta \in \mathbb{R}_{++}$ is clear, while in (ii) it follows from Lemma 5.3 which asserts the existence of $\gamma, \delta \in \mathbb{R}_{++}$ such that

$$(\gamma + \delta)^2 < 4(\gamma + \alpha)(\delta + \beta). \quad (81)$$

Also, Lemma 5.3 implies that $\gamma + \alpha > 0$ and $\delta + \beta > 0$ in both cases (i) and (ii). By employing Proposition 3.7, $J_{\gamma A}$, $J_{\delta B}$ and, hence, $T_{A,B}$ and $T_{B,A}$ are single-valued and have full domain.

Next, Proposition 3.8 implies that R_1 and R_2 are conically θ_1 -averaged and θ_2 -averaged, respectively, where

$$\theta_1 := \frac{\lambda\gamma}{2(\gamma + \alpha)} = \frac{\gamma + \delta}{2(\gamma + \alpha)} \quad \text{and} \quad \theta_2 := \frac{\mu\delta}{2(\delta + \beta)} = \frac{\gamma + \delta}{2(\delta + \beta)}. \quad (82)$$

Now, if (i) holds, then $\theta_1 = \theta_2 = 1$ and $\frac{1}{\kappa^*} = 1$; if (ii) holds, then $\theta_1\theta_2 < 1$ and

$$\frac{1}{\kappa^*} = \frac{2(\gamma + \delta)(\alpha + \beta)}{4(\gamma + \alpha)(\delta + \beta) - (\gamma + \delta)^2} = \frac{\theta_1 + \theta_2 - 2\theta_1\theta_2}{1 - \theta_1\theta_2}. \quad (83)$$

Thus, Corollary 2.6 implies that R_1R_2 and R_2R_1 are conically $\frac{1}{\kappa^*}$ -averaged. By invoking Proposition 2.2(i)&(ii), we conclude that $T_{A,B}$ and $T_{B,A}$ are conically $\frac{\kappa}{\kappa^*}$ -averaged. Finally, due to Proposition 5.2, the proof is complete. $\blacksquare$

Corollary 5.5 (DR for α -comonotone and β -comonotone operators). *Suppose that A and B are maximally α -comonotone and β -comonotone, respectively, that $\gamma = \delta \in \mathbb{R}_{++}$ and $\lambda = \mu = 2$. Suppose further that either*

(i) $\alpha = \beta = 0$ and $\kappa^* := 1$; or

(ii) $\alpha + \beta > 0$ and $\kappa^* := \gamma + \frac{\alpha\beta}{\alpha+\beta} > 0$.

Then the adaptive DR operators $T_{A,B}$ and $T_{B,A}$ are conically $\frac{\kappa}{\kappa^}$ -averaged and have full domain. Moreover, if $\text{zer}(A+B) \neq \emptyset$ and $\kappa < \kappa^*$, then, for any $(T, C) \in \{(T_{A,B}, \gamma A), (T_{B,A}, \delta B)\}$, every sequence $(x_n)_{n \in \mathbb{N}}$ generated by T converges weakly to a point $\bar{x} \in \text{Fix } T$ with $J_C \bar{x} \in \text{zer}(A+B)$ and the rate of asymptotic regularity of T is $o(1/\sqrt{n})$.*

Proof. We invoke Theorem 5.4 with $\gamma = \delta$ and $\lambda = \mu = 2$. $\blacksquare$

5.2. The case of α - and β -monotone operators

Convergence of the adaptive DR algorithm for α -monotone and β -monotone operators was provided in [11, Section 4]. In this section we revisit some of these results by employing conical averagedness. In comparison to [11], our new results (see Theorem 5.7) extend the admissible range for the parameters $\gamma, \delta, \lambda, \mu$, and κ , which guarantees the averagedness of T .

Similarly to Lemma 5.3, we begin our discussion with the existence of parameters.

Lemma 5.6 (existence of parameters). *Let $\alpha, \beta \in \mathbb{R}$ be such that $\alpha + \beta \geq 0$, and let $\gamma, \delta \in \mathbb{R}_{++}$. Set γ_0 as in (75) and set $\Delta := \gamma(1 + \gamma\alpha)(\alpha + \beta)$. Then the following assertions are equivalent:*

- (i) $(\gamma + \delta)^2 \leq 4\gamma\delta(1 + \gamma\alpha)(1 + \delta\beta)$.
- (ii) $\frac{1}{\gamma} > \gamma_0$ and $\frac{1}{\gamma}(1 + 2\gamma\alpha - 2\sqrt{\Delta}) \leq \frac{1}{\delta} \leq \frac{1}{\gamma}(1 + 2\gamma\alpha + 2\sqrt{\Delta})$.

Consequently, if $\alpha + \beta = 0$, then (i) and (ii) are equivalent to $1 + 2\gamma\alpha > 0$ and $\delta = \frac{\gamma}{1 + 2\gamma\alpha}$; if $\alpha + \beta > 0$, then there always exist $\gamma, \delta \in \mathbb{R}_{++}$ such that

$$\frac{1}{\gamma} > \gamma_0 \quad \text{and} \quad \max \left\{ 0, \frac{1}{\gamma}(1 + 2\gamma\alpha - 2\sqrt{\Delta}) \right\} < \frac{1}{\delta} < \frac{1}{\gamma}(1 + 2\gamma\alpha + 2\sqrt{\Delta}), \quad (84)$$

in which case, all inequalities in (i) and (ii) are strict. Moreover, if (i) or (ii) holds, then $1 + \gamma\alpha > 0$ and $1 + \delta\beta > 0$.

Proof. We note that (i) is equivalent to

$$\left(\frac{1}{\gamma} + \frac{1}{\delta} \right)^2 \leq 4 \left(\frac{1}{\gamma} + \alpha \right) \left(\frac{1}{\delta} + \beta \right), \quad (85)$$

while the last two inequalities in (ii) can be written in the form

$$\frac{1}{\gamma} + 2\alpha - 2\sqrt{\left(\frac{1}{\gamma} + \alpha \right) (\alpha + \beta)} \leq \frac{1}{\delta} \leq \frac{1}{\gamma} + 2\alpha + 2\sqrt{\left(\frac{1}{\gamma} + \alpha \right) (\alpha + \beta)}. \quad (86)$$

The proof is complete by invoking Lemma 5.3 with $\frac{1}{\gamma}$ and $\frac{1}{\delta}$. ■

Theorem 5.7 (aDR for α -monotone and β -monotone operators). *Suppose that A and B are maximally α -monotone and β -monotone, respectively, that λ, μ are set by (73), and that either*

- (i) $\alpha + \beta = 0$, $1 + 2\gamma\alpha > 0$, $\delta = \frac{\gamma}{1 + 2\gamma\alpha}$, and $\kappa^* := 1$; or
- (ii) $\alpha + \beta > 0$ and $\kappa^* := \frac{4\gamma\delta(1 + \gamma\alpha)(1 + \delta\beta) - (\gamma + \delta)^2}{2\gamma\delta(\gamma + \delta)(\alpha + \beta)} > 0$.

Then the adaptive DR operators $T_{A,B}$ and $T_{B,A}$ are conically $\frac{\kappa}{\kappa^}$ -averaged and have full domain. Moreover, if $\text{zer}(A + B) \neq \emptyset$ and $\kappa < \kappa^*$, then, for any $(T, C) \in \{(T_{A,B}, \gamma A), (T_{B,A}, \delta B)\}$, every sequence $(x_n)_{n \in \mathbb{N}}$ generated by T converges weakly to a point $\bar{x} \in \text{Fix } T$ with $J_C \bar{x} \in \text{zer}(A + B)$ and the rate of asymptotic regularity of T is $o(1/\sqrt{n})$.*

Proof. The existence of $\gamma, \delta \in \mathbb{R}_{++}$ in (i) is clear while in (ii) it follows from the existence of $\gamma, \delta \in \mathbb{R}_{++}$ such that

$$(\gamma + \delta)^2 < 4\gamma\delta(1 + \gamma\alpha)(1 + \delta\beta) \quad (87)$$

due to Lemma 5.6. Furthermore, in both cases (i) and (ii), Lemma 5.6 implies that $1 + \gamma\alpha > 0$ and $1 + \delta\beta > 0$. On the one hand, by Proposition 3.4, $J_{\gamma A}, J_{\delta B}$ and, hence, $T_{A,B}$ and $T_{B,A}$ are single-valued and have full domain. On the other hand, Proposition 3.5 implies that $\frac{1}{1-\lambda}R_1$ and $\frac{1}{1-\mu}R_2$ are conically θ_1 -averaged and conically θ_2 -averaged, respectively, where

$$\theta_1 := \frac{\lambda}{2(\lambda-1)(1+\gamma\alpha)} = \frac{\gamma+\delta}{2\delta(1+\gamma\alpha)} \quad \text{and} \quad \theta_2 := \frac{\mu}{2(\mu-1)(1+\delta\beta)} = \frac{\gamma+\delta}{2\gamma(1+\delta\beta)}. \quad (88)$$

Next, we observe that if (i) holds, then $\theta_1 = \theta_2 = 1$ and $1 = \frac{1}{\kappa^*}$; if (ii) holds, then $\theta_1\theta_2 < 1$ and

$$\frac{\theta_1 + \theta_2 - 2\theta_1\theta_2}{1 - \theta_1\theta_2} = \frac{\frac{1}{\theta_1} + \frac{1}{\theta_2} - 2}{\frac{1}{\theta_1\theta_2} - 1} = \frac{2\gamma\delta(\gamma+\delta)(\alpha+\beta)}{4\gamma\delta(1+\gamma\alpha)(1+\delta\beta) - (\gamma+\delta)^2} = \frac{1}{\kappa^*}. \quad (89)$$

By applying Corollary 2.6 to $\frac{1}{1-\lambda}R_1$ and $\frac{1}{1-\mu}R_2$, we conclude that $R_2R_1 = \frac{1}{(1-\lambda)(1-\mu)}R_2R_1$ and $R_1R_2 = \frac{1}{(1-\lambda)(1-\mu)}R_1R_2$ are conically $\frac{1}{\kappa^*}$ -averaged. By invoking Proposition 2.2(i)&(ii), $T_{A,B} = (1-\kappa)\text{Id} + \kappa R_2R_1$ and $T_{B,A} = (1-\kappa)\text{Id} + \kappa R_1R_2$ are conically $\frac{\kappa}{\kappa^*}$ -averaged. Finally, by recalling Proposition 5.2, the proof is complete. $\blacksquare$

Theorem 5.7 sheds new light on the convergence analysis of the adaptive DR algorithm for two generalized monotone operators [11, Section 4]. More specifically, Theorem 5.7(i) recovers the convergence of the classical DR algorithm for two maximally monotone operators, see, e.g., [22]. Although Theorem 5.7(i) is the case $\alpha + \beta = 0$ in [11, Theorem 4.5], our current alternative proof employs conical averagedness and applies concurrently to both adaptive DR operators $T_{A,B}$ and $T_{B,A}$. In addition, Theorem 5.7(ii) presents an improvement for [11, Theorem 4.5(i)] on the parameter range, as shown in the following remark.

Remark 5.8 (aDR for the case $\alpha + \beta > 0$). Theorem 5.7(ii) readily implies and extends [11, Theorem 4.5(i)] in terms of parameter ranges. Indeed, in order to obtain convergence of the adaptive DR algorithm in the case where $\alpha + \beta > 0$, in Theorem 5.7(ii) it is only required that

$$0 < \kappa < \kappa^*, \quad (90)$$

while the assumptions of [11, Theorem 4.5(i)] are equivalent to the more restrictive conditions

$$1 + 2\gamma\alpha > 0 \quad \text{and} \quad 0 < \kappa < 1 \leq \kappa^*. \quad (91)$$

In order to verify (91), we first claim that

$$\kappa^* \geq 1 \iff \frac{1-2\gamma\beta}{\gamma} \leq \frac{1}{\delta} \leq \frac{1+2\gamma\alpha}{\gamma} \iff \mu \in [2-2\gamma\beta, 2+2\gamma\alpha]. \quad (92)$$

Since $\alpha + \beta > 0$, it follows that

$$\kappa^* \geq 1 \iff 2\gamma\delta(\gamma+\delta)(\alpha+\beta) \leq 4\gamma\delta(1+\gamma\alpha)(1+\delta\beta) - (\gamma+\delta)^2 \quad (93a)$$

$$\iff (1+2\gamma\alpha-2\gamma\beta-4\gamma^2\alpha\beta)\delta^2 - 2\gamma(1+\gamma\alpha-\gamma\beta)\delta + \gamma^2 \leq 0 \quad (93b)$$

$$\iff (1+2\gamma\alpha-2\gamma\beta-4\gamma^2\alpha\beta) - 2(1+\gamma\alpha-\gamma\beta)\frac{\gamma}{\delta} + \frac{\gamma^2}{\delta^2} \leq 0 \quad (93c)$$

$$\iff 1-2\gamma\beta \leq \frac{\gamma}{\delta} \leq 1+2\gamma\alpha. \quad (93d)$$

Since $\mu = 1 + \frac{\gamma}{\delta}$, we arrive at (92).

In the case where $\alpha + \beta > 0$, we recall that [11, Theorem 4.5(i)] requires $(\gamma, \delta, \lambda, \mu) \in \mathbb{R}_{++}^2 \times]1, +\infty[^2$ such that

$$1 + 2\gamma\alpha > 0, \quad (94a)$$

$$\mu \in [2 - 2\gamma\beta, 2 + 2\gamma\alpha], \quad (94b)$$

$$(\lambda - 1)(\mu - 1) = 1, \quad \delta = (\lambda - 1)\gamma. \quad (94c)$$

and $\kappa \in]0, 1[$. We note that (94c) is equivalent to (73), and that (94b) is equivalent to $\kappa^* \geq 1$ due to (92). Therefore, the assumptions of [11, Theorem 4.5(i)] for the case $\alpha + \beta > 0$ can be rewritten as (91).

The following corollary recovers the main convergence result in [11, Theorem 4.5(ii)].

Corollary 5.9 (DR for α -monotone and β -monotone operators). *Suppose that A and B are maximally α -monotone and β -monotone, respectively, that $\gamma = \delta \in \mathbb{R}_{++}$ and $\lambda = \mu = 2$, and that either*

- (i) $\alpha = \beta = 0$ and $\kappa^* := 1$; or
- (ii) $\alpha + \beta > 0$ and $\kappa^* := 1 + \gamma \frac{\alpha\beta}{\alpha+\beta} > 0$.

Then the DR operators $T_{A,B}$ and $T_{B,A}$ are conically $\frac{\kappa}{\kappa^}$ -averaged and have full domain. Moreover, if $\text{zer}(A + B) \neq \emptyset$ and $\kappa < \kappa^*$, then, for any $(T, C) \in \{(T_{A,B}, \gamma A), (T_{B,A}, \delta B)\}$, every sequence $(x_n)_{n \in \mathbb{N}}$ generated by T converges weakly to a point $\bar{x} \in \text{Fix } T$ with $J_C \bar{x} \in \text{zer}(A + B)$ and the rate of asymptotic regularity of T is $o(1/\sqrt{n})$.*

Proof. We employ Theorem 5.7 with $\gamma = \delta$ and $\lambda = \mu = 2$. ■

Remark 5.10 (DR for the case $\alpha + \beta > 0$). In the classical setup, the DR operator is defined as a strict convex combination of Id and $R_2 R_1$ (or $R_1 R_2$), i.e., $\kappa \in]0, 1[$. In Corollary 5.9, suppose that both operators A and B are strongly monotone ($\alpha > 0$ and $\beta > 0$), then the upper bound for κ is $\kappa^* = 1 + \gamma \frac{\alpha\beta}{\alpha+\beta} > 1$ which implies that κ can be chosen to be larger than 1 and the DR still converges. On the other hand, if either A or B is weakly monotone (i.e., $\alpha < 0$ or $\beta < 0$), then $\kappa^* < 1$, which means that one needs to further restrict κ than the standard range $]0, 1[$ in order to obtain the convergence.

For the remainder of this section we consider the adaptive DR algorithm for the problem of minimizing the sum of two functions. Let $f: X \rightarrow]-\infty, +\infty]$. Then f is *proper* if $\text{dom } f := \{x \in X \mid f(x) < +\infty\} \neq \emptyset$, and *lower semicontinuous* if $\forall x \in \text{dom } f, f(x) \leq \liminf_{z \rightarrow x} f(z)$. Set $\alpha \in \mathbb{R}$. The function f is α -convex (see [31, Definition 4.1]) if $\forall x, y \in \text{dom } f, \forall \kappa \in]0, 1[$,

$$f((1 - \kappa)x + \kappa y) + \frac{\alpha}{2} \kappa(1 - \kappa) \|x - y\|^2 \leq (1 - \kappa)f(x) + \kappa f(y). \quad (95)$$

If (95) holds with $\alpha = 0$, then we say that f is *convex*, $\alpha > 0$, then we say that f is *strongly convex*, $\alpha < 0$, then we say that f is *weakly convex*. The *proximity operator* of a proper function f with parameter $\gamma \in \mathbb{R}_{++}$ is the mapping $\text{Prox}_{\gamma f}: X \rightrightarrows X$ defined by

$$\forall x \in X, \quad \text{Prox}_{\gamma f}(x) := \underset{z \in X}{\operatorname{argmin}} \left(f(z) + \frac{1}{2\gamma} \|z - x\|^2 \right). \quad (96)$$

Finally, we consider the (α, β) -convex minimization problem of the form

$$\min_{x \in X} (f(x) + g(x)), \quad (97)$$

where f and g are α -convex and β -convex functions, respectively. This problem arises in several important applications, see, for example, [19]. The adaptive DR algorithm for problem (97) incorporates the operators

$$T_{f,g} := (1 - \kappa) \text{Id} + \kappa R_g R_f \quad \text{and} \quad T_{g,f} := (1 - \kappa) \text{Id} + \kappa R_f R_g, \quad (98)$$

where

$$R_f := (1 - \lambda) \text{Id} + \lambda \text{Prox}_{\gamma f} \quad \text{and} \quad R_g := (1 - \mu) \text{Id} + \mu \text{Prox}_{\delta g}. \quad (99)$$

We continue our analysis in order to show that the relations between the proximal operators, the resolvents of the subdifferentials and the problem of finding zeros of the sum of the operators are useful in producing a solution of the minimization problem. To this end we recall that the Fréchet subdifferential of f at x is defined by

$$\hat{\partial}f(x) := \left\{ u \in X \mid \liminf_{z \rightarrow x} \frac{f(z) - f(x) - \langle u, z - x \rangle}{\|z - x\|} \geq 0 \right\}. \quad (100)$$

For additional details on various subdifferentials and related properties, we refer the reader to the monograph [25].

Lemma 5.11 (proximity operators of α -convex functions). *Let $f: X \rightarrow]-\infty, +\infty]$ be a proper, lower semicontinuous and α -convex function. Let $\gamma \in \mathbb{R}_{++}$ be such that $1 + \gamma\alpha > 0$. Then*

- (i) $\hat{\partial}f$ is maximally α -monotone.
- (ii) $\text{Prox}_{\gamma f} = J_{\gamma \hat{\partial}f}$ is single-valued and has full domain.

Proof. See [11, Lemma 5.2]. ■

Theorem 5.12 (aDR for (α, β) -convex minimization). *Let $f: X \rightarrow]-\infty, +\infty]$ and $g: X \rightarrow]-\infty, +\infty]$ be proper and lower semicontinuous. Suppose that f and g are α -convex and β -convex, respectively, that λ, μ are set by (73), and that either*

- (i) $\alpha + \beta = 0$, $1 + 2\gamma\alpha > 0$, $\delta = \frac{\gamma}{1+2\gamma\alpha}$, $\kappa^* := 1$; or
- (ii) $\alpha + \beta > 0$, $\kappa^* := \frac{4\gamma\delta(1+\gamma\alpha)(1+\delta\beta) - (\gamma+\delta)^2}{2\gamma\delta(\gamma+\delta)(\alpha+\beta)} > 0$.

Then the adaptive DR operators $T_{f,g}$ and $T_{g,f}$ are conically $\frac{\kappa}{\kappa^}$ -averaged and have full domain. Moreover, if $\text{zer}(\hat{\partial}f + \hat{\partial}g) \neq \emptyset$ and $\kappa < \kappa^*$, then, for any $(T, h) \in \{(T_{f,g}, \gamma f), (T_{g,f}, \delta g)\}$, every sequence $(x_n)_{n \in \mathbb{N}}$ generated by T converges weakly to a point $\bar{x} \in \text{Fix } T$ with $\text{Prox}_h(\bar{x}) \in \text{zer}(\hat{\partial}f + \hat{\partial}g) \subseteq \text{argmin}(f + g)$ and the rate of asymptotic regularity of T is $o(1/\sqrt{n})$.*

Proof. Similarly to the proof of Theorem 5.7, our assumptions imply that $1 + \gamma\alpha > 0$ and $1 + \delta\beta > 0$. Now, Lemma 5.11 implies that $\hat{\partial}f$ and $\hat{\partial}g$ are maximally α -monotone and β -monotone, respectively, where $\text{Prox}_{\gamma f} = J_{\gamma \hat{\partial}f}$ and $\text{Prox}_{\delta g} = J_{\delta \hat{\partial}g}$. We also have from [11, Lemma 5.3] that $\text{zer}(\hat{\partial}f + \hat{\partial}g) \subseteq \text{argmin}(f + g)$. The conclusion then follows by applying Theorem 5.7 to $A = \hat{\partial}f$ and $B = \hat{\partial}g$. ■

As a consequence, we retrieve [11, Theorem 5.4(ii)] which, in turn, unifies and extends [19, Theorems 4.4 and 4.6] to the Hilbert space setting.

Corollary 5.13 (DR for (α, β) -convex minimization). *Let $f: X \rightarrow]-\infty, +\infty]$ and $g: X \rightarrow]-\infty, +\infty]$ be proper and lower semicontinuous. Suppose that f and g are α -convex and β -convex, respectively, that $\gamma = \delta \in \mathbb{R}_{++}$, $\lambda = \mu = 2$, and that either*

- (i) $\alpha = \beta = 0$, $\kappa^* := 1$; or
- (ii) $\alpha + \beta > 0$, $\kappa^* := 1 + \gamma \frac{\alpha\beta}{\alpha+\beta} > 0$.

Then the DR operators $T_{f,g}$ and $T_{g,f}$ are conically $\frac{\kappa}{\kappa^}$ -averaged and have full domain. Moreover, if $\text{zer}(\hat{\partial}f + \hat{\partial}g) \neq \emptyset$ and $\kappa < \kappa^*$, then, for any $(T, h) \in \{(T_{f,g}, \gamma f), (T_{g,f}, \delta g)\}$, every sequence $(x_n)_{n \in \mathbb{N}}$ generated by T converges weakly to a point $\bar{x} \in \text{Fix } T$ with $\text{Prox}_h(\bar{x}) \in \text{zer}(\hat{\partial}f + \hat{\partial}g) \subseteq \text{argmin}(f + g)$ and the rate of asymptotic regularity of T is $o(1/\sqrt{n})$.*

Proof. Apply Theorem 5.12 with $\gamma = \delta$ and $\lambda = \mu = 2$. ■

Acknowledgement

We would like to thank two anonymous referees for their careful review and helpful comments. SB was partially supported by a UMass Lowell faculty startup grant. MND was partially supported by Discovery Projects 160101537 and 190100555 from the Australian Research Council. HMP was partially supported by Autodesk, Inc. via a gift made to the Department of Mathematical Sciences, UMass Lowell.

References

- [1] J.-B. Baillon, R.E. Bruck, and S. Reich, On the asymptotic behavior of nonexpansive mappings and semigroups, *Houston J. Math.* **4**(1), 1–9 (1978).
- [2] H.H. Bauschke and P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed., Springer, Cham (2017).
- [3] H.H. Bauschke, W.M. Moursi, and X. Wang, Generalized monotone operators and their averaged resolvents, *Math. Program., Ser. B* (2020).
- [4] H.H. Bauschke, D. Noll, and H.M. Phan, Linear and strong convergence of algorithms involving averaged nonexpansive operators, *J. Math. Anal. Appl.* **421**(1), 1–20 (2015).
- [5] J. Borwein, S. Reich, and I. Shafrir, Krasnoselskii–Mann iterations in normed spaces, *Canad. Math. Bull.* **35**(1), 21–28 (1992).
- [6] R.E. Bruck and S. Reich, Nonexpansive projections and resolvents of accretive operators in Banach spaces, *Houston J. Math.* **3**(4), 459–470, (1977).
- [7] R.S. Burachik and A.N. Iusem, *Set-Valued Mappings and Enlargements of Monotone Operators*, Springer, New York (2008).
- [8] P.L. Combettes, Solving monotone inclusions via compositions of nonexpansive averaged operators, *Optimization* **53**(5–6), 475–504 (2004).
- [9] P.L. Combettes and T. Pennanen, Proximal methods for cohypomonotone operators, *SIAM J. Control Optim.* **43**(2), 731–742 (2004).

- [10] P.L. Combettes and I. Yamada, Compositions and convex combinations of averaged nonexpansive operators, *J. Math. Anal. Appl.* **425**(1), 55–70 (2015).
- [11] M.N. Dao and H.M. Phan, Adaptive Douglas–Rachford splitting algorithm for the sum of two operators, *SIAM J. Optim.* **29**(4), 2697–2724 (2019).
- [12] M.N. Dao and H.M. Phan, Computing the resolvent of the sum of operators with application to best approximation problems, *Optim. Letters* **14**(5), 1193–1205 (2020).
- [13] M.N. Dao and M.K. Tam, Union averaged operators with applications to proximal algorithms for min-convex functions, *J. Optim. Theory Appl.* **181**(1), 61–94 (2019).
- [14] J. Douglas and H.H. Rachford, On the numerical solution of heat conduction problems in two and three space variables, *Trans. Amer. Math. Soc.* **82**, 421–439 (1956).
- [15] J. Eckstein and D.P. Bertsekas, On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators, *Math. Program., Ser. A* **55**(3), 293–318 (1992).
- [16] P. Giselsson, Tight global linear convergence rate bounds for Douglas–Rachford splitting, *J. Fixed Point Theory Appl.* **19**(4), 2241–2270 (2017).
- [17] K. Goebel and S. Reich, *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings*, Marcel Dekker, New York and Basel (1984).
- [18] A.A. Goldstein, Convex programming in Hilbert space. *Bull. Amer. Math. Soc.* **70**, 709–710 (1964).
- [19] K. Guo, D. Han, and X. Yuan, Convergence analysis of Douglas–Rachford splitting method for “strongly + weakly” convex programming, *SIAM J. Numer. Anal.* **55**(4), 1549–1577 (2017).
- [20] V.I. Kolobov, S. Reich, and R. Zalas, Weak, strong, and linear convergence of a double-layer fixed point algorithm, *SIAM J. Optim.* **27**(3), 1431–1458, (2017).
- [21] M. A. Krasnosel’skiĭ, Two remarks on the method of successive approximations, *Uspekhi Mat. Nauk* **10**, 123–127 (1955).
- [22] P.-L. Lions and B. Mercier, Splitting algorithms for the sum of two nonlinear operators, *SIAM J. Numer. Anal.* **16**(6), 964–979 (1979).
- [23] W.R. Mann, Mean value methods in iteration, *Proc. Amer. Math. Soc.*, **4**, 506–510 (1953).
- [24] B. Martinet, Régularisation d’inéquations variationnelles par approximations successives, *Rev. Française Informat. Recherche Opérationnelle* **4**(R-3), 154–158 (1970).
- [25] B.S. Mordukhovich, *Variational Analysis and Generalized Differentiation I. Basic Theory*, Springer, Berlin (2006).
- [26] N. Ogura and I. Yamada, Non-strictly convex minimization over the fixed point set of an asymptotically shrinking nonexpansive mapping, *Numer. Funct. Anal. Optim.* **23**(1–2) 113–137 (2002).
- [27] G.B. Passty, Ergodic convergence to a zero of the sum of monotone operators in Hilbert space, *J. Math. Anal. Appl.* **72**(2), 383–390 (1979).
- [28] S. Reich, Weak convergence theorems for nonexpansive mappings in Banach spaces, *J. Math. Anal. Appl.* **67**(2), 274–276 (1979).
- [29] R.T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM J. Control Optim.* **14**(5), 877–898 (1976).
- [30] M.K. Tam, Algorithms based on unions of nonexpansive maps, *Optim. Lett.* **12**(5), 1019–1027 (2018).
- [31] J.-P. Vial, Strong and weak convexity of sets and functions. *Math. Oper. Res.* **8**(2), 231–259 (1983).