Giselsson and Moursi *Fixed Point Theory Algorithms Sci Eng* (2021) 2021:25
https://doi.org/10.1186/s13663-021-00709-0

Fixed Point Theory and Algorithms for Sciences and Engineering

RESEARCH **Open Access**

On compositions of special cases of Lipschitz continuous operators

Pontus Giselsson[1] and Walaa M. Moursi[2,3*]

*Correspondence:
walaa.moursi@uwaterloo.ca
[2]Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada
[3]Mathematics Department, Faculty of Science, Mansoura University, Mansoura 35516, Egypt
Full list of author information is available at the end of the article

Abstract

Many iterative optimization algorithms involve compositions of special cases of Lipschitz continuous operators, namely firmly nonexpansive, averaged, and nonexpansive operators. The structure and properties of the compositions are of particular importance in the proofs of convergence of such algorithms. In this paper, we systematically study the compositions of further special cases of Lipschitz continuous operators. Applications of our results include compositions of scaled conically nonexpansive mappings, as well as the Douglas–Rachford and forward–backward operators, when applied to solve certain structured monotone inclusion and optimization problems. Several examples illustrate and tighten our conclusions.

MSC: Primary 47H05; 90C25; secondary 47H09; 49M27; 65K05; 65K10

Keywords: Compositions of operators; Conically nonexpansive operators; Douglas–Rachford algorithm; Forward-backward algorithm; Hypoconvex function; Maximally monotone operator; Proximal operator; Resolvent

1 Introduction

In this paper, we assume that

$$\boxed{X \text{ is a real Hilbert space}}$$

with the inner product $\langle \cdot \mid \cdot \rangle$ and the induced norm $\|\cdot\|$. Let $L > 0$ and let $T\colon X \to X$. Then T is *L-Lipschitz continuous* if $(\forall (x,y) \in X \times X)$ $\|Tx - Ty\| \le L\|x - y\|$, and T is *nonexpansive* if T is 1-Lipschitz continuous, i.e., $(\forall (x,y) \in X \times X)$ $\|Tx - Ty\| \le \|x - y\|$. In this paper, we study compositions of what we call (see Definition 3.1) identity-nonexpansive decompositions (I-N decompositions for short) of Lipschitz continuous operators. Let $(\alpha, \beta) \in \mathbb{R}^2$ and let $\mathrm{Id}\colon X \to X$ be the *identity operator* on X. A Lipschitz continuous operator R admits an (α, β)-I-N decomposition if $R = \alpha \mathrm{Id} + \beta N$ for some nonexpansive operator $N\colon X \to X$.

For instance, averaged,[1] conically nonexpansive,[2] and cocoercive[3] operators are all Lipschitz continuous operators that admit special I-N decompositions.

We consider compositions of the form

$$R = R_m \cdots R_1, \tag{1}$$

where $m \in \{2, 3, \ldots\}$, $I = \{1, \ldots, m\}$, and $(R_i)_{i\in I}$ is a family of Lipschitz continuous operators such that, for each $i \in I$, R_i admits an (α_i, β_i)-I-N decomposition. That is, $R_i = \alpha_i \operatorname{Id} + \beta_i N_i$ for all $i \in I$, where α_i and β_i are real numbers, and $N_i \colon X \to X$ are nonexpansive for all $i \in I$. A straightforward (and naive) conclusion is that the composition is Lipschitz continuous with a constant $\Pi_{i\in I}(|\alpha_i| + |\beta_i|)$. However, such a conclusion can be further refined when, for instance, each R_i is an averaged operator. Indeed, in this case it is known that the composition is an averaged (and not just Lipschitz continuous) operator (see, e.g., [2, Proposition 4.46], [6, Lemma 2.2], and [21, Theorem 3]). In this paper, we provide a systematic study of the structure of R under additional assumptions on the decomposition parameters.

Our main result is stated in Theorem 3.4. We show that, for $m = 2$, under a mild assumption on $(\alpha_1, \alpha_2, \beta_1, \beta_2)$ composition (1) is a scalar multiple of a conically nonexpansive operator. As a consequence of Theorem 3.4, we show in Theorem 4.2 that, under additional assumptions on the decomposition parameters, compositions of scaled conically nonexpansive mappings are scaled conically nonexpansive mappings, see also [1] for a relevant result.[4] Special cases of Theorem 4.2 include, e.g., compositions of averaged operators [2, Proposition 4.46] and compositions of averaged and negatively averaged operators [12].

Of particular interest are compositions R that are averaged, conically nonexpansive, or contractive. Let $x_0 \in X$. For an averaged (respectively contractive) operator R, the sequence $(R^k x_0)_{k\in\mathbb{N}}$ converges weakly (respectively strongly) towards a fixed point of R (if one exists) [2, Theorem 5.14]. For conically nonexpansive operators, a simple averaging trick gives an averaged operator with the same fixed point set as the conically nonexpansive operator. Iterating the new averaged operator yields a sequence that converges weakly to a fixed point of the conically nonexpansive operator. These properties have been instrumental in proving convergence for the Douglas–Rachford algorithm and the forward–backward algorithm. In this paper, we apply our composition result Theorem 4.2 to prove convergence of these splitting methods in new settings.

The Douglas–Rachford and forward–backward methods traditionally solve monotone inclusion problems of the form

$$\text{Find } x \in X \text{ such that } 0 \in Ax + Bx, \tag{2}$$

where $A \colon X \rightrightarrows X$ and $B \colon X \rightrightarrows X$ are maximally monotone, and, in the case of the forward–backward method, A is additionally assumed to be cocoercive. The Douglas–Rachford

---

[1] Let $T \colon X \to X$. Then T is α-*averaged* if $\alpha \in]0, 1[$ and nonexpansive $N \colon X \to X$ exists such that $T = (1 - \alpha) \operatorname{Id} + \alpha N$.

[2] Let $T \colon X \to X$. Then T is α-*conically nonexpansive* if $\alpha \in]0, \infty[$ and nonexpansive $N \colon X \to X$ exists such that $T = (1 - \alpha) \operatorname{Id} + \alpha N$.

[3] Let $T \colon X \to X$, and let $\beta > 0$. Then T is $\frac{1}{\beta}$-*cocoercive* if nonexpansive $N : X \to X$ exists such that $T = \frac{\beta}{2}(\operatorname{Id} + N)$.

[4] The paper [1] appeared online while putting the finishing touches on this paper. Partial results of this work were presented by the second author at the *Numerical Algorithms in Nonsmooth Optimization* workshop at Erwin Schrödinger International Institute for Mathematics and Physics (ESI) in Vienna in February 2019 and at the *Operator Splitting Methods in Data Analysis* workshop at the Flatiron Institute, in New York in March 2019. Both workshops predate [1].

method iterates the Douglas–Rachford map $T = \frac{1}{2}(\mathrm{Id} + R_{\gamma B}R_{\gamma A})$, where[5] $\gamma > 0$ is a positive step-size. The Douglas–Rachford map is an averaged map of the composition of reflected resolvents. The forward–backward method iterates the forward–backward map $T = J_{\gamma B}(\mathrm{Id} - \gamma A)$, where $\gamma > 0$ is a positive step-size. The forward–backward map is a composition of a resolvent and a forward-step.

In this paper, we show that for Douglas–Rachford splitting we need not impose monotonicity on the individual operators, but only on the sum, provided the sum is strongly monotone. The reflected resolvents $R_{\gamma A}$ and $R_{\gamma B}$ are negatively conically nonexpansive, the composition is conically nonexpansive, and a sufficient averaging gives an averaged map that converges to a fixed point when iterated. Relevant work appears in [9, 16], and [17].

More striking, for the forward–backward method, we show that it is sufficient that the sum is monotone (not strongly monotone as for DR). More specifically, we show that identity can be shifted between the two operators, while still guaranteeing averagedness of the forward–backward map $T = J_{\gamma B}(\mathrm{Id} - \gamma A)$. Indeed, the resolvent $J_{\gamma B}$ is cocoercive and the forward-step $(\mathrm{Id} - \gamma A)$ is scaled averaged. This implies that the composition is averaged (given restrictions on the cocoercivity and averagedness parameters). Moreover, when the sum is strongly monotone, again with no assumptions on monotonicity of the individual operators, we show that the forward–backward map is contractive. We also prove tightness of our contraction factor.

We also provide, in Theorem 4.7, a generalization of Theorem 4.2 to the setting in (1) of compositions of more than two operators. We assume that all R_i are scaled conically nonexpansive operators and provide conditions on the parameters that give a specific scaled conically nonexpansive representation of R. Our condition is symmetric in the individual operators and allows for one of them to be scaled conic, while the rest must be scaled averaged. This is in compliance with the $m = 2$ case in Theorem 4.2.

Finally, in Sect. 8, we provide graphical 2D-representations of different operator classes that admit I-N decompositions such as Lipschitz continuous operators, averaged operators, and cocoercive operators. We also provide 2D-representations of compositions of two such operator classes. Illustrations of the firmly nonexpansive ($\frac{1}{2}$-averaged) and nonexpansive operator classes have previously appeared in [10, 11], and illustrations of more operator classes that admit particular I-N decompositions and their compositions have appeared in [14, 24] and in early preprints of [15].

1.1 Organization and notation

The remainder of this paper is organized as follows: Sect. 2 presents useful facts and auxiliary results that are used throughout the paper. In Sect. 3, we present the main abstract results of the paper. Section 4 presents the main composition results of Lipschitz continuous operators that admit I-N decompositions, under mild assumptions on the decomposition parameters, as well as illustrative and limiting examples. In Sect. 5 and Sect. 6, we present applications of our composition results to the Douglas–Rachford and forward–backward algorithms, respectively. In Sect. 7 we present applications of our results to optimization problems. Finally, in Sect. 8, we provide graphical representations of many different I-N decompositions and their compositions.

---

[5] Let $A\colon XX$ be an operator. The *resolvent* of A, denoted by J_A, is defined by $J_A = (\mathrm{Id} + A)^{-1}$, and the *reflected resolvent* of A, denoted by R_A, is defined by $R_A = 2J_A - \mathrm{Id}$

The notation we use is standard and follows, e.g., [2] or [23].

2 Facts and auxiliary results

Let $\rho \in \mathbb{R}$. Let $A\colon X \to X$. Recall that A is *ρ-monotone* if $(\forall (x,u) \in \operatorname{gra} A)$ $(\forall (y,v) \in \operatorname{gra} A)$

$$\langle x - y \mid u - v\rangle \geq \rho \|x - y\|^2 \tag{3}$$

and is maximally ρ-monotone if any proper extension of $\operatorname{gra} A$ will violate (3). In passing we point out that A is (maximally) monotone (respectively ρ-hypomonotone, ρ-strongly monotone) if $\rho = 0$ (respectively $\rho < 0$, $\rho > 0$) see, e.g., [2, Chap. 20], [4, Definition 6.9.1], [7, Definition 2.2], and [23, Example 12.28].

Fact 2.1 *Let $A\colon X \rightrightarrows X$, let $B\colon X \rightrightarrows X$, let $\lambda \in \mathbb{R} \smallsetminus \{0\}$, and suppose that* $\operatorname{zer}(A + B) = (A + B)^{-1}(0) \neq \varnothing$. *Suppose that J_A and J_B are single-valued and that* $\operatorname{dom} J_A = \operatorname{dom} J_B = X$. *Set*

$$T = (1 - \lambda)\operatorname{Id} + \lambda R_B R_A. \tag{4}$$

Then T is single-valued, $\operatorname{dom} T = X$, *and*

$$\operatorname{zer}(A + B) = J_A(\operatorname{Fix} R_B R_A) = J_A(\operatorname{Fix} T). \tag{5}$$

Proof See [9, Lemma 4.1]. □

Proposition 2.2 *Let $A\colon X \to X$, let $B\colon X \rightrightarrows X$, and suppose that* $\operatorname{zer}(A + B) = (A + B)^{-1}(0) \neq \varnothing$. *Suppose that J_B is single-valued and that* $\operatorname{dom} J_B = X$. *Set*

$$T = J_B(\operatorname{Id} - A). \tag{6}$$

Then T is single-valued, $\operatorname{dom} T = X$, *and*

$$\operatorname{zer}(A + B) = \operatorname{Fix} T. \tag{7}$$

Proof The proof is similar to the proof of [2, Proposition 26.1(iv)].[6] Indeed, let $x \in X$. Then $x \in \operatorname{zer}(A + B) \Leftrightarrow -Ax \in Bx \Leftrightarrow (\operatorname{Id} - A)x \in (\operatorname{Id} + B)x \Leftrightarrow x = J_B(\operatorname{Id} - A)x = Tx$. □

Lemma 2.3 *Let $\lambda \in \mathbb{R}$, let $R_1\colon X \to X$, let $R_2\colon X \to X$, and set*

$$R(\lambda) = (1 - \lambda)\operatorname{Id} + \lambda R_2 R_1. \tag{8}$$

Let $(x, y) \in X \times X$. Then

$$\begin{aligned}
&\bigl\langle R(\lambda)x - R(\lambda)y \mid \bigl(\operatorname{Id} - R(\lambda)\bigr)x - \bigl(\operatorname{Id} - R(\lambda)\bigr)y\bigr\rangle \\
&\quad = (1 - 2\lambda)\bigl\langle x - y \mid \bigl(\operatorname{Id} - R(\lambda)\bigr)x - \bigl(\operatorname{Id} - R(\lambda)\bigr)y\bigr\rangle
\end{aligned}$$

[6] In passing, we mention that [2, Proposition 26.1(iv)] assume that A and B are maximally monotone, which is not required here. However, the proof is the same.

$$+ \lambda^2\big\langle(\mathrm{Id}+R_1)x-(\mathrm{Id}+R_1)y \mid (\mathrm{Id}-R_1)x-(\mathrm{Id}-R_1)y\big\rangle$$
$$+ \lambda^2\big\langle(\mathrm{Id}+R_2)R_1x-(\mathrm{Id}+R_2)R_1y \mid (\mathrm{Id}-R_2)R_1x-(\mathrm{Id}-R_2)R_1y\big\rangle. \tag{9}$$

Proof See Appendix A. □

Proposition 2.4 *Let $\alpha \in \mathbb{R}$, let $\beta \in \mathbb{R}$, let $N\colon X \to X$, and set $T = \alpha\mathrm{Id} + \beta N$. Let $(x,y) \in X \times X$. Then the following hold*:

$$\begin{aligned}
&\beta^2\big(\|x-y\|^2-\|Nx-Ny\|^2\big)\\
&\quad=\big(\beta^2-\alpha^2\big)\|x-y\|^2-\|Tx-Ty\|^2+2\alpha\langle x-y \mid Tx-Ty\rangle && \text{(10a)}\\
&\quad=\big(\beta^2-\alpha^2\big)\|x-y\|^2-(1-2\alpha)\|Tx-Ty\|^2+2\alpha\big\langle Tx-Ty \mid (\mathrm{Id}-T)x-(\mathrm{Id}-T)y\big\rangle && \text{(10b)}\\
&\quad=\big(\beta^2-\alpha(\alpha-1)\big)\|x-y\|^2-\Big((1-\alpha)\|Tx-Ty\|^2+\alpha\big\|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\big\|^2\Big). && \text{(10c)}
\end{aligned}$$

Proof Indeed, we have

$$\begin{aligned}
&\beta^2\big(\|x-y\|^2-\|Nx-Ny\|^2\big)\\
&\quad=\beta^2\|x-y\|^2-\big\|(Tx-\alpha x)-(Ty-\alpha y)\big\|^2 && \text{(11a)}\\
&\quad=\beta^2\|x-y\|^2-\big(\|Tx-Ty\|^2+\alpha^2\|x-y\|^2-2\alpha\langle Tx-Ty \mid x-y\rangle\big) && \text{(11b)}\\
&\quad=\big(\beta^2-\alpha^2\big)\|x-y\|^2-\big(\|Tx-Ty\|^2-2\alpha\langle Tx-Ty \mid x-y\rangle\big) && \text{(11c)}\\
&\quad=\big(\beta^2-\alpha^2+\alpha\big)\|x-y\|^2-\big((1-\alpha)\|Tx-Ty\|^2\\
&\qquad+\alpha\|Tx-Ty\|^2-2\alpha\langle Tx-Ty \mid x-y\rangle+\alpha\|x-y\|^2\big) && \text{(11d)}\\
&\quad=\big(\beta^2-\alpha(\alpha-1)\big)\|x-y\|^2-\Big((1-\alpha)\|Tx-Ty\|^2+\alpha\big\|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\big\|^2\Big). && \text{(11e)}
\end{aligned}$$

This proves (10a) and (10c) in view of (11c) and (11e). Finally, note that $(\beta^2-\alpha^2)\|x-y\|^2 - \|Tx-Ty\|^2+2\alpha\langle x-y \mid Tx-Ty\rangle = (\beta^2-\alpha^2)\|x-y\|^2-(1-2\alpha)\|Tx-Ty\|^2-2\alpha\|Tx-Ty\|^2 + 2\alpha\langle x-y \mid Tx-Ty\rangle = (\beta^2-\alpha^2)\|x-y\|^2-(1-2\alpha)\|Tx-Ty\|^2+2\alpha\langle Tx-Ty \mid (\mathrm{Id}-T)x-(\mathrm{Id}-T)y\rangle$. This proves (10b). □

Proposition 2.5 *Let $\alpha \in \mathbb{R}$, let $\beta \in \mathbb{R}$, let $N\colon X \to X$, and set $T = \alpha\mathrm{Id} + \beta N$. Let $(x,y) \in X \times X$. Then the following are equivalent*:

(i) *N is nonexpansive.*

(ii) $\|Tx-Ty\|^2-2\alpha\langle x-y \mid Tx-Ty\rangle \le (\beta^2-\alpha^2)\|x-y\|^2$.

(iii) $(1-2\alpha)\|Tx-Ty\|^2-2\alpha\langle Tx-Ty \mid (\mathrm{Id}-T)x-(\mathrm{Id}-T)y\rangle \le (\beta^2-\alpha^2)\|x-y\|^2$.

(iv) $(2\alpha-1)\|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\|^2-2(1-\alpha)\langle Tx-Ty \mid (\mathrm{Id}-T)x-(\mathrm{Id}-T)y\rangle \le (\beta^2-(1-\alpha)^2)\|x-y\|^2$.

(v) $(1-\alpha)\|Tx-Ty\|^2+\alpha\|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\|^2 \le (\beta^2-\alpha(\alpha-1))\|x-y\|^2$.

Proof (i)⇔(ii)⇔(iii)⇔(v): This is a direct consequence of Proposition 2.4. (i)⇔(iv): Applying (10b) with (T,α,β) replaced by $(\mathrm{Id}-T, 1-\alpha, -\beta)$ yields $\beta^2(\|x-y\|^2-\|Nx-Ny\|^2) = (\beta^2-(1-\alpha)^2)\|x-y\|^2-(2\alpha-1)\|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\|^2+2(1-\alpha)\langle Tx-Ty \mid (\mathrm{Id}-T)x-(\mathrm{Id}-T)y\rangle$. The proof is complete. □

Proposition 2.6 *Let $\alpha \in \mathbb{R}$, let $N: X \to X$, and set $T = (1-\alpha)\,\mathrm{Id} + \alpha N$. Let $(x,y) \in X \times X$. Then the following are equivalent*:

(i) *N is nonexpansive.*
(ii) $\|Tx - Ty\|^2 - 2(1-\alpha)\langle x - y \mid Tx - Ty\rangle \le (2\alpha - 1)\|x - y\|^2$.
(iii) $(2\alpha - 1)\|Tx - Ty\|^2 - 2(1-\alpha)\langle Tx - Ty \mid (\mathrm{Id} - T)x - (\mathrm{Id} - T)y\rangle \le (2\alpha - 1)\|x - y\|^2$.
(iv) $(1 - 2\alpha)\|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 \le 2\alpha\langle Tx - Ty \mid (\mathrm{Id} - T)x - (\mathrm{Id} - T)y\rangle$.
(v) $(1-\alpha)\|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 \le \alpha\|x - y\|^2 - \alpha\|Tx - Ty\|^2$.

Proof Apply Proposition 2.5 with (α, β) replaced by $(1-\alpha, \alpha)$. □

Lemma 2.7 *Let $\lambda < 1$. Then*

$$\|x\|^2 - \lambda\|y\|^2 \ge -\frac{\lambda}{1-\lambda}\|x + y\|^2. \tag{12}$$

Proof Let $\delta > 0$. By Young's inequality, $\|x + y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \ge (1-\delta)\|x\|^2 + (1-\delta^{-1})\|y\|^2$. Equivalently, $\|x + y\|^2 - (1-\delta)\|x\|^2 \ge (1-\delta^{-1})\|y\|^2$. Now, replace (x, y, δ) by $(-y, x + y, 1 - \lambda)$. □

Proposition 2.8 *Let $\alpha \in \,]0,1[$, let $\beta > 0$, and let $T: X \to X$. Then T is α-averaged if and only if $T = (1-\beta)\,\mathrm{Id} + \beta M$ and M is $\frac{\alpha}{\beta}$-conically nonexpansive.*

Proof Indeed, T is α-averaged if and only if there exists a nonexpansive mapping $N: X \to X$ such that $T = (1-\alpha)\mathrm{Id} + \alpha N$. Equivalently,

$$T = (1-\alpha)\mathrm{Id} + \alpha N = (1-\beta)\,\mathrm{Id} + \beta\left(\left(1 - \frac{\alpha}{\beta}\right)\mathrm{Id} + \frac{\alpha}{\beta}N\right),$$

and the conclusion follows by setting $M = (1 - \frac{\alpha}{\beta})\,\mathrm{Id} + \frac{\alpha}{\beta}N$. □

The following three lemmas can be directly verified, hence we omit the proof.

Lemma 2.9 *Let $\alpha > 0$, and let $T: X \to X$. Then T is α-conically nonexpansive $\Leftrightarrow$ $\mathrm{Id} - T$ is $\frac{1}{2\alpha}$-cocoercive $\Rightarrow$ $\mathrm{Id} - T$ is maximally monotone.*

Lemma 2.10 *Let $\beta > 0$, let $\mu \in \mathbb{R}$, and let $A: X \to X$. Suppose that A is maximally μ-monotone and $\frac{1}{\beta}$-cocoercive. Then $\mu \le \frac{1}{\beta}$.*

Lemma 2.11 *Let $\beta > 0$, let $T: X \to X$, and let $\overline{\beta} \ge \beta$. Suppose that T is $\frac{1}{\beta}$-cocoercive. Then T is $\frac{1}{\overline{\beta}}$-cocoercive.*

Lemma 2.12 *Let $\beta > 0$, and let $A: X \to X$. Suppose that A is β-Lipschitz continuous. Then the following hold*:

(i) *A is maximally $(-\beta)$-monotone.*
(ii) *$A + \beta\,\mathrm{Id}$ is $\frac{1}{2\beta}$-cocoercive.*

Proof See Appendix B. □

Lemma 2.13 *Let $\beta > \delta > 0$, let $T_1 \colon X \to X$, and let $T_2 \colon X \to X$. Suppose that T_1 (respectively T_2) is $\frac{1}{\beta}$-cocoercive (respectively $\frac{1}{\delta}$-cocoercive). Then $T_1 - T_2$ is β-Lipschitz continuous.*

Proof See Appendix C. □

As a corollary, we obtain the following result which was stated in [27, page 4].

Corollary 2.14 *Let $f_1 \colon X \to \mathbb{R}$, $f_2 \colon X \to \mathbb{R}$ be Frechét differentiable convex functions, and let $\beta > \delta > 0$. Suppose that ∇f_1 (respectively ∇f_2) is β-Lipschitz continuous (respectively δ-Lipschitz continuous). Then the following hold:*
 (i) *$\nabla f_1 - \nabla f_2$ is β-Lipschitz continuous.*
 (ii) *Suppose that $f_1 - f_2$ is convex. Then $\nabla f_1 - \nabla f_2$ is $\frac{1}{\beta}$-cocoercive.*

Proof See Appendix D. □

Lemma 2.15 *Let $\alpha \in]0,1[$, let $\delta \in]0,1]$, and let $T \colon X \to X$. Suppose that T is α-averaged. Then the following hold:*
 (i) *δT is $(1 - \delta(1-\alpha))$-averaged.*
 (ii) *Suppose that $\delta \in]0,1[$. Then δT is a Banach contraction with constant δ.*

Proof See Appendix E. □

Let A be maximally *ρ-monotone*, where $\rho > -1$. Then (see [9, Proposition 3.4] and [3, Corollary 2.11 and Proposition 2.12]) we have

$$J_A \text{ is single-valued and } \operatorname{dom} J_A = X. \tag{13}$$

The following result involves resolvents and reflected resolvents of ρ-monotone operators.

Proposition 2.16 *Let A be ρ-monotone, where $\rho > -1$. Then the following hold:*
 (i) *J_A is $(1+\rho)$- cocoercive, in which case J_A is Lipschitz continuous with constant $\frac{1}{1+\rho}$.*
 (ii) *$-R_A$ is $\frac{1}{1+\rho}$-conically nonexpansive.*
 (iii) *Suppose that $\rho \le 0$. Then R_A is Lipschitz continuous with constant $\frac{1-\rho}{1+\rho}$.*

Proof (i): See [9, Lemma 3.3(ii)]. Alternatively, it follows from [3, Corollary 3.8(ii)] that $\mathrm{Id} - T$ is $\frac{1}{2(1+\rho)}$-averaged. Now apply Lemma 2.9 with T replaced by $\mathrm{Id} - J_A$. (ii): It follows from (i) that there exists a nonexpansive operator $N \colon X \to X$ such that $J_A = \frac{1}{2(1+\rho)}(\mathrm{Id} + N)$. Now, $-R_A = \mathrm{Id} - 2J_A = \mathrm{Id} - \frac{1}{1+\rho}(\mathrm{Id} + N) = (1 - \frac{1}{1+\rho})\mathrm{Id} + \frac{1}{1+\rho}N$. (iii): Indeed, let $(x,y) \in X \times X$ and let N be as defined above. We have

$$\|R_A x - R_A y\| = \left\| -\frac{\rho}{1+\rho}(x-y) - \frac{1}{1+\rho}(Nx - Ny) \right\| \le -\frac{\rho}{1+\rho}\|x-y\| + \frac{1}{1+\rho}\|Nx - Ny\| \tag{14a}$$

$$\le \frac{1-\rho}{1+\rho}\|x-y\|. \tag{14b}$$

The proof is complete. □

3 Compositions

Definition 3.1 ((α,β)-I-N decomposition) Let $R\colon X \to X$ be Lipschitz continuous, and let[7] $(\alpha,\beta) \in \mathbb{R} \times \mathbb{R}_+$. We say that R admits an (α,β)-identity-nonexpansive (I-N) decomposition[8] if there exists a nonexpansive operator $N\colon X \to X$ such that $R = \alpha\,\mathrm{Id} + \beta N$.

Throughout the rest of this paper, we assume that

$$\boxed{R_1\colon X \to X \text{ and } R_2\colon X \to X \text{ are Lipschitz continuous operators.}}$$

Proposition 3.2 *Let $\alpha_1 \in \left]-\infty,1\right[$, let $\alpha_2 \in \left]-\infty,1\right[$, let $\beta_1 \in \mathbb{R}_+$, let $\beta_2 \in \mathbb{R}_+$, and suppose that $\alpha_2(\alpha_2-1) \le \beta_2^2$. Set*

$$\delta_1 = \frac{\alpha_1}{1-\alpha_1}\left(1 - \frac{(1-\alpha_2)^2-\beta_2^2}{1-\alpha_2}\right), \tag{15a}$$

$$\delta_2 = \frac{\alpha_2}{1-\alpha_2}, \tag{15b}$$

$$\delta_3 = 1 - \left(\frac{(1-\alpha_1)^2-\beta_1^2}{1-\alpha_1}\left(1 - \frac{(1-\alpha_2)^2-\beta_2^2}{1-\alpha_2}\right) + \frac{(1-\alpha_2)^2-\beta_2^2}{(1-\alpha_2)}\right). \tag{15c}$$

Suppose that R_1 admits an (α_1,β_1)-I-N decomposition and that R_2 admits an (α_2,β_2)-I-N decomposition. Then $(\forall (x,y) \in X \times X)$ we have

$$\begin{aligned}&\|R_2R_1x - R_2R_1y\|^2 + \delta_1\big\|(\mathrm{Id}-R_1)x - (\mathrm{Id}-R_1)y\big\|^2\\ &\quad + \delta_2\big\|(\mathrm{Id}-R_2)R_1x - (\mathrm{Id}-R_2)R_1y\big\|^2 \le \delta_3\|x-y\|^2.\end{aligned} \tag{16}$$

Proof Set $T_i = \frac{1}{2}(\mathrm{Id}+R_i) = \frac{1+\alpha_i}{2}\,\mathrm{Id} + \frac{\beta_i}{2}N_i$, and observe that by Proposition 2.5 applied with (T,α,β) replaced by $(T_i, \frac{1+\alpha_i}{2}, \frac{\beta_i}{2})$, $i \in \{1,2\}$, we have $(\forall (x,y) \in X \times X)$

$$\begin{aligned}&\big\langle T_ix - T_iy \mid (\mathrm{Id}-T_i)x - (\mathrm{Id}-T_i)y\big\rangle\\ &\quad \ge \frac{\alpha_i}{1-\alpha_i}\big\|(\mathrm{Id}-T_i)x - (\mathrm{Id}-T_i)y\big\|^2 + \frac{(1-\alpha_i)^2-\beta_i^2}{4(1-\alpha_i)}\|x-y\|^2.\end{aligned} \tag{17}$$

Equivalently,

$$\begin{aligned}&\big\langle (\mathrm{Id}+R_i)x - (\mathrm{Id}+R_i)y \mid (\mathrm{Id}-R_i)x - (\mathrm{Id}-R_i)y\big\rangle\\ &\quad \ge \frac{\alpha_i}{1-\alpha_i}\big\|(\mathrm{Id}-R_i)x - (\mathrm{Id}-R_i)y\big\|^2 + \frac{(1-\alpha_i)^2-\beta_i^2}{1-\alpha_i}\|x-y\|^2.\end{aligned} \tag{18}$$

Observe also that, because $\alpha_2 < 1$, we have

$$\alpha_2(\alpha_2-1) \le \beta_2^2 \quad \Leftrightarrow \quad 1 - \frac{(1-\alpha_2)^2-\beta_2^2}{1-\alpha_2} \ge 0. \tag{19}$$

---

[7] Here and elsewhere, we use $\mathbb{R}_+$ to denote the interval $[0,+\infty[$.

[8] The assumption that $\beta \in \mathbb{R}_+$ is not restrictive. Indeed, since N is nonexpansive, an operator admits an (α,β)-I-N decomposition if and only if it admits an $(\alpha,-\beta)$-I-N decomposition. This is the reason why we define it only for nonnegative β.

It follows from (18), applied with $i = 2$ and (x, y) replaced by (R_1x, R_1y) in (20c) and by $i = 1$ in (20f), in view of (19) that

$$
\begin{aligned}
&\|x-y\|^2 - \|R_2R_1x - R_2R_1y\|^2 \\
&= \|x-y\|^2 - \|R_1x - R_1y\|^2 + \|R_1x - R_1y\|^2 - \|R_2R_1x - R_2R_1y\|^2 && \text{(20a)} \\
&= \langle (\mathrm{Id} + R_1)x - (\mathrm{Id} + R_1)y \mid (\mathrm{Id} - R_1)x - (\mathrm{Id} - R_1)y \rangle \\
&\quad + \langle (\mathrm{Id} + R_2)R_1x - (\mathrm{Id} + R_2)R_1y \mid (\mathrm{Id} - R_2)R_1x - (\mathrm{Id} - R_2)R_1y \rangle && \text{(20b)} \\
&\geq \langle (\mathrm{Id} + R_1)x - (\mathrm{Id} + R_1)y \mid (\mathrm{Id} - R_1)x - (\mathrm{Id} - R_1)y \rangle \\
&\quad + \frac{\alpha_2}{1-\alpha_2}\big\|(\mathrm{Id} - R_2)R_1x - (\mathrm{Id} - R_2)R_1y\big\|^2 + \frac{(1-\alpha_2)^2 - \beta_2^2}{1-\alpha_2}\|R_1x - R_1y\|^2 && \text{(20c)} \\
&= \langle (\mathrm{Id} + R_1)x - (\mathrm{Id} + R_1)y \mid (\mathrm{Id} - R_1)x - (\mathrm{Id} - R_1)y \rangle \\
&\quad + \frac{\alpha_2}{1-\alpha_2}\big\|(\mathrm{Id} - R_2)R_1x - (\mathrm{Id} - R_2)R_1y\big\|^2 \\
&\quad + \frac{(1-\alpha_2)^2 - \beta_2^2}{1-\alpha_2}\big(\|x-y\|^2 - \langle (\mathrm{Id} + R_1)x - (\mathrm{Id} + R_1)y \mid (\mathrm{Id} - R_1)x - (\mathrm{Id} - R_1)y \rangle\big) \\
& && \text{(20d)} \\
&= \left(1 - \frac{(1-\alpha_2)^2 - \beta_2^2}{1-\alpha_2}\right)\langle (\mathrm{Id} + R_1)x - (\mathrm{Id} + R_1)y \mid (\mathrm{Id} - R_1)x - (\mathrm{Id} - R_1)y \rangle \\
&\quad + \frac{\alpha_2}{1-\alpha_2}\big\|(\mathrm{Id} - R_2)R_1x - (\mathrm{Id} - R_2)R_1y\big\|^2 + \frac{(1-\alpha_2)^2 - \beta_2^2}{1-\alpha_2}\|x-y\|^2 && \text{(20e)} \\
&\geq \left(1 - \frac{(1-\alpha_2)^2 - \beta_2^2}{1-\alpha_2}\right)\left(\frac{\alpha_1}{1-\alpha_1}\big\|(\mathrm{Id} - R_1)x - (\mathrm{Id} - R_1)y\big\|^2 + \frac{(1-\alpha_1)^2 - \beta_1^2}{1-\alpha_1}\|x-y\|^2\right) \\
&\quad + \frac{\alpha_2}{1-\alpha_2}\big\|(\mathrm{Id} - R_2)R_1x - (\mathrm{Id} - R_2)R_1y\big\|^2 + \frac{(1-\alpha_2)^2 - \beta_2^2}{1-\alpha_2}\|x-y\|^2 && \text{(20f)} \\
&= \frac{\alpha_1}{1-\alpha_1}\left(1 - \frac{(1-\alpha_2)^2 - \beta_2^2}{1-\alpha_2}\right)\big\|(\mathrm{Id} - R_1)x - (\mathrm{Id} - R_1)y\big\|^2 \\
&\quad + \frac{\alpha_2}{1-\alpha_2}\big\|(\mathrm{Id} - R_2)R_1x - (\mathrm{Id} - R_2)R_1y\big\|^2 \\
&\quad + \left(\frac{(1-\alpha_1)^2 - \beta_1^2}{1-\alpha_1}\left(1 - \frac{(1-\alpha_2)^2 - \beta_2^2}{1-\alpha_2}\right) + \frac{(1-\alpha_2)^2 - \beta_2^2}{1-\alpha_2}\right)\|x-y\|^2. && \text{(20g)}
\end{aligned}
$$

Rearranging yields the desired result. □

Theorem 3.3 *Let $\alpha_1 \in \left]-\infty, 1\right[$, let $\alpha_2 \in \left]-\infty, 1\right[$, let $\beta_1 \in \mathbb{R}_+$, let $\beta_2 \in \mathbb{R}_+$, and suppose that $\alpha_2(\alpha_2 - 1) \leq \beta_2^2$. Let δ_1, δ_2, and δ_3 be defined as in (15a)–(15c). Set*

$$\delta_4 = \frac{\delta_1\delta_2}{\delta_1 + \delta_2}, \tag{21}$$

and suppose that $\delta_1 + \delta_2 > 0$, that $\delta_3 - \delta_4 + \delta_3\delta_4 \geq 0$, and that $\delta_4 > -1$. Suppose that R_1 admits an (α_1, β_1)-I-N decomposition, and that R_2 admits an (α_2, β_2)-I-N decomposition. Then R_2R_1 admits an (α, β)-I-N decomposition, where

$$\alpha = \frac{\delta_4}{1 + \delta_4}, \qquad \beta = \frac{\sqrt{\delta_3 - \delta_4 + \delta_3\delta_4}}{1 + \delta_4}. \tag{22}$$

Proof Let $\underline{\delta} := \min(\delta_1, \delta_2)$, let $\bar{\delta} := \max(\delta_1, \delta_2)$, and let $\lambda := -\underline{\delta}/\bar{\delta}$ (i.e., $\lambda = -\delta_1/\delta_2$ if $\delta_1 \le \delta_2$, and $\lambda = -\delta_2/\delta_1$ if $\delta_1 \ge \delta_2$). Then Proposition 3.2 and Lemma 2.7 imply that

$$\begin{aligned}
&\delta_3\|x-y\|^2 - \|R_2R_1x - R_2R_1y\|^2 \\
&\quad \ge \delta_1\big\|(\mathrm{Id}-R_1)x - (\mathrm{Id}-R_1)y\big\|^2 + \delta_2\big\|(\mathrm{Id}-R_2)R_1x - (\mathrm{Id}-R_2)R_1y\big\|^2 && \text{(23a)}\\
&\quad = \bar{\delta}\Big(\frac{\delta_1}{\bar{\delta}}\big\|(\mathrm{Id}-R_1)x - (\mathrm{Id}-R_1)y\big\|^2 + \frac{\delta_2}{\bar{\delta}}\big\|(\mathrm{Id}-R_2)R_1x - (\mathrm{Id}-R_2)R_1y\big\|^2 && \text{(23b)}\\
&\quad \ge \bar{\delta}\left(-\frac{\lambda}{1-\lambda}\big\|(\mathrm{Id}-R_1)x - (\mathrm{Id}-R_1)y + (\mathrm{Id}-R_2)R_1x - (\mathrm{Id}-R_2)R_1y\big\|^2\right) && \text{(23c)}\\
&\quad = -\frac{\lambda\bar{\delta}}{1-\lambda}\big\|(\mathrm{Id}-R_2R_1)x - (\mathrm{Id}-R_2R_1)y\big\|^2 && \text{(23d)}\\
&\quad = \frac{\underline{\delta}\bar{\delta}}{\bar{\delta}+\underline{\delta}}\big\|(\mathrm{Id}-R_2R_1)x - (\mathrm{Id}-R_2R_1)y\big\|^2 && \text{(23e)}\\
&\quad = \delta_4\big\|(\mathrm{Id}-R_2R_1)x - (\mathrm{Id}-R_2R_1)y\big\|^2. && \text{(23f)}
\end{aligned}$$

Comparing (23a)–(23f) to Proposition 2.5 applied with T replaced by R_2R_1, we learn that there exist a nonexpansive operator $N\colon X \to X$ and $(\alpha, \beta) \in \mathbb{R}^2$ such that $R_2R_1 = \alpha\,\mathrm{Id} + \beta N$, where $\delta_3 = \frac{\beta^2+\alpha(1-\alpha)}{1-\alpha}$ and $\delta_4 = \frac{\alpha}{1-\alpha}$. Equivalently, $\alpha = \frac{\delta_4}{1+\delta_4}$, hence $\beta = \frac{\sqrt{\delta_3-\delta_4+\delta_3\delta_4}}{1+\delta_4}$, as claimed. □

Theorem 3.4 *Let* $\alpha_1 \in \mathbb{R}$, *let* $\alpha_2 \in \mathbb{R}$, *let* $\beta_1 > 0$, *let* $\beta_2 > 0$, *suppose that* $\alpha_1 + \beta_1 > 0$, *that* $\alpha_2 + \beta_2 > 0$, *and that either* $\frac{\beta_1\beta_2}{(\alpha_1+\beta_1)(\alpha_2+\beta_2)} < 1$ *or* $\max\{\frac{\beta_1}{\alpha_1+\beta_1}, \frac{\beta_2}{\alpha_2+\beta_2}\} = 1$. *Set*

$$\kappa = (\alpha_1+\beta_1)(\alpha_2+\beta_2), \tag{24a}$$

$$\theta = \begin{cases} \frac{\beta_1\alpha_2+\beta_2\alpha_1}{\alpha_1\alpha_2+\alpha_1\beta_2+\alpha_2\beta_1}, & \frac{\beta_1\beta_2}{(\alpha_1+\beta_1)(\alpha_2+\beta_2)} < 1; \\ 1, & \max\{\frac{\beta_1}{\alpha_1+\beta_1}, \frac{\beta_2}{\alpha_2+\beta_2}\} = 1. \end{cases} \tag{24b}$$

Suppose that R_1 *admits an* (α_1, β_1)*-I-N decomposition, and that* R_2 *admits an* (α_2, β_2)*-I-N decomposition. Then* $\theta \in\,]0, +\infty[$ *and* R_2R_1 *admits a* $(\kappa(1-\theta), \kappa\theta)$*-I-N decomposition, i.e.,* R_2R_1 *is* κ*-scaled* θ*-conically nonexpansive. That is, there exists a nonexpansive operator* $N\colon X \to X$ *such that*

$$R_2R_1 = \kappa(1-\theta)\,\mathrm{Id} + \kappa\theta N. \tag{25}$$

Proof Let $\theta_i = \frac{\beta_i}{\alpha_i+\beta_i} > 0$, and observe that

$$R_i = (\alpha_i+\beta_i)\big((1-\theta_i)\,\mathrm{Id} + \theta_i N_i\big), \quad i \in \{1,2\}. \tag{26}$$

Next, let $\widetilde{N}_2 = \frac{1}{\alpha_1+\beta_1}N_2 \circ (\alpha_1+\beta_1)\,\mathrm{Id}$, and note that $\widetilde{N}_2$ is nonexpansive. Now, set

$$\widetilde{R}_1 = (1-\theta_1)\,\mathrm{Id} + \theta_1 N_1, \qquad \widetilde{R}_2 = (1-\theta_2)\,\mathrm{Id} + \theta_2\widetilde{N}_2. \tag{27}$$

Then (26) and (27) yield

$$R_2R_1 = \big((\alpha_2+\beta_2)\big((1-\theta_2)\mathrm{Id}+\theta_2N_2\big)\big)\big((\alpha_1+\beta_1)\big((1-\theta_1)\mathrm{Id}+\theta_1N_1\big)\big) \tag{28a}$$

$$= (\alpha_1+\beta_1)(\alpha_2+\beta_2)\left(\frac{1}{\alpha_1+\beta_1}\big((1-\theta_2)\mathrm{Id}+\theta_2N_2\big)\right)\big((\alpha_1+\beta_1)\widetilde{R}_1\big) \tag{28b}$$

$$= (\alpha_1+\beta_1)(\alpha_2+\beta_2)\widetilde{R}_2\widetilde{R}_1. \tag{28c}$$

We proceed by cases. Case I: $\alpha_1\alpha_2 = 0$. Observe that $0 \in \{\alpha_1,\alpha_2\} \Leftrightarrow \max\{\frac{\beta_1}{\alpha_1+\beta_1},\frac{\beta_2}{\alpha_2+\beta_2}\} = \max\{\theta_1,\theta_2\} = 1$. The conclusion follows by observing that $\widetilde{R}_i$ is nonexpansive, $i \in \{1,2\}$.

Case II: $\alpha_1\alpha_2 \neq 0$. By assumption we must have $\frac{\beta_1}{\alpha_1+\beta_1}\frac{\beta_2}{\alpha_2+\beta_2} = \theta_1\theta_2 < 1$. We claim that $\widetilde{R}_i$, $i \in \{1,2\}$, satisfy the conditions of Theorem 3.3 with (α_i,β_i) replaced by $(1-\theta_i,\theta_i)$. Indeed, observe that $(1-\theta_2)(1-\theta_2-1) \le \theta_2^2 \Leftrightarrow \theta_2(\theta_2-1) \le \theta_2^2 \Leftrightarrow \theta_2 - 1 \le \theta_2$, which is always true. Moreover, replacing (α_i,β_i) by $(1-\theta_i,\theta_i)$ yields $\delta_1 = \frac{1-\theta_1}{\theta_1}$, $\delta_2 = \frac{1-\theta_2}{\theta_2}$, $\delta_3 = 1$, and, consequently, $\delta_4 = \frac{\theta_2(1-\theta_1)+\theta_1(1-\theta_2)}{(1-\theta_1)(1-\theta_2)}$. We claim that

$$\theta_1+\theta_2-2\theta_1\theta_2 > 0. \tag{29}$$

Indeed, recall that $\theta_1+\theta_2-2\theta_1\theta_2 = \theta_1\theta_2(\frac{1}{\theta_1}+\frac{1}{\theta_2}-2) > \theta_1\theta_2(\frac{1}{\theta_1}+\theta_1-2) = \theta_1\theta_2(\sqrt{\theta_1}-\frac{1}{\sqrt{\theta_1}})^2 > 0$. This implies that $\delta_1+\delta_2 = \frac{\theta_1+\theta_2-2\theta_1\theta_2}{\theta_1\theta_2} > 0$. Moreover,

$$\delta_4 = \frac{(1-\theta_1)(1-\theta_2)}{\theta_2(1-\theta_1)+\theta_1(1-\theta_2)} = \frac{1-\theta_1-\theta_2+\theta_1\theta_2}{\theta_1+\theta_2-2\theta_1\theta_2} = -1+\frac{1-\theta_1\theta_2}{\theta_1+\theta_2-2\theta_1\theta_2} > -1. \tag{30}$$

Therefore, by Theorem 3.3, we conclude that there exists a nonexpansive operator $N\colon X \to X$ such that $\widetilde{R}_2\widetilde{R}_1 = \alpha\,\mathrm{Id}+\beta N$, $\alpha = \frac{\delta_4}{1+\delta_4} = \frac{1-\theta_1-\theta_2+\theta_1\theta_2}{1-\theta_1\theta_2} = \frac{\alpha_1\alpha_2}{\alpha_1\alpha_2+\alpha_1\beta_2+\alpha_2\beta_1}$, and $\beta = \frac{1}{1+\delta_4} = \frac{\theta_1+\theta_2-2\theta_1\theta_2}{1-\theta_1\theta_2} = \frac{\beta_1\alpha_2+\beta_2\alpha_1}{\alpha_1\alpha_2+\alpha_1\beta_2+\alpha_2\beta_1}$. Now combine with (28a)–(28c). □

4 Applications to special cases

We start this section by recording the following simple lemma which can be easily verified, hence we omit the proof.

Lemma 4.1 *Set* $(\widetilde{R}_1,\widetilde{R}_2) = (-R_1, R_2\circ(-\mathrm{Id}))$. *Then the following hold*:

(i) $R_2R_1 = \widetilde{R}_2\widetilde{R}_1$.

(ii) *Let* $\alpha_i > 0$, *let* $\delta_i \in \mathbb{R}\smallsetminus\{0\}$, *and suppose that* $-\frac{1}{\delta_i}R_i$ *is* α_i*-conically nonexpansive. Then* $\frac{1}{\delta_i}\widetilde{R}_i$ *is* α_i*-conically nonexpansive.*

Theorem 4.2 *Let* $i \in \{1,2\}$, *let* $\alpha_i > 0$, *let* $\delta_i \in \mathbb{R}\smallsetminus\{0\}$, *let* $R_i\colon X \to X$ *be such that* $\frac{1}{\delta_i}R_i$ *is* α_i*-conically nonexpansive. Suppose that either* $\alpha_1\alpha_2 < 1$ *or* $\max\{\alpha_1,\alpha_2\} = 1$. *Set*

$$\begin{cases} \frac{\alpha_1+\alpha_2-2\alpha_1\alpha_2}{1-\alpha_1\alpha_2}, & \alpha_1\alpha_2 < 1; \\ 1, & \max\{\alpha_1,\alpha_2\} = 1. \end{cases} \tag{31}$$

Then there exists a nonexpansive operator $N\colon X \to X$ *such that*

$$R_2R_1 = \delta_1\delta_2\big((1-\alpha)\,\mathrm{Id}+\alpha N\big). \tag{32}$$

Furthermore, $\alpha < 1 \Leftrightarrow [\alpha_1 < 1$ *and* $\alpha_2 < 1]$.

Proof Set $(\widetilde{R}_1, \widetilde{R}_2) = (-R_1, R_2 \circ (-\operatorname{Id}))$ and set $R = R_2R_1$. The proof proceeds by cases.

CASE I: $\delta_i > 0$, $i \in \{1,2\}$. By assumption, there exist nonexpansive operators $N_i : X \to X$ such that $R_i = \delta_i(1-\alpha_i)\operatorname{Id} + \delta_i\alpha_i N_i$. Moreover, one can easily check that R_i satisfy the assumptions of Theorem 3.4 with (α_i, β_i) replaced by $(\delta_i(1-\alpha_i), \delta_i\alpha_i)$. Applying Theorem 3.4, with (α_i, β_i) replaced by $(\delta_i(1-\alpha_i), \delta_i\alpha_i)$, we learn that there exists a nonexpansive operator $N : X \to X$ such that $R = (\delta_1(1-\alpha_1) + \delta_1\alpha_1)(\delta_2(1-\alpha_2) + \delta_2\alpha_2)((1-\alpha)\operatorname{Id} + \alpha N) = \delta_1\delta_2((1-\alpha)\operatorname{Id} + \alpha N)$, where

$$\alpha = \frac{\delta_1(1-\alpha_1)\delta_2\alpha_2 + \delta_2(1-\alpha_2)\delta_1\alpha_1}{\delta_1(1-\alpha_1)\delta_2\alpha_2 + \delta_2(1-\alpha_2)\delta_1\alpha_1 + \delta_1(1-\alpha_1)\delta_2(1-\alpha_2)} = \frac{\alpha_1 + \alpha_2 - 2\alpha_1\alpha_2}{1-\alpha_1\alpha_2}. \tag{33}$$

Finally, observe that $\alpha < 1 \Leftrightarrow [\alpha_1\alpha_2 < 1$ and $\frac{\alpha_1+\alpha_2-2\alpha_1\alpha_2}{1-\alpha_1\alpha_2} < 1] \Leftrightarrow [\alpha_1\alpha_2 < 1$ and $1 - \alpha_1\alpha_2 > \alpha_1 + \alpha_2 - 2\alpha_1\alpha_2] \Leftrightarrow [\alpha_1\alpha_2 < 1$ and $(1-\alpha_1)(1-\alpha_2) > 0] \Leftrightarrow [\alpha_1 < 1$ and $\alpha_2 < 1]$.

CASE II: $\delta_i < 0$, $i \in \{1,2\}$. Observe that $\frac{1}{\delta_i}R_i = -\frac{1}{|\delta_i|}R_i$ is α_i-conically nonexpansive. Therefore, Lemma 4.1(ii), applied with δ_i replaced by $|\delta_i|$, implies that $\frac{1}{|\delta_i|}\widetilde{R}_i$ are α_i-conically nonexpansive. Now combine Lemma 4.1(i) and CASE I applied with (R_i, δ_i) replaced by $(\widetilde{R}_i, |\delta_i|)$.

CASE III: $\delta_1 < 0$ and $\delta_2 > 0$: Observe that $\frac{1}{\delta_1}R_1 = -\frac{1}{|\delta_1|}R_1$ is α_1-conically nonexpansive. Now, using Lemma 4.1(i)&(ii), we have $-R = -R_2R_1 = -\widetilde{R}_2\widetilde{R}_1$, and $-\frac{1}{\delta_2}\widetilde{R}_2$ is α_2-conically nonexpansive. Now combine with CASE II, applied with (R_1, R_2, δ_1) replaced by $(\widetilde{R}_1, -\widetilde{R}_2, |\delta_1|)$, to learn that there exists a nonexpansive mapping $N : X \to X$ such that $-R = |\delta_1|\delta_2((1-\alpha)\operatorname{Id} + \alpha N)$, and the conclusion follows.

CASE IV: $\delta_1 > 0$ and $\delta_2 < 0$: Indeed, $-R = -R_2R_1$. Now combine with CASE I applied with R_2 replaced by $-R_2$, in view of Lemma 4.1(ii). □

Corollary 4.3 *Let $\alpha \in \,]0,1[$, let $\beta > 0$, let $\delta \in \mathbb{R} \smallsetminus \{0\}$, let $\{i,j\} = \{1,2\}$, and suppose that $\frac{1}{\delta}R_i$ is α-averaged, and that R_j is $\frac{1}{\beta}$-cocoercive. Set $\overline{\alpha} = \frac{1}{2-\alpha}$. Then $\overline{\alpha} \in \,]0,1[$, and there exists a nonexpansive operator $N : X \to X$ such that*

$$R_2R_1 = \beta\delta\big((1-\overline{\alpha})\operatorname{Id} + \overline{\alpha}N\big). \tag{34}$$

Proof Suppose first that $(i,j) = (1,2)$, and observe that there exists a nonexpansive operator $\overline{N}$ such that $R_2 = \frac{\beta}{2}(\operatorname{Id} + \overline{N})$. Applying Theorem 4.7 with $m = 2$, $(\alpha_1, \alpha_2, \delta_1, \delta_2)$ replaced by $(\alpha, 1/2, \delta, \beta)$ yields that there exists a nonexpansive operator N such that $R_2R_1 = \beta\delta((1-\overline{\alpha})\operatorname{Id} + \overline{\alpha}N)$, where

$$\overline{\alpha} = \frac{\alpha + \frac{1}{2} - 2\frac{\alpha}{2}}{1 - \frac{\alpha}{2}} = \frac{1}{2-\alpha} \in \,]0,1[. \tag{35}$$

The case $(i,j) = (2,1)$ follows similarly. □

The assumption $\alpha_1\alpha_2 < 1$ is critical in the conclusion of Theorem 4.2 as we illustrate below.

Example 4.4 ($\alpha_1 = \alpha_2 > 1$) Let $\alpha > 1$, and set $R_1 = R_2 = (1-\alpha)\operatorname{Id} - \alpha\operatorname{Id} = (1-2\alpha)\operatorname{Id}$. Then

$$R_2R_1 = (1-2\alpha)^2\operatorname{Id} = \big(1 - 4\alpha + 4\alpha^2\big)\operatorname{Id}. \tag{36}$$

Hence, $\mathrm{Id} - R_2R_1 = 4\alpha(1-\alpha)\,\mathrm{Id}$. That is, $\mathrm{Id} - R_2R_1$ is not monotone; hence, R_2R_1 is *not* conically nonexpansive by Lemma 2.9 applied with T replaced by R_2R_1.

The following proposition provides an abstract framework to construct a family of operators R_1 and R_2 such that R_1 is α_1-conically nonexpansive, R_2 is α_2-conically nonexpansive, $\alpha_1\alpha_2 > 1$, and the composition R_2R_1 fails to be conically nonexpansive.

Proposition 4.5 *Let $\theta \in \mathbb{R}$, let $\alpha_1 > 0$, let $\alpha_2 > 0$, let*

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \tag{37}$$

set

$$R_1 = (1-\alpha_1)\,\mathrm{Id} + \alpha_1 R_\theta, \qquad R_2 = (1-\alpha_2)\,\mathrm{Id} - \alpha_2 R_\theta, \tag{38}$$

and set

$$\kappa = \alpha_1 + \alpha_2 - 2\alpha_1\alpha_2 \sin^2\theta - (\alpha_1 - \alpha_2)\cos\theta. \tag{39}$$

Then R_1 is α_1-conically nonexpansive, and R_2 is α_2-conically nonexpansive. Moreover, we have the implication $\kappa < 0 \Rightarrow R_2R_1$ is not *conically nonexpansive.*

Proof Set $S = R_{\pi/2}$, and observe that $S^2 = -\mathrm{Id}$ and that $R_\theta = (\cos\theta)\,\mathrm{Id} + (\sin\theta)S$. Now,

$$\begin{aligned} R_2R_1 &= \big((1-\alpha_1)\,\mathrm{Id} + \alpha_1 R_\theta\big)\big((1-\alpha_2)\,\mathrm{Id} - \alpha_2 R_\theta\big) && \text{(40a)}\\ &= (1-\alpha_1-\alpha_2+\alpha_1\alpha_2)\,\mathrm{Id} + (\alpha_1-\alpha_2)R_\theta - \alpha_1\alpha_2 R_{2\theta} && \text{(40b)}\\ &= \big(1-\alpha_1-\alpha_2+\alpha_1\alpha_2+(\alpha_1-\alpha_2)\cos\theta - \alpha_1\alpha_2\cos(2\theta)\big)\,\mathrm{Id} \\ &\quad + \big((\alpha_1-\alpha_2)\sin\theta - \alpha_1\alpha_2\sin(2\theta)\big)S && \text{(40c)}\\ &= \big(1-\alpha_1-\alpha_2+\alpha_1\alpha_2+(\alpha_1-\alpha_2)\cos\theta - \alpha_1\alpha_2\big(2\cos^2\theta - 1\big)\big)\,\mathrm{Id} \\ &\quad + \big((\alpha_1-\alpha_2)\sin\theta - \alpha_1\alpha_2\sin(2\theta)\big)S && \text{(40d)}\\ &= \big(1-\alpha_1-\alpha_2+2\alpha_1\alpha_2\sin^2\theta+(\alpha_1-\alpha_2)\cos\theta\big)\,\mathrm{Id} \\ &\quad + \big((\alpha_1-\alpha_2)\sin\theta - \alpha_1\alpha_2\sin(2\theta)\big)S. && \text{(40e)} \end{aligned}$$

Consequently,

$$\begin{aligned} \mathrm{Id} - R_2R_1 &= \big(\alpha_1+\alpha_2-2\alpha_1\alpha_2\sin^2\theta - (\alpha_1-\alpha_2)\cos\theta\big)\,\mathrm{Id} \\ &\qquad - \big((\alpha_1-\alpha_2)\sin\theta - \alpha_1\alpha_2\sin(2\theta)\big)S. \end{aligned} \tag{41}$$

Hence, $(\forall x \in \mathbb{R}^2)$

$$\big\langle (\mathrm{Id} - R_2R_1)x \mid x\big\rangle = \big(\alpha_1+\alpha_2-2\alpha_1\alpha_2\sin^2\theta - (\alpha_1-\alpha_2)\cos\theta\big)\|x\|^2 = \kappa\|x\|^2. \tag{42}$$

Now, R_2R_1 is conically nonexpansive $\Rightarrow$ $\mathrm{Id} - R_2R_1$ is monotone by Lemma 2.9, and the conclusion follows in view of (42). ☐

The following example provides two concrete instances where: (i) $\alpha_1 > 1$, $\alpha_2 > 1$, hence $\alpha_1\alpha_2 > 1$, (ii) $\alpha_1 > 1$, $\alpha_2 < 1$, $\alpha_1\alpha_2 > 1$. In both cases, R_2R_1 is *not* conically nonexpansive.

Example 4.6 Suppose that one of the following holds:

(i) $\theta \in]0, \pi/2[$, $\epsilon \geq 0$, $\delta \geq 0$, $\alpha_1 = \frac{1+\epsilon}{\sin^2\theta}$, and $\alpha_2 = \frac{1+\delta}{\sin^2\theta}$.

(ii) $\theta \in]\pi/4, \pi/2[$, $\epsilon > \frac{\cos^2\theta(2-\cos^2\theta)}{(1-2\cos^2\theta)(1+\cos\theta)+\cos\theta}$, $\alpha_1 = \frac{1+\epsilon}{\sin^2\theta}$, and $\alpha_2 = \sin^2\theta$.

Let R_θ be defined as in (37), let $R_1 = (1-\alpha_1)\,\mathrm{Id} + \alpha_1 R_\theta$, and let $R_2 = (1-\alpha_2)\,\mathrm{Id} - \alpha_2 R_\theta$. Then $\alpha_1\alpha_2 > 1$, and R_2R_1 is *not* conically nonexpansive.

Proof Let κ be defined as in (39). In view of Proposition 4.5, it is sufficient to show that $\kappa < 0$. (i): Note that $\kappa < 0 \Leftrightarrow \kappa\sin^2\theta < 0$. Now,

$$\begin{aligned}
\kappa\sin^2\theta &= 2 + \epsilon + \delta - (\epsilon - \delta)\cos\theta - 2 - 2\epsilon - 2\delta - 2\epsilon\delta && \text{(43a)}\\
&= -\big(\epsilon(1+\cos\theta) + \delta(1-\cos\theta) + 2\epsilon\delta\big) < 0. && \text{(43b)}
\end{aligned}$$

(ii): We have

$$\begin{aligned}
\kappa &= \frac{1+\epsilon+\sin^4\theta}{\sin^2\theta} - 2(1+\epsilon)\sin^2\theta - \frac{1+\epsilon-\sin^4\theta}{\sin^2\theta}\cos\theta && \text{(44a)}\\
&= -\frac{1}{\sin^2\theta}\big(2(1+\epsilon)\sin^4\theta - (1+\epsilon+\sin^4\theta) + (1+\epsilon-\sin^4\theta)\cos\theta\big) && \text{(44b)}\\
&= -\frac{1}{1-\cos^2\theta}\big((2\sin^4\theta + \cos\theta - 1)\epsilon + \sin^4\theta(1-\cos\theta) - (1-\cos\theta)\big) && \text{(44c)}\\
&= -\frac{1-\cos\theta}{1-\cos^2\theta}\big((2(1+\cos\theta)(1-\cos^2\theta) - 1)\epsilon + 1 - 2\cos^2\theta + \cos^4\theta - 1\big) && \text{(44d)}\\
&= -\frac{1}{1+\cos\theta}\big((1 + 2\cos\theta - 2\cos^2\theta - 2\cos^3\theta)\epsilon - \cos^2\theta(2-\cos^2\theta)\big) && \text{(44e)}\\
&= -\frac{1}{1+\cos\theta}\big(((1-2\cos^2\theta)(1+\cos\theta) + \cos\theta)\epsilon - \cos^2\theta(2-\cos^2\theta)\big). && \text{(44f)}
\end{aligned}$$

Now, observe that $(\forall\theta \in]\frac{\pi}{4}, \frac{\pi}{2}[)$ $1 - 2\cos^2\theta = -\cos(2\theta) > 0$. Consequently, $(1-2\cos^2\theta)(1+\cos\theta) + \cos\theta > \cos\theta > 0$. Now use the assumption $\epsilon > \frac{\cos^2\theta(2-\cos^2\theta)}{(1-2\cos^2\theta)(1+\cos\theta)+\cos\theta}$ to learn that $(1-2\cos^2\theta)(1+\cos\theta) + \cos\theta)\epsilon - \cos^2\theta(2-\cos^2\theta) > 0$, hence $\kappa < 0$, and the conclusion follows. ☐

Theorem 4.7 (composition of m scaled conically nonexpansive operators) *Let $m \geq 2$ be an integer, set $I = \{1, \dots, m\}$, let $(R_i)_{i\in I}$ be a family of operators from X to X, let $r \in I$, let α_i be real numbers such that $\{\alpha_i \mid i \in I \smallsetminus \{r\}\} \subseteq]0,1[$ and $\alpha_r > 0$, let δ_i be real numbers in $\mathbb{R} \smallsetminus \{0\}$, and suppose that, for every $i \in I$, $\frac{1}{\delta_i}R_i$ is α_i-conically nonexpansive. Set*

$$\overline{\alpha} = \frac{\sum_{\substack{i=1\\ i\neq r}}^{m} \frac{\alpha_i}{1-\alpha_i}}{1 + \sum_{\substack{i=1\\ i\neq r}}^{m} \frac{\alpha_i}{1-\alpha_i}}. \tag{45}$$

Suppose that $\alpha_r\overline{\alpha} < 1$, *and set*

$$\alpha = \begin{cases} \frac{\sum_{i=1}^{m} \frac{\alpha_i}{1-\alpha_i}}{1+\sum_{i=1}^{m} \frac{\alpha_i}{1-\alpha_i}}, & \alpha_r \neq 1; \\ 1, & \alpha_r = 1. \end{cases} \tag{46}$$

Then there exists a nonexpansive operator $N\colon X \to X$ *such that*

$$R_m \cdots R_1 = \delta_m \cdots \delta_1\big((1-\alpha)\,\mathrm{Id} + \alpha N\big). \tag{47}$$

Proof First, observe that $(\forall i \in I \smallsetminus \{r\})$, $\frac{1}{\delta_i}R_i$ is nonexpansive. If $\alpha_r = 1$, then $(\forall i \in \{1,\dots,m\})$ R_i is $|\delta_i|$-Lipschitz continuous and the conclusion readily follows. Now, suppose that $\alpha_r \neq 1$. We proceed by induction on $k \in \{2,\dots,m\}$. At $k = 2$, the claim holds by Theorem 4.2. Now, suppose that the claim holds for some $k \in \{2,\dots,m-1\}$. Let $(R_i)_{1\le i\le k+1}$ be a family of operators from X to X, let $\overline{r} \in \{1,\dots,k,k+1\}$, let α_i be real numbers such that $\{\alpha_i \mid i \in \{1,\dots,k,k+1\} \smallsetminus \{\overline{r}\}\} \subseteq]0,1[$ and $\alpha_{\overline{r}} \in]0,+\infty[\smallsetminus\{1\}$, let δ_i be real numbers in $\mathbb{R} \smallsetminus \{0\}$, and suppose that, for every $i \in \{1,\dots,k+1\}$, $\frac{1}{\delta_i}R_i$ is α_i-conically nonexpansive. Set $\overline{\beta} = \frac{\sum_{i=1, i\neq\overline{r}}^{k+1} \frac{\alpha_i}{1-\alpha_i}}{1+\sum_{i=1, i\neq\overline{r}}^{k+1} \frac{\alpha_i}{1-\alpha_i}}$, and suppose that $\alpha_{\overline{r}}\overline{\beta} < 1$. We examine two cases.

Case I: $\alpha_{k+1} = \alpha_{\overline{r}}$. In this case the conclusion follows by applying Theorem 4.2 in view of the inductive hypothesis with (R_1,R_2) replaced by $(R_k \dots R_1, R_{k+1})$ and $(\delta_1,\delta_2,\alpha_1,\alpha_2)$ replaced by $(\delta_1\dots\delta_k, \delta_{k+1}, (\sum_{i=1}^{k} \frac{\alpha_i}{1-\alpha_i})/(1+\sum_{i=1}^{k} \frac{\alpha_i}{1-\alpha_i}), \alpha_{k+1})$.

Case II: $\alpha_{k+1} \neq \alpha_{\overline{r}}$. We claim that

$$\alpha_{k+1} \frac{\sum_{i=1}^{k} \frac{\alpha_i}{1-\alpha_i}}{1+\sum_{i=1}^{k} \frac{\alpha_i}{1-\alpha_i}} < 1. \tag{48}$$

To this end, set $\hat{\alpha} = \frac{\sum_{i=1, i\neq\overline{r}}^{k} \frac{\alpha_i}{1-\alpha_i}}{1+\sum_{i=1, i\neq\overline{r}}^{k} \frac{\alpha_i}{1-\alpha_i}}$, and observe that $\hat{\alpha} < \overline{\beta}$. By assumption we have $\alpha_{\overline{r}}\overline{\beta} < 1$. Altogether, we conclude that $\alpha_{\overline{r}}\hat{\alpha} < 1$. It follows from the inductive hypothesis that

$$\frac{1}{\delta_1\dots\delta_k}(R_k\dots R_1) \text{ is } \frac{\sum_{i=1}^{k} \frac{\alpha_i}{1-\alpha_i}}{1+\sum_{i=1}^{k} \frac{\alpha_i}{1-\alpha_i}}\text{-conically nonexpansive.} \tag{49}$$

Next note that

$$\frac{\sum_{i=1}^{k} \frac{\alpha_i}{1-\alpha_i}}{1+\sum_{i=1}^{k} \frac{\alpha_i}{1-\alpha_i}} = \frac{\frac{\sum_{i=1, i\neq\overline{r}}^{k} \frac{\alpha_i}{1-\alpha_i} + \frac{\alpha_{\overline{r}}}{1-\alpha_{\overline{r}}}}{1+\sum_{i=1, i\neq\overline{r}}^{k} \frac{\alpha_i}{1-\alpha_i}}}{\frac{1+\sum_{i=1, i\neq\overline{r}}^{k} \frac{\alpha_i}{1-\alpha_i} + \frac{\alpha_{\overline{r}}}{1-\alpha_{\overline{r}}}}{1+\sum_{i=1, i\neq\overline{r}}^{k} \frac{\alpha_i}{1-\alpha_i}}} = \frac{\hat{\alpha} + \frac{\alpha_{\overline{r}}}{(1-\alpha_{\overline{r}})(1+\sum_{i=1, i\neq\overline{r}}^{k} \frac{\alpha_i}{1-\alpha_i})}}{1 + \frac{\alpha_{\overline{r}}}{(1-\alpha_{\overline{r}})(1+\sum_{i=1, i\neq\overline{r}}^{k} \frac{\alpha_i}{1-\alpha_i})}} \tag{50a}$$

$$= \frac{\hat{\alpha}(1-\alpha_{\overline{r}})(1+\sum_{i=1, i\neq\overline{r}}^{k} \frac{\alpha_i}{1-\alpha_i}) + \alpha_{\overline{r}}}{(1-\alpha_{\overline{r}})(1+\sum_{i=1, i\neq\overline{r}}^{k} \frac{\alpha_i}{1-\alpha_i}) + \alpha_{\overline{r}}} \tag{50b}$$

$$= \frac{\alpha_{\bar{r}}(1-\hat{\alpha}(1+\sum_{\substack{i=1\\i\neq\bar{r}}}^{k}\frac{\alpha_i}{1-\alpha_i})) + \hat{\alpha}(1+\sum_{\substack{i=1\\i\neq\bar{r}}}^{k}\frac{\alpha_i}{1-\alpha_i})}{1+(1-\alpha_{\bar{r}})\sum_{\substack{i=1\\i\neq\bar{r}}}^{k}\frac{\alpha_i}{1-\alpha_i}}. \tag{50c}$$

Because $\alpha_{\bar{r}}\overline{\beta} < 1$, we learn that $1+(1-\alpha_{\bar{r}})\sum_{\substack{i=1\\i\neq\bar{r}}}^{k}\frac{\alpha_i}{1-\alpha_i} > 0$. Moreover, because $\hat{\alpha} < 1$, we have $\alpha_{k+1}\hat{\alpha} < 1$. Therefore, (50a)–(50c) implies

$$\alpha_{k+1}\frac{\sum_{i=1}^{k}\frac{\alpha_i}{1-\alpha_i}}{1+\sum_{i=1}^{k}\frac{\alpha_i}{1-\alpha_i}} < 1 \tag{51a}$$

$$\begin{aligned}
\Leftrightarrow\quad & \alpha_{k+1}\left(\alpha_{\bar{r}}\left(1-\hat{\alpha}\left(1+\sum_{\substack{i=1\\i\neq\bar{r}}}^{k}\frac{\alpha_i}{1-\alpha_i}\right)\right)+\hat{\alpha}\left(1+\sum_{\substack{i=1\\i\neq\bar{r}}}^{k}\frac{\alpha_i}{1-\alpha_i}\right)\right) \\
& < 1+(1-\alpha_{\bar{r}})\sum_{\substack{i=1\\i\neq\bar{r}}}^{k}\frac{\alpha_i}{1-\alpha_i} \qquad (51\text{b})\\
\Leftrightarrow\quad & \alpha_{\bar{r}}\left(\alpha_{k+1}\left(1-\hat{\alpha}\left(1+\sum_{\substack{i=1\\i\neq\bar{r}}}^{k}\frac{\alpha_i}{1-\alpha_i}\right)\right)+\sum_{\substack{i=1\\i\neq\bar{r}}}^{k}\frac{\alpha_i}{1-\alpha_i}\right) \\
& < \left(1+\sum_{\substack{i=1\\i\neq\bar{r}}}^{k}\frac{\alpha_i}{1-\alpha_i}\right)(1-\alpha_{k+1}\hat{\alpha}) \qquad (51\text{c})\\
\Leftrightarrow\quad & \alpha_{\bar{r}}\left(\alpha_{k+1}\left(1-\sum_{\substack{i=1\\i\neq\bar{r}}}^{k}\frac{\alpha_i}{1-\alpha_i}\right)+\sum_{\substack{i=1\\i\neq\bar{r}}}^{k}\frac{\alpha_i}{1-\alpha_i}\right) \\
& < \left(1+\sum_{\substack{i=1\\i\neq\bar{r}}}^{k}\frac{\alpha_i}{1-\alpha_i}\right)(1-\alpha_{k+1}\hat{\alpha}) \qquad (51\text{d})\\
\Leftrightarrow\quad & \alpha_{\bar{r}}\frac{\alpha_{k+1}(1-\sum_{\substack{i=1\\i\neq\bar{r}}}^{k}\frac{\alpha_i}{1-\alpha_i})+\sum_{\substack{i=1\\i\neq\bar{r}}}^{k}\frac{\alpha_i}{1-\alpha_i}}{(1+\sum_{\substack{i=1\\i\neq\bar{r}}}^{k}\frac{\alpha_i}{1-\alpha_i})(1-\alpha_{k+1}\hat{\alpha})} < 1. \qquad (51\text{e})
\end{aligned}$$

Now, observe that

$$\begin{aligned}
\alpha_{k+1}\left(1-\sum_{\substack{i=1\\i\neq\bar{r}}}^{k}\frac{\alpha_i}{1-\alpha_i}\right)+\sum_{\substack{i=1\\i\neq\bar{r}}}^{k}\frac{\alpha_i}{1-\alpha_i} &= \left(\sum_{\substack{i=1\\i\neq\bar{r}}}^{k}\frac{\alpha_i}{1-\alpha_i}+\frac{\alpha_{k+1}}{1-\alpha_{k+1}}\right)(1-\alpha_{k+1}) \\
&= \sum_{\substack{i=1\\i\neq\bar{r}}}^{k+1}\frac{\alpha_i}{1-\alpha_i}(1-\alpha_{k+1}) \qquad (52)
\end{aligned}$$

and

$$\left(1+\sum_{\substack{i=1\\ i\neq\overline{r}}}^{k}\frac{\alpha_i}{1-\alpha_i}\right)(1-\alpha_{k+1}\hat{\alpha}) = 1+\sum_{\substack{i=1\\ i\neq\overline{r}}}^{k}\frac{\alpha_i}{1-\alpha_i} - \alpha_{k+1}\sum_{\substack{i=1\\ i\neq\overline{r}}}^{k}\frac{\alpha_i}{1-\alpha_i} \tag{53a}$$

$$= \left(1+\sum_{\substack{i=1\\ i\neq\overline{r}}}^{k}\frac{\alpha_i}{1-\alpha_i}+\frac{\alpha_{k+1}}{1-\alpha_{k+1}}\right)(1-\alpha_{k+1}) \tag{53b}$$

$$= \left(1+\sum_{\substack{i=1\\ i\neq\overline{r}}}^{k+1}\frac{\alpha_i}{1-\alpha_i}\right)(1-\alpha_{k+1}). \tag{53c}$$

In view of (52) and (53a)–(53c), (51a)–(51e) becomes

$$\alpha_{k+1}\frac{\sum_{i=1}^{k}\frac{\alpha_i}{1-\alpha_i}}{1+\sum_{i=1}^{k}\frac{\alpha_i}{1-\alpha_i}} < 1 \quad\Leftrightarrow\quad \alpha_{\overline{r}}\frac{\sum_{i=1}^{k+1}\frac{\alpha_1}{1-\alpha_i}}{1+\sum_{i=1}^{k+1}\frac{\alpha_1}{1-\alpha_i}} = \alpha_{\overline{r}}\overline{\beta} < 1. \tag{54}$$

This proves (48). Now proceed similar to Case I in view of (48) and (49). □

The assumption $\alpha_r\overline{\alpha} < 1$ is critical in the conclusion of the above theorem as we illustrate in the following example.

Example 4.8 Let $\epsilon > 0$, let $\delta > 1$, let $\alpha_1 \in\,]0, \frac{1}{2}(\sqrt{(\epsilon+\delta)^2+4}-(\epsilon+\delta))[$, let $\alpha_2 = \alpha_1+\delta+\epsilon$, and let

$$S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \tag{55}$$

Set $R_1 = (1-\alpha_1)\operatorname{Id} - \alpha_1 S$, $R_2 = (1-\alpha_2)\operatorname{Id} + \alpha_2 S$, $R_3 = -\frac{1}{\delta}S$, and

$$R = R_3R_2R_1. \tag{56}$$

Then $R = R_3R_1R_2 = R_1R_2R_3 = R_1R_3R_2 = R_2R_3R_1 = R_2R_1R_3$. Moreover, the following hold:

(i) $\alpha_1 \in\,]0,1[$, $\alpha_2 > 1$, and $\alpha_1\alpha_2 < 1$.

(ii) R_3 is α_3-conically nonexpansive where $\alpha_3 = \frac{1+\delta}{2\delta} \in\,]1/2, 1]$.

(iii) $\frac{\alpha_1+\alpha_2-2\alpha_1\alpha_2}{1-\alpha_1\alpha_2}\alpha_3 > 1$.

(iv) $R = (\frac{\epsilon+\delta}{\delta})\operatorname{Id} + (\frac{\alpha_1+\alpha_2-2\alpha_1\alpha_2-1}{\delta})S$.

(v) $\operatorname{Id} - R = -\frac{\epsilon}{\delta}\operatorname{Id} - (\frac{\alpha_1+\alpha_2-2\alpha\alpha_2-1}{\delta})S$. Hence, $\operatorname{Id} - R$ is not monotone.

(vi) R is *not* conically nonexpansive.

Proof It is straightforward to verify that $R = R_3R_1R_2 = R_1R_2R_3 = R_1R_3R_2 = R_2R_3R_1 = R_2R_1R_3$. (i): It is clear that $\alpha_1 \in\,]0,1[$ and that $\alpha_2 > 1$. Note that $\alpha_1\alpha_2 < 1 \Leftrightarrow \alpha_1^2 + (\epsilon+\delta)\alpha_1 - 1 < 0 \Leftrightarrow \alpha_1$ lies between the roots of the quadratic $x^2 + (\epsilon+\delta)x - 1$, and the conclusion follows from the quadratic formula. (ii): This follows from [2, Proposition 4.38]. (iii): Indeed, in view of (i) we have

$$
\begin{aligned}
&\frac{\alpha_1+\alpha_2-2\alpha_1\alpha_2}{1-\alpha_1\alpha_2}\alpha_3>1\\
&\quad\Leftrightarrow\quad (\alpha_1+\alpha_2-2\alpha_1\alpha_2)\alpha_3>1-\alpha_1\alpha_2 && \text{(57a)}\\
&\quad\Leftrightarrow\quad (\alpha_1+\alpha_2-2\alpha_1\alpha_2)(1+\delta)>2(1-\alpha_1\alpha_2)\delta && \text{(57b)}\\
&\quad\Leftrightarrow\quad (\alpha_1+\alpha_2)(1+\delta)-2\alpha_1\alpha_2-2\alpha_1\alpha_2\delta>2\delta-2\alpha_1\alpha_2\delta && \text{(57c)}\\
&\quad\Leftrightarrow\quad (\alpha_1+\alpha_2)(1+\delta)-2\alpha_1\alpha_2>2\delta && \text{(57d)}\\
&\quad\Leftrightarrow\quad (2\alpha_1+\epsilon+\delta)(1+\delta)-2\alpha_1(\alpha_1+\epsilon+\delta)>2\delta && \text{(57e)}\\
&\quad\Leftrightarrow\quad 2\alpha_1(1+\delta-\alpha_1-\epsilon-\delta)+\delta^2+\delta(1+\epsilon)+\epsilon>2\delta && \text{(57f)}\\
&\quad\Leftrightarrow\quad 2\alpha_1(\alpha_1-1+\epsilon)<\delta^2-\delta+\epsilon\delta+\epsilon=\delta^2-\delta+(1+\delta)\epsilon. && \text{(57g)}
\end{aligned}
$$

Now, because $\alpha_1<1$, $\delta\geq 1$, we learn that $2\alpha_1(\alpha_1-1+\epsilon)<2\alpha_1\epsilon<(1+\delta)\epsilon<(1+\delta)\epsilon+\delta^2-\delta$, and the conclusion follows. (iv): It is straightforward, by noting that $S^2=-\mathrm{Id}$, to verify that $R_2R_1=R_1R_2=(1-\alpha_1-\alpha_2+\alpha_1\alpha_2)\,\mathrm{Id}+(\alpha_2(1-\alpha_1)-\alpha_1(1-\alpha_2))S-\alpha_1\alpha_2S^2=(1-\alpha_1-\alpha_2+2\alpha_1\alpha_2)\,\mathrm{Id}+(\alpha_2-\alpha_1)S$. Consequently, $R_3R_2R_1=\frac{1}{\delta}(-(1-\alpha_1-\alpha_2+2\alpha_1\alpha_2)S-(\alpha_2-\alpha_1)S^2)=\frac{1}{\delta}((\alpha_2-\alpha_1)\,\mathrm{Id}-(1-\alpha_1-\alpha_2+2\alpha_1\alpha_2)S)=\frac{\epsilon+\delta}{\delta}\,\mathrm{Id}+\frac{\alpha_1+\alpha_2-2\alpha_1\alpha_2-1}{\delta}S$. (v): This is a direct consequence of (iv). (vi): Combine (v) and Lemma 2.9. □

Theorem 4.9 (Composition of cocoercive operators) *Let $m\geq 1$ be an integer, set $I=\{1,\dots,m\}$, let $(R_i)_{i\in I}$ be a family of operators from X to X, let β_i be real numbers in $]0,+\infty[$, and suppose that, for every $i\in I$, R_i is $\frac{1}{\beta_i}$-cocoercive. Then there exists a nonexpansive operator $N\colon X\to X$ such that*

$$
R_m\cdots R_1=\beta_m\cdots\beta_1\left(\frac{1}{1+m}\,\mathrm{Id}+\frac{m}{1+m}N\right). \tag{58}
$$

Proof Apply Theorem 4.7 with (α_i,δ_i) replaced by $(1/2,\beta_i)$, $i\in\{1,\dots,m\}$. □

5 Application to the Douglas–Rachford algorithm

Theorem 5.1 (Averagedness of the Douglas–Rachford operator) *Let $\mu>\omega\geq 0$, and let $\gamma\in\,]0,(\mu-\omega)/(2\mu\omega)[$. Suppose that one of the following holds:*

(i) *A is maximally $(-\omega)$-monotone and B is maximally μ-monotone.*
(ii) *A is maximally μ-monotone and B is maximally $(-\omega)$-monotone.*

Set

$$
T=\frac{1}{2}(\mathrm{Id}+R_{\gamma B}R_{\gamma A}),\quad \textit{and}\quad \alpha=\frac{\mu-\omega}{2(\mu-\omega-\gamma\mu\omega)}. \tag{59}
$$

Then $\alpha\in\,]0,1[$ and T is α-averaged.

Proof Suppose that (i) holds. Note that γA is $-\gamma\omega$-monotone, and

$$
-\gamma\omega>-\frac{\mu-\omega}{2\mu}\geq-\frac{\mu}{2\mu}>-1. \tag{60}
$$

Using (13) and Fact 2.1 we learn that $J_{\gamma A}$ and, in turn, T are single-valued and $\operatorname{dom} J_{\gamma A} = \operatorname{dom} T = X$. It follows from [3, Proposition 4.3 and Table 1] that $-R_{\gamma A}$ is $\frac{1}{1+\gamma\mu}$-conically nonexpansive and $-R_{\gamma B}$ is $\frac{1}{1-\gamma\omega}$-conically nonexpansive. It follows from Theorem 4.2, applied with $(\alpha_1,\beta_1,\delta_1,\alpha_2,\beta_2,\delta_2)$ replaced by $(1-\frac{1}{1+\gamma\mu},\frac{1}{1+\gamma\mu},-1,1-\frac{1}{1-\gamma\omega},\frac{1}{1-\gamma\omega},-1)$, that $R_{\gamma B}R_{\gamma A}$ is $\frac{\mu-\omega}{\mu-\omega-\gamma\mu\omega}$-conically nonexpansive. Therefore, there exists a nonexpansive mapping $N\colon X\to X$ such that

$$R_{\gamma B}R_{\gamma A} = (1-\delta)\operatorname{Id} + \delta N, \qquad \delta = \frac{\mu-\omega}{\mu-\omega-\gamma\mu\omega}. \tag{61}$$

The conclusion now follows by applying Proposition 2.8 with (β,N) replaced by $(\frac{\alpha}{\delta}, R_{\gamma B}R_{\gamma A})$. Finally, notice that $\gamma < \frac{\mu-\omega}{2\mu\omega}$, which implies that $0 < \mu-\omega < 2(\mu-\omega-\gamma\mu\omega)$. Therefore,

$$\alpha = \frac{\mu-\omega}{2(\mu-\omega-\gamma\mu\omega)} \in \left]0,1\right[. \tag{62}$$

The proof of (ii) follows similarly. □

Corollary 5.2 ([9, Theorem 4.5(ii)]) *Let $\mu > \omega \geq 0$, and let $\gamma \in \left]0,(\mu-\omega)/(2\mu\omega)\right[$. Suppose that one of the following holds:*

(i) *A is maximally $(-\omega)$-monotone and B is maximally μ-monotone.*

(ii) *A is maximally μ-monotone and B is maximally $(-\omega)$-monotone.*

Set $T = \frac{1}{2}(\operatorname{Id} + R_{\gamma B}R_{\gamma A})$ and let $x_0 \in X$. Then $(\exists \overline{x} \in \operatorname{Fix} T = \operatorname{Fix} R_{\gamma B}R_{\gamma A})$ such that $T^n x_0 \rightharpoonup \overline{x}$.

Proof Combine Theorem 5.1 and [2, Theorem 5.15]. □

Remark 5.3 In view of (13), one might think that the scaling factor γ is required *only* to guarantee the single-valuedness and the full domain of T. However, it is actually critical to guarantee convergence as well, as we illustrate in Example 5.4.

Example 5.4 Let $\mu > \omega \geq 0$, let U be a closed linear subspace of X, suppose that[9]

$$A = N_U + \mu \operatorname{Id}, \qquad B = -\omega \operatorname{Id}. \tag{63}$$

Then A is μ-monotone, B is $-\omega$-monotone, and $(\forall \gamma \in [1/(2\omega), 1/\omega[)$ $J_{\gamma B}$ is single-valued. Furthermore, we have

$$T = \frac{1}{2}(\operatorname{Id} + R_{\gamma B}R_{\gamma A}) = \frac{1+\gamma\omega}{(1-\gamma\omega)(1+\gamma\mu)} P_U - \frac{\gamma\omega}{1-\gamma\omega}\operatorname{Id}, \tag{64}$$

and $(\forall x_0 \in U^{\perp})$ $(T^n x_0)_{n\in\mathbb{N}}$ does not converge.

Proof Indeed, one can verify that

$$J_{\gamma A} = \frac{1}{1-\gamma\omega}\operatorname{Id}, \qquad J_{\gamma B} = \frac{1}{1+\gamma\mu} P_U. \tag{65}$$

[9]Let C be a nonempty, closed convex subset of X. Here and elsewhere, we shall use N_C to denote the *normal cone operator* associated with C, defined by $N_C(x) = \{u \in X \mid \sup\langle C - x \mid u\rangle \leq 0\}$ if $x \in C$; and $N_C(x) = \varnothing$, otherwise.

Consequently,

$$R_{\gamma A} = \frac{1+\gamma\omega}{1-\gamma\omega}\operatorname{Id}, \qquad R_{\gamma B} = \frac{2}{1+\gamma\mu}P_U - \operatorname{Id}, \tag{66}$$

and (64) follows. Therefore,

$$T_{|U^\perp} = -\frac{\gamma\omega}{1-\gamma\omega}\operatorname{Id} \quad \text{and} \quad -\frac{\gamma\omega}{1-\gamma\omega} \in \left]-\infty,-1\right]. \tag{67}$$

Hence, $(\forall x_0 \in U^\perp)$ $(T^n x_0)_{n\in\mathbb{N}}$ does not converge. □

Before we proceed to the convergence analysis, we recall that if T is averaged and $\operatorname{Fix} T \neq \varnothing$ then $(\forall x \in X)$ we have (see, e.g., [22, Theorem 3.7])

$$T^n x - T^{n+1}x \to 0. \tag{68}$$

We conclude this section by proving the strong convergence of the shadow sequence of the Douglas–Rachford algorithm.

Theorem 5.5 (Convergence analysis of the Douglas–Rachford algorithm) *Let $\mu > \omega \geq 0$, and let $\gamma \in \left]0, (\mu-\omega)/(2\mu\omega)\right[$. Suppose that one of the following holds*:

(i) *A is maximally μ-monotone and B is maximally $(-\omega)$-monotone.*
(ii) *A is maximally $(-\omega)$-monotone and B is maximally μ-monotone.*

Set

$$T = \frac{1}{2}(\operatorname{Id} + R_{\gamma B}R_{\gamma A}), \tag{69}$$

and let $x_0 \in X$. Then $\operatorname{zer}(A+B) \neq \varnothing$. Moreover, there exist $\overline{x} \in \operatorname{Fix} T = \operatorname{Fix} R_{\gamma B}R_{\gamma A}$, $\operatorname{zer}(A+B) = \{J_{\gamma A}\overline{x}\} = \{J_{\gamma B}R_{\gamma A}\overline{x}\}$, $T^n x_0 \rightharpoonup \overline{x}$, $J_{\gamma A}T^n x_0 \to J_{\gamma A}\overline{x}$, and $J_{\gamma B}R_{\gamma A}T^n x_0 \to J_{\gamma B}R_{\gamma A}\overline{x}$.

Proof Suppose that (i) holds. Since $A+B$ is $(\mu-\omega)$-monotone and $\mu-\omega > 0$, we conclude from [2, Proposition 23.35] that $\operatorname{zer}(A+B)$ is a singleton. Combining with Fact 2.1 with (A,B) replaced by $(\gamma A, \gamma B)$ yields $\operatorname{zer}(A+B) = \operatorname{zer}(\gamma A + \gamma B) = \{J_{\gamma A}\overline{x}\} = \{J_{\gamma B}R_{\gamma A}\overline{x}\}$. The claim that $T^n x_0 \rightharpoonup \overline{x}$ follows from Corollary 5.2. It remains to show that $J_{\gamma A}T^n x_0 \to J_{\gamma A}\overline{x}$ and $J_{\gamma B}R_{\gamma A}T^n x_0 \to J_{\gamma B}R_{\gamma A}\overline{x}$. To this end, note that $(T^n x_0)_{n\in\mathbb{N}}$ is bounded; consequently, since $J_{\gamma A}$ and $J_{\gamma B}R_{\gamma A}$ are Lipschitz continuous (see Proposition 2.16(i)&(ii)), we learn that

$$\big(J_{\gamma A}T^n x_0\big)_{n\in\mathbb{N}} \text{ and } \big(J_{\gamma B}R_{\gamma A}T^n x_0\big)_{n\in\mathbb{N}} \text{ are bounded.} \tag{70}$$

On the one hand, in view of (68) we have

$$(\operatorname{Id} - T)T^n x_0 = T^n x_0 - T^{n+1}x_0 = J_{\gamma A}T^n x_0 - J_{\gamma B}R_{\gamma A}T^n x_0 \to 0. \tag{71}$$

Combining (70) and (71) yields

$$\big\|J_{\gamma A}T^n x_0 - J_{\gamma A}\overline{x}\big\|^2 - \big\|J_{\gamma B}R_{\gamma A}T^n x_0 - J_{\gamma B}R_{\gamma A}\overline{x}\big\|^2 \tag{72a}$$

$$= \big\langle J_{\gamma A}T^n x_0 - J_{\gamma B}R_{\gamma A}T^n x_0 \mid J_{\gamma A}T^n x_0 + J_{\gamma B}R_{\gamma A}T^n x_0 - J_{\gamma A}\overline{x} - J_{\gamma B}R_{\gamma A}\overline{x}\big\rangle \tag{72b}$$

$$= \big\langle T^n x_0 - T^{n+1}x_0 \mid J_{\gamma A}T^n x_0 + J_{\gamma B}R_{\gamma A}T^n x_0 - J_{\gamma A}\overline{x} - J_{\gamma B}R_{\gamma A}\overline{x}\big\rangle \to 0. \tag{72c}$$

On the other hand, combining Lemma 2.3, applied with $(R_1, R_2, R(\lambda), \lambda)$ replaced by $(R_{\gamma A}, R_{\gamma B}, T, 1/2)$ and (x, y) replaced by $(T^n x_0, \overline{x})$, in view of (68) yields

$$0 \leftarrow \langle T^{n+1}x_0 - \overline{x} \mid T^n x_0 - T^{n+1}x_0 \rangle \tag{73a}$$

$$\geq \gamma\mu \left(\|J_{\gamma A} T^n x_0 - J_{\gamma A}\overline{x}\|^2 - \frac{\omega}{\mu} \|J_{\gamma B} R_{\gamma A} T^n x_0 - J_{\gamma B} R_{\gamma A}\overline{x}\|^2 \right) \tag{73b}$$

$$\geq -\frac{\gamma\mu\omega}{\mu - \omega} \|T^n x_0 - T^{n+1}x_0\|^2 \to 0. \tag{73c}$$

Therefore,

$$\|J_{\gamma A} T^n x_0 - J_{\gamma A}\overline{x}\|^2 - \frac{\omega}{\mu} \|J_{\gamma B} R_{\gamma A} T^n x_0 - J_{\gamma B} R_{\gamma A}\overline{x}\|^2 \to 0. \tag{74}$$

Combining (72a)–(72c) and (74) and noting that $\frac{\omega}{\mu} < 1$ yields $\|J_{\gamma A} T^n x_0 - J_{\gamma A}\overline{x}\|^2 \to 0$ and $\|J_{\gamma B} R_{\gamma A} T^n x_0 - J_{\gamma B} R_{\gamma A}\overline{x}\|^2 \to 0$, which proves (i). The proof of (ii) proceeds similarly. □

Remark 5.6 (Relaxed Douglas–Rachford algorithm) A careful look at the proofs of Theorem 5.1 and Theorem 5.5 reveals that analogous conclusions can be drawn for the relaxed Douglas–Rachford operator defined by $T_\lambda = (1-\lambda)\operatorname{Id} + \lambda R_{\gamma B} R_{\gamma A}$, $\lambda \in]0, 1[$. In this case, we choose $\gamma \in]0, ((1-\lambda)(\mu - \omega))/(\mu\omega)[$. One can verify that the corresponding averagedness constant is $\alpha = \frac{\lambda(\mu-\omega)}{\mu-\omega-\gamma\mu\omega} \in]0, 1[$.

6 Application to the forward–backward algorithm

Throughout this section we assume that

$$\boxed{A\colon X \to X, \qquad B\colon X \rightrightarrows X, \qquad \mu \geq 0, \qquad \omega \geq 0, \quad \text{and} \quad \beta > 0.}$$

In the rest of this section, we prove that the forward–backward operator is averaged, hence its iterates form a weakly convergent sequence in each of the following situations:

- A is maximally μ-monotone, $A - \mu \operatorname{Id}$ is $\frac{1}{\beta}$-cocoercive, B is maximally $(-\omega)$-monotone, and $\mu \geq \omega$.
- A is maximally $(-\omega)$-monotone, $A + \omega \operatorname{Id}$ is $\frac{1}{\beta}$-cocoercive, B is maximally μ-monotone, and $\mu \geq \omega$.
- A is β-Lipschitz continuous, B is maximally μ-monotone, and $\mu \geq \beta$.

That is, we do not require A and B to be monotone. Instead, it is enough that the sum $A + B$ is monotone to have an averaged forward–backward map. In addition, we show that the forward–backward map is contractive if the sum $A + B$ is strongly monotone, and we prove the tightness of our contraction factor.

Theorem 6.1 (Case I: A is μ-monotone) *Let $\mu \geq \omega \geq 0$, and let $\beta > 0$. Suppose that A is maximally μ-monotone, $A - \mu \operatorname{Id}$ is $\frac{1}{\beta}$-cocoercive, and B is maximally $(-\omega)$-monotone. Let $\gamma \in]0, 2/(\beta + 2\mu)[$. Set $T = J_{\gamma B}(\operatorname{Id} - \gamma A)$, set $\nu = \gamma\beta/(2(1-\gamma\mu))$, set $\delta = (1-\gamma\mu)/(1-\gamma\omega)$, and let $x_0 \in X$. Then $\delta \in]0, 1]$ and $\nu \in]0, 1[$. Moreover, the following hold:*

(i) *$T = \delta((1-\nu)\operatorname{Id} + \nu N)$, N is nonexpansive.*

(ii) *T is $(1 - (\delta(1-\nu))/(2-\nu))$-averaged.*

(iii) *T is δ-Lipschitz continuous.*

(iv) *There exists* $\overline{x} \in \operatorname{Fix} T = \operatorname{zer}(A+B)$ *such that* $T^n x_0 \rightharpoonup \overline{x}$.

Suppose that $\mu > \omega$. *Then we additionally have*:

(v) T *is a Banach contraction with a constant* $\delta < 1$.

(vi) $\operatorname{zer}(A+B) = \{\overline{x}\}$ *and* $T^n x_0 \to \overline{x}$ *with a linear rate* $\delta < 1$.

Proof Clearly, $\delta \in]0,1]$ and $\nu > 0$. Moreover, we have $\nu < 1 \Leftrightarrow \gamma\beta < 2(1-\gamma\mu) \Leftrightarrow \gamma < 2/(\beta + 2\mu\beta)$. Hence, $\nu \in]0,1[$ as claimed. Next note that $\mu < (\beta + 2\mu)/2$, hence $\gamma\omega < \gamma\mu < (2\gamma)/(\beta+2\mu) < 1$. It follows from Proposition 2.2 that $J_{\gamma B}$ and, in turn, T are single-valued and $\operatorname{dom} J_{\gamma B} = \operatorname{dom} T = X$. The assumption on A implies that there exists $\overline{N}\colon X \to X$, $\overline{N}$ is nonexpansive, such that $A - \mu \operatorname{Id} = \frac{\beta}{2}\operatorname{Id} + \frac{\beta}{2}\overline{N}$. Therefore,

$$\operatorname{Id} - \gamma A = \operatorname{Id} - \gamma(A - \mu \operatorname{Id}) - \gamma\mu \operatorname{Id} = (1-\gamma\mu)\operatorname{Id} - \frac{\gamma\beta}{2}(\operatorname{Id} + \overline{N}) \tag{75a}$$

$$= (1-\gamma\mu)\big((1-\nu)\operatorname{Id} + \nu(-\overline{N})\big). \tag{75b}$$

Moreover, Proposition 2.16(i) implies that

$$J_{\gamma B} \text{ is } (1-\gamma\omega)\text{-cocoercive.} \tag{76}$$

(i): It follows from Corollary 4.3 applied with (R_1, R_2) replaced by $(\operatorname{Id} - \gamma A, J_{\gamma B})$ and (α, β, δ) replaced by $(\nu, 1/(1-\gamma\omega), 1-\gamma\mu)$, in view of (75a)–(75b) and (76), that there exists a nonexpansive operator N such that $T = J_{\gamma B}(\operatorname{Id} - \gamma A) = \delta((1-\nu)\operatorname{Id} + \nu N)$. (ii): Combine (i) and Lemma 2.15(i). (iii): Combine (i) and (ii). (iv): Applying Proposition 2.2 with (A,B) replaced by $(\gamma A, \gamma B)$ yields $\operatorname{zer}(A+B) = \operatorname{zer}(\gamma A + \gamma B) = \operatorname{Fix} T$. The claim that $T^n x_0 \rightharpoonup \overline{x}$ follows from combining (ii) and [2, Theorem 5.15]. (v): Observe that $\delta < 1 \Leftrightarrow \mu > \omega$. Now, combine with (iii). (vi): Note that $A+B$ is maximally $(\mu - \omega)$-monotone and $\mu - \omega > 0$, we conclude from [2, Proposition 23.35] that $\operatorname{zer}(A+B)$ is a singleton. Alternatively, use (iii) to learn that T is a Banach contraction with a constant $\delta < 1$, hence $\operatorname{zer}(A+B) = \operatorname{Fix} T$ is a singleton, and the conclusion follows. □

Theorem 6.2 *Let* $\mu > \omega \geq 0$, *and let* $\beta > 0$. *Suppose that* A *is maximally* μ*-monotone,* $A - \mu \operatorname{Id}$ *is* $\frac{1}{\beta}$*-cocoercive, and* B *is maximally* $(-\omega)$*-monotone. Let* $\gamma \in [2/(\beta+2\mu), 2/(\beta+\mu)[$. *Set* $T = J_{\gamma B}(\operatorname{Id} - \gamma A)$, *set* $\nu = \gamma\beta/(2(\gamma(\mu+\beta)-1))$, *set* $\delta = (1-\gamma(\mu+\beta))/(1-\gamma\omega)$, *and let* $x_0 \in X$. *Then* $\delta \in]-1,0]$ *and* $\nu \in]0,1[$. *Moreover, the following hold*:

(i) $T = \delta((1-\nu)\operatorname{Id} + \nu N)$, N *is nonexpansive.*

(ii) T *is a Banach contraction with a constant* $|\delta| < 1$.

(iii) *There exists* $\overline{x} \in X$ *such that* $\operatorname{Fix} T = \operatorname{zer}(A+B) = \{\overline{x}\}$ *and* $T^n x_0 \to \overline{x}$ *with a linear rate* $|\delta| < 1$.

Proof We proceed similar to the proof of Theorem 6.1 to verify that T is single-valued, $\operatorname{dom} T = X$, $\nu \in]0,1[$, and $\delta \in]-1,0]$. The assumption on A implies that there exists $\overline{N}\colon X \to X$, $\overline{N}$ is nonexpansive such that $A - \mu \operatorname{Id} = \frac{\beta}{2}\operatorname{Id} + \frac{\beta}{2}\overline{N}$. Therefore,

$$\operatorname{Id} - \gamma A = \operatorname{Id} - \gamma(A - \mu \operatorname{Id}) - \gamma\mu \operatorname{Id} = (1-\gamma\mu)\operatorname{Id} - \frac{\gamma\beta}{2}(\operatorname{Id} + \overline{N}) \tag{77a}$$

$$= \big(1-\gamma(\mu+\beta)\big)\big((1-\nu)\operatorname{Id} + \nu(\overline{N})\big). \tag{77b}$$

Now, proceed similar to the proof of Theorem 6.1(i), (v), and (vi) in view of (76). □

Corollary 6.3 *Let $\mu > \omega \geq 0$, and let $\beta > 0$. Suppose that A is maximally μ-monotone, $A - \mu \operatorname{Id}$ is $\frac{1}{\beta}$-cocoercive, and B is maximally $(-\omega)$-monotone. Let $\gamma \in]0, 2/(\beta + \mu)[$. Set $T = J_{\gamma B}(\operatorname{Id} - \gamma A)$, set $\delta = \max(1 - \gamma\mu, \gamma(\mu + \beta) - 1)/(1 - \gamma\omega)$, and let $x_0 \in X$. Then $\delta \in [0, 1[$, T is a Banach contraction with a constant δ, and there exists $\overline{x} \in X$ such that $\operatorname{Fix} T = \operatorname{zer}(A + B) = \{\overline{x}\}$ and $T^n x_0 \to \overline{x}$.*

Proof Combine Theorem 6.1 and Theorem 6.2. □

Remark 6.4 (Tightness of the Lipschitz constant)

(i) Suppose that the setting of Theorem 6.1 holds. Set $(A, B) = (\mu \operatorname{Id}, -\omega \operatorname{Id})$. Then $T = \frac{1-\gamma\mu}{1-\gamma\omega} \operatorname{Id}$. Hence, the claimed Lipschitz constant is tight.

(ii) Suppose that the setting of Theorem 6.2 holds. Set $(A, B) = ((\mu + \beta) \operatorname{Id}, -\omega \operatorname{Id})$. Then $T = \frac{\gamma(\mu+\beta)-1}{1-\gamma\omega} \operatorname{Id}$. Hence, the claimed contraction factor is tight.

Note in particular that the worst cases are subgradients of convex functions. Hence, the worst cases are attained by the proximal gradient method.

Theorem 6.5 (Case II: $A + \omega \operatorname{Id}$ is cocoercive) *Let $\mu \geq \omega \geq 0$, let $\beta > 0$, and let $\overline{\beta} \in]\max\{\beta, \mu + \omega\}, +\infty[$. Suppose that A is maximally $(-\omega)$-monotone, $A + \omega \operatorname{Id}$ is β-cocoercive, and B is maximally μ-monotone. Let $\gamma \in]0, 2/(\overline{\beta} - 2\omega)[$. Set $T = J_{\gamma B}(\operatorname{Id} - \gamma A)$, set $\nu = \gamma\overline{\beta}/(2(1 + \gamma\omega))$, set $\delta = (1 + \gamma\omega)/(1 + \gamma\mu)$, and let $x_0 \in X$. Then $\delta \in]0, 1]$ and $\nu \in]0, 1[$. Moreover, the following hold:*

(i) *$T = \delta((1 - \nu) \operatorname{Id} + \nu N)$, N is nonexpansive.*

(ii) *T is $(1 - (\delta(1 - \nu))/(2 - \nu))$-averaged.*

(iii) *T is δ-Lipschitz continuous.*

(iv) *There exists $\overline{x} \in \operatorname{Fix} T = \operatorname{zer}(A + B)$, and $T^n x_0 \rightharpoonup \overline{x}$.*

Suppose that $\mu > \omega$. Then we additionally have:

(v) *T is a Banach contraction with a constant $\delta < 1$.*

(vi) *$\operatorname{zer}(A + B) = \{\overline{x}\}$ and $T^n x_0 \to \overline{x}$ with a linear rate $\delta < 1$.*

Proof Observe that the assumption on A and Lemma 2.11 applied with T replaced by $A + \omega \operatorname{Id}$ imply that there exists $\overline{N}: X \to X$, $\overline{N}$ is nonexpansive, such that $A + \omega \operatorname{Id} = \frac{\overline{\beta}}{2} \operatorname{Id} + \frac{\overline{\beta}}{2} \overline{N}$.

$$\operatorname{Id} - \gamma A = \operatorname{Id} - \gamma(A + \omega \operatorname{Id}) + \gamma\omega \operatorname{Id} = (1 + \gamma\omega) \operatorname{Id} - \frac{\gamma\overline{\beta}}{2}(\operatorname{Id} + \overline{N}) \tag{78a}$$

$$= (1 + \gamma\omega)\big((1 - \nu) \operatorname{Id} + \nu(-\overline{N})\big). \tag{78b}$$

Moreover, Proposition 2.16(i) implies that

$$J_{\gamma B} \text{ is } (1 + \gamma\mu)\text{-cocoercive.} \tag{79}$$

Now proceed similar to the proof of Theorem 6.1 but use (78a)–(78b) and (79). □

Theorem 6.6 *Let $\mu > \omega \geq 0$, let $\beta > 0$, and let $\overline{\beta} \in]\max\{\beta, \mu + \omega\}, +\infty[$. Suppose that A is maximally $(-\omega)$-monotone, $A + \omega \operatorname{Id}$ is β-cocoercive, and B is maximally μ-monotone. Let $\gamma \in [2/(\overline{\beta} - 2\omega), 2/(\overline{\beta} - \mu - \omega)[$. Set $T = J_{\gamma B}(\operatorname{Id} - \gamma A)$, set $\nu = \gamma\overline{\beta}/(2(\gamma\overline{\beta} - \gamma\omega - 1))$, set $\delta = (1 + \gamma\omega - \gamma\overline{\beta})/(1 + \gamma\mu)$, and let $x_0 \in X$. Then $\delta \in]-1, 0]$ and $\nu \in]0, 1[$. Moreover, the following hold:*

(i) $T = \delta((1-\nu)\operatorname{Id} + \nu N)$, N *is nonexpansive.*
(ii) T *is a Banach contraction with a constant* $|\delta| < 1$.
(iii) *There exists* $\overline{x} \in X$ *such that* $\operatorname{Fix} T = \operatorname{zer}(A+B) = \{\overline{x}\}$ *and* $T^n x_0 \to \overline{x}$ *with a linear rate* $|\delta| < 1$.

Proof Observe that the assumption on A and Lemma 2.11 applied with T replaced by $A + \omega \operatorname{Id}$ implies that there exists $\overline{N} \colon X \to X$, $\overline{N}$ is nonexpansive, such that $A + \omega \operatorname{Id} = \frac{\overline{\beta}}{2}\operatorname{Id} + \frac{\overline{\beta}}{2}\overline{N}$.

$$\operatorname{Id} - \gamma A = \operatorname{Id} - \gamma(A + \omega \operatorname{Id}) + \gamma\omega \operatorname{Id} = (1+\gamma\omega)\operatorname{Id} - \frac{\gamma\overline{\beta}}{2}(\operatorname{Id} + \overline{N}) \tag{80a}$$

$$= (1 + \gamma\omega - \gamma\overline{\beta})\big((1-\nu)\operatorname{Id} + \nu\overline{N}\big). \tag{80b}$$

Now proceed similar to the proof of Theorem 6.5 in view of (79). □

Corollary 6.7 *Let* $\mu > \omega \geq 0$, *let* $\beta > 0$, *and let* $\overline{\beta} \in \,]\max\{\beta, \mu + \omega\}, +\infty[$. *Suppose that* A *is maximally* $(-\omega)$*-monotone,* $A + \omega \operatorname{Id}$ *is* β*-cocoercive, and* B *is maximally* μ*-monotone. Let* $\gamma \in [0, 2/(\overline{\beta} - \mu - \omega)[$. *Set* $T = J_{\gamma B}(\operatorname{Id} - \gamma A)$, *set* $\delta = \max\{1 + \gamma\mu, \gamma\overline{\beta} - \gamma\omega - 1\}/(1+\gamma\mu)$, *and let* $x_0 \in X$. *Then* $\delta \in \,]-1, 0]$ *and* $\nu \in \,]0,1[$. *Then* $\delta \in [0,1[$, T *is a Banach contraction with a constant* δ, *and there exists* $\overline{x} \in X$ *such that* $\operatorname{Fix} T = \operatorname{zer}(A+B) = \{\overline{x}\}$ *and* $T^n x_0 \to \overline{x}$.

Proof Combine Theorem 6.5 and Theorem 6.6. □

Theorem 6.8 (Case III: A is β-Lipschitz continuous) *Let* $\mu \geq \beta > 0$. *Suppose that* A *is* β*-Lipschitz continuous and that* B *is maximally* μ*-monotone. Let* $\overline{\beta} \in \,]2\beta, +\infty[$, *and let* $\gamma \in \,]0, 2/(\overline{\beta} - 2\beta)\}[$. *Set* $T = J_{\gamma B}(\operatorname{Id} - \gamma A)$, *set* $\nu = \gamma\overline{\beta}/(2(1+\gamma\beta))$, *set* $\delta = (1+\gamma\beta)/(1+\gamma\mu)$, *and let* $x_0 \in X$. *Then* $\delta \in \,]0,1]$ *and* $\nu \in \,]0,1[$. *Moreover, the following hold*:
(i) $T = \delta((1-\nu)\operatorname{Id} + \nu N)$, N *is nonexpansive.*
(ii) T *is* $(1 - (\delta(1-\nu))/(2-\nu))$*-averaged.*
(iii) T *is* δ*-Lipschitz continuous.*
(iv) *There exists* $\overline{x} \in \operatorname{Fix} T = \operatorname{zer}(A+B)$, *and* $T^n x_0 \rightharpoonup \overline{x}$.

Suppose that $\mu > 1/\beta$. *Then we additionally have*:
(v) T *is a Banach contraction with a constant* $\delta < 1$.
(vi) $\operatorname{zer}(A+B) = \{\overline{x}\}$ *and* $T^n x_0 \to \overline{x}$ *with a linear rate* $\delta < 1$.

Proof Combine Lemma 2.12 and Theorem 6.5 applied with (ω, β) replaced by $(\beta, 2\beta)$. □

Theorem 6.9 *Let* $\mu > \beta > 0$. *Suppose that* A *is* β*-Lipschitz continuous and that* B *is maximally* μ*-monotone. Let* $\overline{\beta} \in \,]\mu + \beta, +\infty[$, *and let* $\gamma \in [2/(\overline{\beta} - 2\beta), 2/(\overline{\beta} - \mu - \beta)[$. *Set* $T = J_{\gamma B}(\operatorname{Id} - \gamma A)$, *set* $\nu = \gamma\overline{\beta}/(2(\gamma\overline{\beta} - \gamma\beta - 1))$, *set* $\delta = (1 + \gamma\beta - \gamma\overline{\beta})/(1+\gamma\mu)$, *and let* $x_0 \in X$. *Then* $\delta \in \,]-1, 0]$ *and* $\nu \in \,]0,1[$. *Moreover, the following hold*:
(i) $T = \delta((1-\nu)\operatorname{Id} + \nu N)$, N *is nonexpansive.*
(ii) T *is a Banach contraction with a constant* $|\delta| < 1$.
(iii) *There exists* $\overline{x} \in X$ *such that* $\operatorname{Fix} T = \operatorname{zer}(A+B) = \{\overline{x}\}$ *and* $T^n x_0 \to \overline{x}$ *with a linear rate* $|\delta| < 1$.

Proof Combine Lemma 2.12 and Theorem 6.6 applied with (ω, β) replaced by $(\beta, 2\beta)$. □

7 Applications to optimization problems

Let $f: X \to]-\infty, +\infty]$, and let $g: X \to]-\infty, +\infty]$. Throughout this section, we shall assume that

$$\boxed{f \text{ and } g \text{ are proper lower semicontinuous functions.}}$$

We shall use ∂f to denote the subdifferential mapping from convex analysis.

Definition 7.1 (see [3, Definition 6.1]) An abstract subdifferential $\partial_{\#}$ associates a subset $\partial_{\#} f(x)$ of X with f at $x \in X$, and it satisfies the following properties:

(i) $\partial_{\#} f = \partial f$ if f is a proper lower semicontinuous convex function;
(ii) $\partial_{\#} f = \nabla f$ if f is continuously differentiable;
(iii) $0 \in \partial_{\#} f(x)$ if f attains a local minimum at $x \in \operatorname{dom} f$;
(iv) for every $\beta \in \mathbb{R}$,

$$\partial_{\#}\left(f + \beta \frac{\|\cdot - x\|^2}{2}\right) = \partial_{\#} f + \beta(\operatorname{Id} - x).$$

The Clarke–Rockafellar subdifferential, Mordukhovich subdifferential, and Frechét subdifferential all satisfy Definition 7.1(i)–(iv), see, e.g., [5, 19, 20], so they are $\partial_{\#}$.

Let $\lambda > 0$. Recall that f is λ-hypoconvex (see [23, 26]) if

$$f\big((1-\tau)x + \tau y\big) \le (1-\tau) f(x) + \tau f(y) + \frac{\lambda}{2}\tau(1-\tau)\|x - y\|^2 \tag{81}$$

for all $(x, y) \in X \times X$ and $\tau \in]0, 1[$ or, equivalently,

$$f + \frac{\lambda}{2}\|\cdot\|^2 \text{ is convex.} \tag{82}$$

For $\gamma > 0$, the *proximal mapping* $\operatorname{Prox}_{\gamma f}$ is defined at $x \in X$ by

$$\operatorname{Prox}_{\gamma f}(x) = \operatorname*{argmin}_{y \in X}\left(f(y) + \frac{\gamma}{2}\|x - y\|^2\right). \tag{83}$$

Fact 7.2 *Suppose that $f: X \to]-\infty, +\infty]$ is a proper lower semicontinuous λ-hypoconvex function. Then*

$$\partial_{\#} f = \partial\left(f + \frac{\lambda}{2}\|\cdot\|^2\right) - \lambda \operatorname{Id}. \tag{84}$$

Moreover, we have:

(i) *The Clarke–Rockafellar, Mordukhovich, and Frechét subdifferential operators of f all coincide.*
(ii) *$\partial_{\#} f$ is maximally $-\lambda$-monotone.*
(iii) $(\forall \gamma \in]0, 1/\lambda[)$ $\operatorname{Prox}_{\gamma f}$ *is single-valued and* $\operatorname{dom} \operatorname{Prox}_{\gamma f} = X$.

Proof See [3, Proposition 6.2 and Proposition 6.3]. □

Proposition 7.3 *Let $\mu \geq \omega \geq 0$. Suppose that* $\operatorname{argmin}(f+g) \neq \varnothing$ *and that one of the following conditions is satisfied*:

(i) *f is μ- strongly convex, g is ω- hypoconvex.*

(ii) *f is ω- hypoconvex, and g is μ- strongly convex.*

Then $f+g$ is convex and $\partial_{\#}(f+g) = \partial(f+g)$.

If, in addition, one of the following conditions is satisfied:

(a) $0 \in \operatorname{sri}(\operatorname{dom} f - \operatorname{dom} g)$.

(b) *X is finite dimensional and* $0 \in \operatorname{ri}(\operatorname{dom} f - \operatorname{dom} g)$.

(c) *X is finite dimensional, f and g are polyhedral, and* $\operatorname{dom} f \cap \operatorname{dom} g \neq \varnothing$.

Then

$$\partial_{\#}(f+g) = \partial(f+g) = \partial_{\#} f + \partial_{\#} g, \tag{85}$$

and

$$\operatorname{zer} \partial_{\#}(f+g) = \operatorname{zer}(\partial_{\#} f + \partial_{\#} g) = \operatorname{argmin}(f+g). \tag{86}$$

Proof It is clear that either (i) or (ii) implies that $f+g$ is convex, and the identity follows in view of Definition 7.1(i). Now, suppose that (i) holds along with one of the assumptions (a)–(c). Rewrite f and g as $(f,g) = (\overline{f} + \frac{\mu}{2}\|\cdot\|^2, \overline{g} - \frac{\omega}{2}\|\cdot\|^2)$ and observe that both $\overline{f}$ and $\overline{g}$ are convex, as is $\overline{f} + \overline{g}$. Moreover, we have $\operatorname{dom} f = \operatorname{dom} \overline{f}$ and $\operatorname{dom} g = \operatorname{dom} \overline{g}$. Now,

$$\begin{aligned}
\partial_{\#}(f+g) &= \partial_{\#}\left(\overline{f} + \overline{g} + \frac{\mu-\omega}{2}\|\cdot\|^2\right) && (87\text{a})\\
&= \partial_{\#}(\overline{f} + \overline{g}) + (\mu-\omega)\operatorname{Id} = \partial(\overline{f} + \overline{g}) + (\mu-\omega)\operatorname{Id} && (87\text{b})\\
&= \partial\overline{f} + \partial\overline{g} + (\mu-\omega)\operatorname{Id} = \partial\overline{f} + \mu\operatorname{Id} + \partial\overline{g} - \omega\operatorname{Id} && (87\text{c})\\
&= \partial f + \partial_{\#} g = \partial_{\#} f + \partial_{\#} g. && (87\text{d})
\end{aligned}$$

Here, (87b) follows from applying Definition 7.1(iv) to $\overline{f} + \overline{g}$, (87c) follows from [2, Theorem 16.47] applied to $\overline{f}$ and $\overline{g}$, and (87c) follows from applying Fact 7.2 to f and g and using Definition 7.1(i), which verify (85). Finally, (86) follows from combining (85) and [2, Theorem 16.3]. □

The following theorem provides an alternative proof to [17, Theorem 4.4] and [9, Theorem 5.4(ii)].

Theorem 7.4 *Let $\mu > \omega \geq 0$, and let $\gamma \in]0, (\mu-\omega)/(2\mu\omega)[$. Suppose that one of the following holds*:

(i) *f is μ- strongly convex, g is ω- hypoconvex.*

(ii) *f is ω-hypoconvex, and g is μ-strongly convex,*

and that $0 \in \partial_{\#} f + \partial_{\#} g$ (see Proposition 7.3 for sufficient conditions). Set

$$T = \frac{1}{2}\big(\operatorname{Id} + (2\operatorname{Prox}_{\gamma g} - \operatorname{Id})(2\operatorname{Prox}_{\gamma f} - \operatorname{Id})\big) \quad \textit{and} \quad \alpha = \frac{\mu-\omega}{2(\mu-\omega-\gamma\mu\omega)}, \tag{88}$$

and let $x_0 \in X$. Then $\alpha \in]0,1[$, and T is α-averaged. Moreover, $(\exists \overline{x} \in \operatorname{Fix} T)$ such that $T^n x_0 \rightharpoonup \overline{x}$, $\operatorname{argmin}(f+g) = \{\operatorname{Prox}_f \overline{x}\}$, *and* $\operatorname{Prox}_f T^n x_0 \to \operatorname{Prox}_f \overline{x}$.

Proof Suppose that (i) holds. Then [2, Example 22.4] (respectively Fact 7.2(ii)) implies that $\partial_{\#} f = \partial f$ (respectively $\partial_{\#} g$) is maximally μ-monotone (respectively maximally $(-\omega)$-monotone). The conclusion follows from applying Theorem 5.5(i) with (A, B) replaced by $(\partial_{\#} f, \partial_{\#} g)$. The proof for (ii) follows similarly by using Theorem 5.5(ii). □

Before we proceed further, we recall the following useful fact.

Fact 7.5 (Baillon–Haddad) *Let $f : X \to \mathbb{R}$ be a Frechét differentiable convex function, and let $\beta > 0$. Then ∇f is β-Lipschitz continuous if and only if ∇f is $\frac{1}{\beta}$-cocoercive.*

Proof See, e.g., [2, Corollary 18.17]. □

Lemma 7.6 *Let $\mu \geq 0$, let $\beta > 0$, and let $f : X \to \mathbb{R}$ be a Frechét differentiable function. Suppose that f is μ-strongly convex with a β-Lipschitz continuous gradient. Then the following hold*:

(i) *$f - \frac{\mu}{2}\|\cdot\|^2$ is convex.*
(ii) *∇f is maximally μ-monotone.*
(iii) *$\nabla f - \mu \operatorname{Id}$ is $\frac{1}{\beta}$-cocoercive.*

Proof (i): See, e.g., [2, Proposition 10.8]. (ii): See, e.g., [2, Example 22.4(iv)]. (iii): Combine (i), Lemma 2.10, and Corollary 2.14(ii) applied with (f_1, f_2) replaced by $(f, \frac{\mu}{2}\|\cdot\|^2)$. □

Theorem 7.7 (The forward–backward algorithm when f is μ-strongly convex) *Let $\mu \geq \omega \geq 0$, and let $\beta > 0$. Let f be μ-strongly convex and Frechét differentiable with a β-Lipschitz continuous gradient, and let g be ω-hypoconvex. Suppose that* $\operatorname{argmin}(f + g) \neq \varnothing$. *Let $\gamma \in]0, 2/(\beta + 2\mu)[$, and set $\delta = (1 - \gamma\mu)/(1 - \gamma\omega)$. Set $T = \operatorname{Prox}_{\gamma g}(\operatorname{Id} - \gamma \nabla f)$, and let $x_0 \in X$. Then the following hold*:

(i) *There exists $\overline{x} \in \operatorname{Fix} T = \operatorname{zer}(A + B) = \operatorname{argmin}(f + g)$ such that $T^n x_0 \rightharpoonup \overline{x}$.*

Suppose that $\mu > \omega$. Then we additionally have:

(ii) $\operatorname{Fix} T = \operatorname{argmin}(f + g) = \{\overline{x}\}$ *and $T^n x_0 \to \overline{x}$ with a linear rate $\delta < 1$.*

Proof Note that Definition 7.1(ii) implies that $\partial_{\#} f = \nabla f$. Set $(A, B) = (\nabla f, \partial_{\#} g)$ and observe that Proposition 7.3 and Proposition 2.2 imply that $\operatorname{Fix} T = \operatorname{zer}(A + B) = \operatorname{argmin}(f + g)$. It follows from [2, Example 22.4] (respectively Fact 7.2(ii)) that A (respectively B) is maximally μ-monotone (respectively maximally $(-\omega)$-monotone). Moreover, Lemma 7.6(iii) implies that $A - \mu \operatorname{Id}$ is $\frac{1}{\beta}$-cocoercive. (i)–(ii): Apply Theorem 6.1(iv)&(vi). □

To proceed to the next result, we need the following lemma.

Lemma 7.8 *Let $\omega \geq 0$, let $\beta > 0$, and let $f : X \to \mathbb{R}$ be a Frechét differentiable function. Suppose that g is ω-hypoconvex with a $\frac{1}{\beta}$-Lipschitz continuous gradient. Then $\nabla f + \omega \operatorname{Id}$ is $\beta/(1 + \omega\beta)$-cocoercive.*

Theorem 7.9 (The forward–backward algorithm when f is ω-hypoconvex) *Let $\mu \geq \omega \geq 0$, let $\beta > 0$, and let $\overline{\beta} \in]\max\{\beta, 2\omega\}, +\infty[$. Let f be ω-hypoconvex, and let g be μ-strongly convex and Frechét differentiable with a β-Lipschitz continuous gradient. Suppose that* $\operatorname{argmin}(f + g) \neq \varnothing$. *Let $\gamma \in]0, 2/(\overline{\beta} - 2\omega)[$, and set $\delta = (1 + \gamma\omega)/(1 + \gamma\mu)$. Set $T = \operatorname{Prox}_{\gamma g}(\operatorname{Id} - \gamma \nabla f)$, and let $x_0 \in X$. Then the following hold*:

(i) *There exists $\overline{x} \in \operatorname{Fix} T = \operatorname{argmin}(f + g)$ such that $T^n x_0 \rightharpoonup \overline{x}$.*

Suppose that $\mu > \omega$. Then we additionally have:

(ii) $\operatorname{Fix} T = \operatorname{argmin}(f+g) = \{\overline{x}\}$ *and* $T^n x_0 \to \overline{x}$ *with a linear rate* $\delta < 1$.

Proof Proceed similar to the proof of Theorem 7.7 but use Theorem 6.5(iv)&(vi). □

Theorem 7.10 (The forward–backward algorithm when f is $1/\beta$-hypoconvex) *Let* $\mu \geq \beta > 0$, *and let* $\overline{\beta} \in]2\beta, +\infty[$. *Let* f *be* μ*-strongly convex, and let* g *be Frechét differentiable with a* β*-Lipschitz continuous gradient. Suppose that* $\operatorname{argmin}(f+g) \neq \varnothing$. *Let* $\gamma \in]0, 2/(\overline{\beta} - 2\beta)\}[$, *and set* $\delta = (1+\gamma\beta)/(1+\gamma\mu)$. *Set* $T = \operatorname{Prox}_{\gamma g}(\operatorname{Id} - \gamma\nabla f)$, *and let* $x_0 \in X$. *Then the following hold*:

(i) *There exists* $\overline{x} \in \operatorname{Fix} T = \operatorname{argmin}(f+g)$ *such that* $T^n x_0 \rightharpoonup \overline{x}$.

Suppose that $\mu > 1/\beta$. *Then we additionally have*:

(ii) $\operatorname{Fix} T = \operatorname{argmin}(f+g) = \{\overline{x}\}$ *and* $T^n x_0 \to \overline{x}$ *with a linear rate* $\delta < 1$.

Proof Combine Lemma 2.12 applied with A replaced by ∇f and Theorem 7.9 applied with (ω, β) replaced by $(\beta, 2\beta)$. □

Remark 7.11 The results of Theorem 6.2, Theorem 6.6, and Theorem 6.9 can be directly applied to optimization settings in a similar fashion à la Theorem 7.7, Theorem 7.9, and Theorem 7.10.

8 Graphical characterizations

This section contains 2D-graphical representations of different Lipschitz continuous operator classes that admit I-N decompositions and of their composition classes. We illustrate exact shapes of the composition classes in 2D and conservative estimates from Theorem 3.4 and Theorem 4.2. Similar graphical representations have appeared before in the literature. In [10, 11], nonexpansiveness and firm nonexpansiveness ($\frac{1}{2}$-averagedness) are characterized. Early preprints of [15] have more 2D graphical representations, and the lecture notes [14] contain many such characterizations with the purpose of illustrating how different properties relate to each other and to provide intuition on why different algorithms converge. This has been further extended and formalized in [24]. Not only do these illustrations provide intuition. Indeed, it is a straightforward consequence of, e.g., [24, 25] that for compositions of two operator classes that admit I-N decompositions, there always exists a 2D-worst case. Hence, if the 2D illustration implies that the composition class admits a specific (α, β)-I-N decomposition, so does the full operator class.

In Sect. 8.1, we characterize many well-known special cases of operator classes that admit I-N decompositions. In Sect. 8.2, we characterize classes obtained by compositions of such operator classes and highlight differences between the true composition classes and their characterizations using Theorem 3.4.

8.1 Single operators

We consider classes of (α, β)-I-N decomposition of Lipschitz continuous operators. We graphically illustrate properties of some special cases. The illustrations should be read as follows. Assume that $x - y$ is represented by the marker in the figure. The diagram then shows where $Rx - Ry$ can end up in relation to $x - y$. If the point $x - y$ is rotated in the picture, the rest of the picture rotates with it. The characterization is, by construction of (α, β)-I-N decompositions, always a circle of radius $\beta\|x - y\|$ shifted $\alpha\|x - y\|$ along the line defined by the origin and the point $x - y$.

Lipschitz continuous operators Let $\beta > 0$ and let $R\colon X \to X$. Then R is β-Lipschitz continuous if and only if R admits an (α,β)-I-N decomposition, with α chosen as 0. Figure 1 shows the case $\beta = 0.8$. The radius of the Lipschitz circle is $\beta\|x-y\|$.

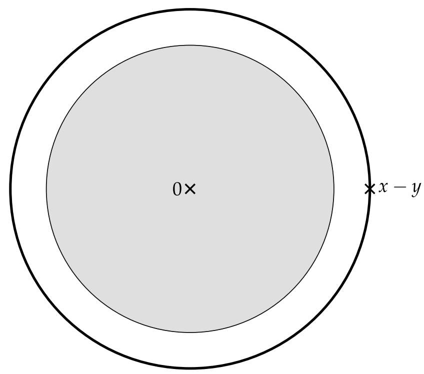

Figure 1 Illustration of β-Lipschitz continuous operator with $\beta = 0.8$

Cocoercive operators Let $\beta > 0$, and let $R\colon X \to X$. Then R is $\frac{1}{\beta}$-cocoercive if and only if R admits an (α,β)-I-N decomposition, with (α,β) chosen as $(\frac{\beta}{2},\frac{\beta}{2})$. Figure 2 shows the cases $\beta = 1.4$ and $\beta = 0.7$. The diameter is $\beta\|x-y\|$. The figure clearly illustrates that $\frac{1}{\beta}$-cocoercive operators are also β-Lipschitz (but not necessarily the other way around).

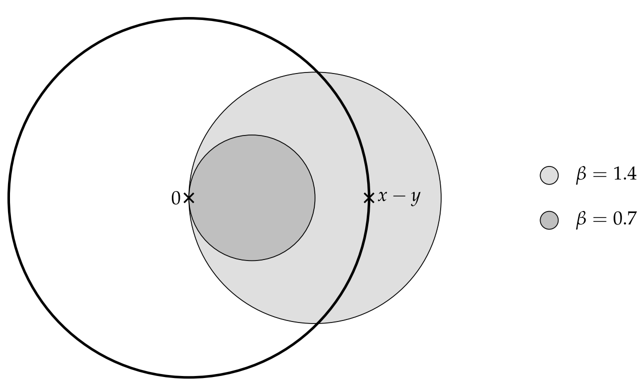

Figure 2 Illustration of $\frac{1}{\beta}$-cocoercive operators with $\beta = 0.7$ and $\beta = 1.4$

Averaged operators Let $\alpha \in]0,1[$, and let $R\colon X \to X$. Then R is α-averaged if and only if R admits an (α,β)-I-N decomposition, with (α,β) chosen as $(1-\alpha,\alpha)$. Figure 3 shows the cases $\alpha = 0.25$ and $\alpha = 0.5$, and $\alpha = 0.75$. All averaged operators are nonexpansive.

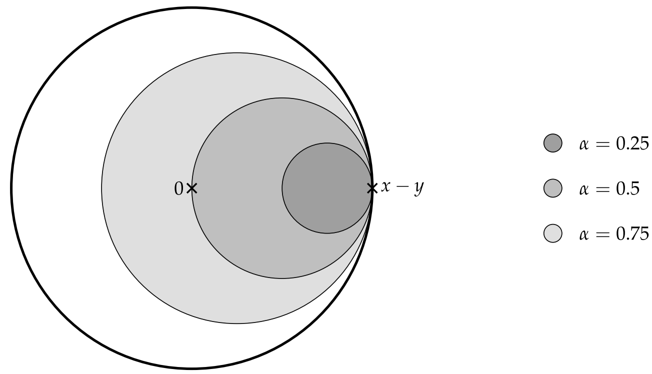

Figure 3 Illustration of α-averaged operators with $\alpha = 0.25$, $\alpha = 0.5$, and $\alpha = 0.75$

Conic operators Let $\alpha > 0$, and let $R: X \to X$. Then R is α-conically nonexpansive if and only if R admits an (α, β)-I-N decomposition, with (α, β) chosen as $(1 - \alpha, \alpha)$. Figure 4 shows the cases $\alpha = 1.2$ and $\alpha = 1.5$. Conically nonexpansive operators fail to be nonexpansive for $\alpha > 1$.

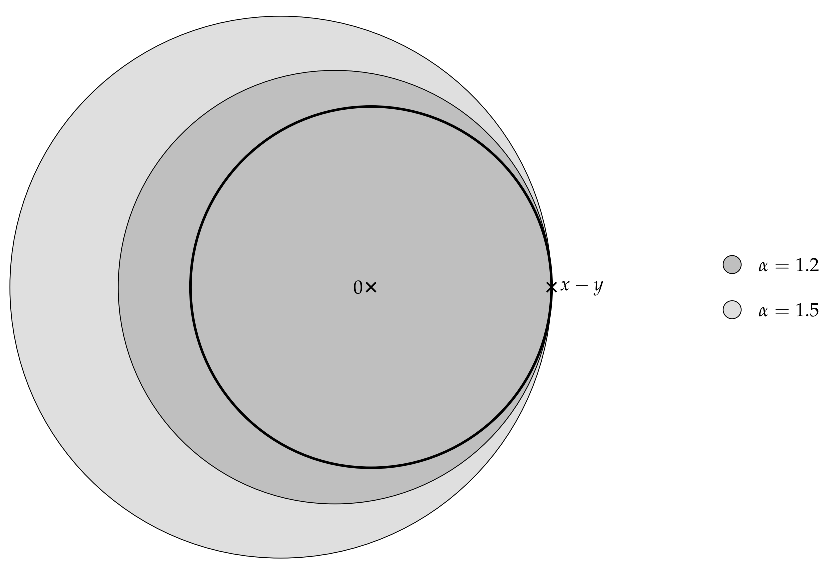

Figure 4 Illustration of α-conically nonexpansive operators with $\alpha = 1.2$ and $\alpha = 1.5$

μ-Monotone operators Let $\mu \in \mathbb{R}$, and suppose that $A: X \rightrightarrows X$ is μ-monotone. The shortest distance between the vertical line and the origin in the illustration is $|\mu|\|x - y\|$. Figure 5 shows the case $\mu = 0.2$.

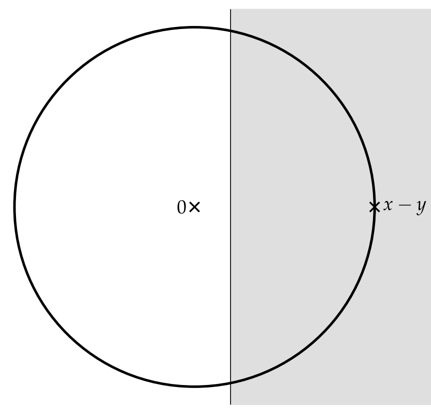

Figure 5 Illustration of μ-monotone operator with $\mu = 0.2$

8.2 Compositions of two operators

In this section, we provide illustrations of compositions of different classes of Lipschitz continuous operators. We consider compositions of the form

$$R = R_2R_1, \quad \text{where } R_i \text{ admits an } (\alpha_i, \beta_i)\text{-I-N decomposition,}$$

$\forall i \in \{1,2\}$. Let $(x,y) \in X \times X$. We illustrate the regions within which $R_2R_1x - R_2R_1y$ can end up. For most considered composition classes, we provide two illustrations. The left illustration explicitly shows how the composition is constructed. It shows the region within which $R_1x - R_1y$ must end up. The second operator R_2 is applied at a subset, marked by crosses, of boundary points of that region. Given these as starting points for R_2 application, the dashed circles show where $R_2R_1x - R_2R_1y$ can end up for this subset. The right illustration shows, in gray, the resulting exact shape of the composition. It also contains the estimate from Theorem 3.4 that provides an I-N decomposition of the composition. From these illustrations, it is obvious that many different I-N decomposition are valid. The illustrations also reveal that the specific I-N decompositions provided in Theorem 3.4 indeed are suitable for our purpose of characterizing the composition as averaged, conic, or contractive.

Averaged-averaged composition We first consider α_i-averaged R_i with $\alpha_i \in]0,1[$. A special case is the forward–backward splitting operator $T = J_{\gamma B}(\mathrm{Id} - \gamma A)$ with $\frac{1}{\beta}$-cocoercive A and maximally monotone B. This implies that $(\mathrm{Id} - \gamma A)$ is $\frac{\gamma\beta}{2}$-averaged for $\gamma \in]0, \frac{2}{\beta}[$ and that $J_{\gamma B}$ is $\frac{1}{2}$-averaged. The example in Fig. 6 has individual averagedness parameters $\alpha_1 = 0.5$ and $\alpha_2 = 0.5$, i.e., $R = R_2R_1$ with $R_1 = 0.5\mathrm{Id} + 0.5N_1$ and $R_2 = 0.5\mathrm{Id} + 0.5N_2$. Theorem 3.4 shows that the composition is of the form $0.33\mathrm{Id} + 0.67N$, where N is nonexpansive, i.e., it is 0.67-averaged. The fact that the composition is averaged is already known, see [8, 12].

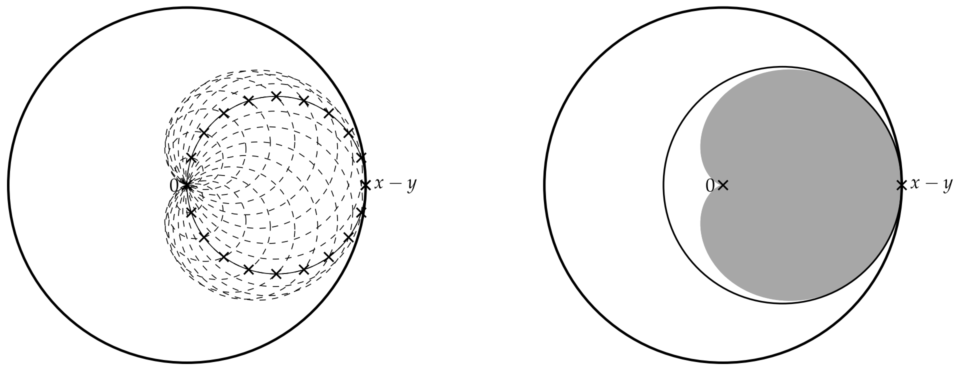

Figure 6 Illustration of composition of α_1-averaged and α_2-averaged operators with $\alpha_1 = \alpha_2 = 0.5$

The example in Fig. 7 shows $\alpha_1 = 0.7$ and $\alpha_2 = 0.6$. Theorem 3.4 shows that the composition is of the form $0.21\mathrm{Id} + 0.79N$, where N is nonexpansive, i.e., it is 0.79-averaged.

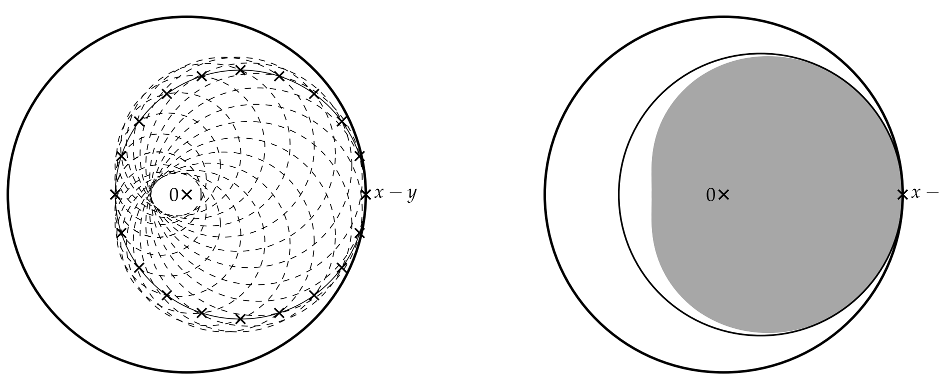

Figure 7 Illustration of composition of α_1-averaged and α_2-averaged operators with $\alpha_1 = 0.7$ and $\alpha_2 = 0.6$

Conic-conic composition We consider α_i-averaged R_i with $\alpha_i > 0$. Several examples with this setting are considered in for Douglas-Rachford splitting and forward–backward splitting in Sect. 5 and Sect. 6. We know from Theorem 4.2 that the composition is conic if $\alpha_1\alpha_2 < 1$. The example in Fig. 8 has $\alpha_1 = 1.7$ and $\alpha_2 = 0.45$, that satisfies $\alpha_1\alpha_2 = 0.76 < 1$. Theorem 4.2 shows that the composition is of the form $-1.64\mathrm{Id} + 2.64N$, where N is nonexpansive, i.e., it is 2.64-conic.

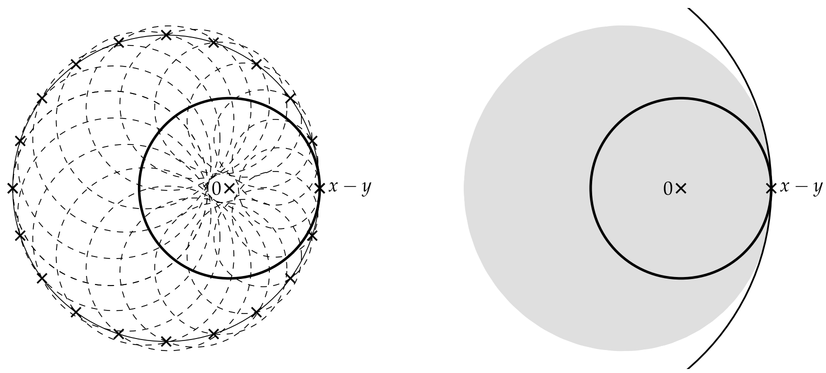

Figure 8 Illustration of composition of α_1-conic operator and α_2-averaged operator with $\alpha_1 = 1.7$ and $\alpha_2 = 0.45$

In Example 4.6, we have shown that the assumption $\alpha_1\alpha_2 < 1$ is critical for the composition to be conic. Figure 9 illustrates the case $\alpha_1 = 1.7$ and $\alpha_2 = 0.7$, which satisfies $\alpha_1\alpha_2 = 1.19 > 1$, hence Theorem 4.2 cannot be used to deduce that the composition is conic. Indeed, we see from the figure that the composition is not conic. It is impossible to draw a circle that touches the marker at $x - y$ and extends only to the left.

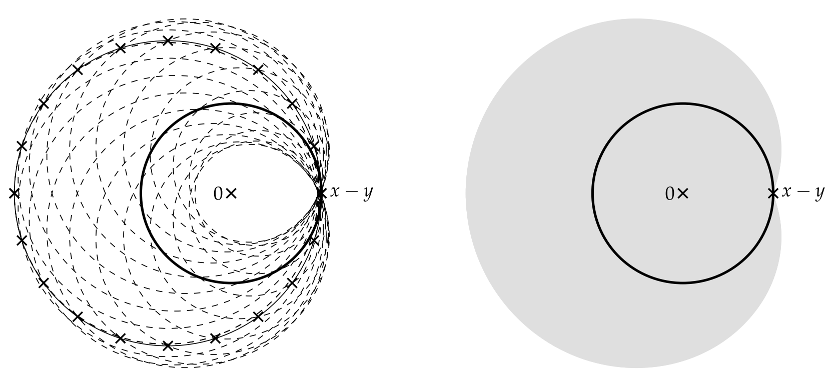

Figure 9 Illustration of composition of α_1-conic operator and α_2-averaged operator with $\alpha_1 = 1.7$ and $\alpha_2 = 0.7$

We conclude the conic composed with conic examples with a forward–backward example. The forward–backward splitting operator $J_{\gamma B}(\mathrm{Id} - \gamma A)$ with A $\frac{1}{\beta}$-cocoercive and B (maximally) monotone is composed of $\frac{1}{2}$-averaged resolvent $J_{\gamma B}$ and $\frac{\gamma\beta}{2}$-conic forward step $(\mathrm{Id} - \gamma A)$. The composition $R = R_2R_1$ with R_i α_i-conic is conic if $\alpha_1\alpha_2 < 1$, Theorem 4.2. In the forward–backward setting, this corresponds to $\gamma \in (0, \frac{4}{\beta})$, which doubles the allowed range compared to guaranteeing an averaged composition. This extended range has been shown before, e.g., in [13, 18].

In Fig. 10, we illustrate the forward–backward setting with $\gamma = \frac{3.9}{\beta}$. This corresponds to conic parameters $\alpha_1 = 1.95$ and $\alpha_2 = 0.5$, i.e., $R = R_2R_1$ with $R_1 = -0.95\mathrm{Id} + 1.95N_1$ and $R_2 = 0.5\mathrm{Id} + 0.5N_2$. The composition is of the form $-18.99\mathrm{Id} + 19.99N$, where N is nonexpansive, i.e., it is 19.99-conic, Theorem 4.2. The left figure shows the resulting composition and (parts of) the conic approximation. The conic approximation is very large compared to the actual region. This is due to the local behavior around the point $x - y$, where it is almost vertical. As $\gamma \nearrow 4\beta$, the exact shape approaches being vertical around $x-y$ and the conic circle approaches to have an infinite radius. For $\gamma > 4\beta$, the exact shape extends to the right of $x-y$ (as in the figure above), and the composition will not be conic.

In the right figure, we consider the relaxed forward–backward map $(1-\theta)\mathrm{Id} + \theta J_{\gamma B}(\mathrm{Id} - \gamma A)$ with $\theta > 0$. If the composition $J_{\gamma B}(\mathrm{Id} - \gamma A)$ is α-conic, it is straightforward to verify that the relaxed map is $\theta\alpha$-conic. Therefore, any $\theta \in (0, \alpha^{-1})$ gives an $\theta\alpha$-averaged relaxed forward–backward map. An averaged map is needed to guarantee convergence to a fixed-point when iterated. In the figure, we let $\theta = 0.04$, which satisfies $\theta < \alpha^{-1} \approx 0.05$. The approximation is indeed averaged, but the region within which the composition can end up is very small compared to the conic approximation.

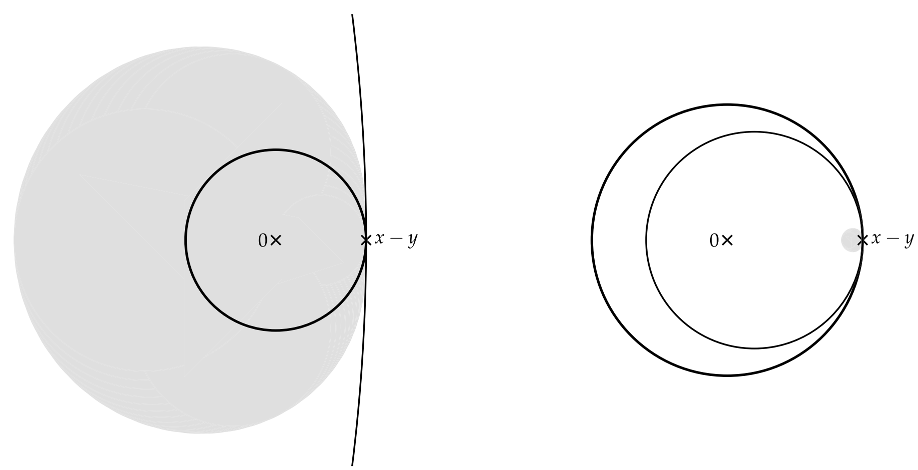

Figure 10 To the left is an illustration of the forward–backward composition $J_{\gamma B}(\mathrm{Id} - \gamma A)$ with $\gamma = \frac{3.9}{\beta}$, where $\frac{1}{\beta}$ is the cocoercivity constant of A. It is a composition between an α_1-conic operator and an α_2-averaged operator with $\alpha_1 = 1.95$ and $\alpha_2 = 0.5$. To the right is an illustration of a θ-relaxation of the same forward-backward map with $\theta = 0.04$

Scaled averaged and cocoercive compositions Compositions of scaled averaged and cocoercive operators are also special cases of scaled conic composed with scaled conic operators treated in Theorem 4.2. It covers the forward backward examples in Sect. 6, where identity is shifted between the operators and the sum is (strongly) monotone. The operators in the composition are of the form $R_1 = \delta_1((1-\alpha_1)\mathrm{Id} + \alpha_1 N_1)$ and $R_2 = \frac{\beta_2}{2}(\mathrm{Id} + N_2)$, where $\alpha_1 \in (0,1)$, $\delta_1 > 0$, and $\beta_2 > 0$.

In Fig. 11, we consider the forward–backward setting in Theorem 6.5. The forward backward map is $J_{\gamma B}(\mathrm{Id} - \gamma A)$ and we let $A + 0.3\mathrm{Id}$ be 1-cocoercive, B be maximally 0.3-monotone. That is, we have shifted 0.3Id from A to B and the sum is monotone. We use step-length $\gamma = 2$. The proof of Theorem 6.5 shows that, in our setting, R_1 is 1.6-scaled 0.62-averaged and that R_2 is 1.6-cocoercive. Theorem 3.4 implies that the composition is of the form $0.27\mathrm{Id} + 0.73N$, where N is nonexpansive, i.e., it is 0.73-averaged.

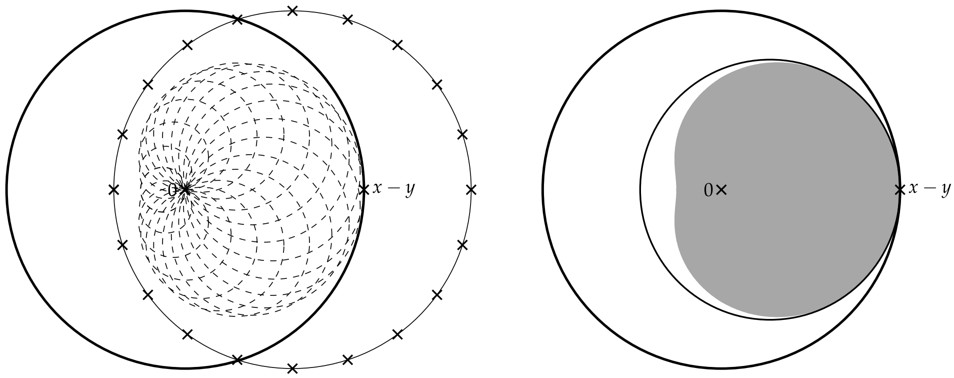

Figure 11 Illustration of composition of 1.6-scaled 0.62-averaged operator with 1.6-cocoercive operator. The composition comes from the forward–backward map $J_{\gamma B}(\mathbf{Id} - \gamma A)$ with $A + 0.3\mathbf{Id}$ 1-cocoercive, B 0.3-monotone, and $\gamma = 2$

Figure 12 considers a similar forward–backward setting, but with a strongly monotone sum. We let $A + 0.2\mathrm{Id}$ be 1-cocoercive, B be maximally 0.3-monotone, which implies that the sum is 0.1-strongly monotone. We keep step-length $\gamma = 2$. The proof of Theorem 6.5 shows that, in our setting, R_1 is 1.4-scaled 0.62-averaged and that R_2 is 1.6-cocoercive. Theorem 3.4 implies that the composition is of the form $0.19\mathrm{Id} + 0.68N$, where N is nonexpansive, i.e., it is 0.87-contractive.

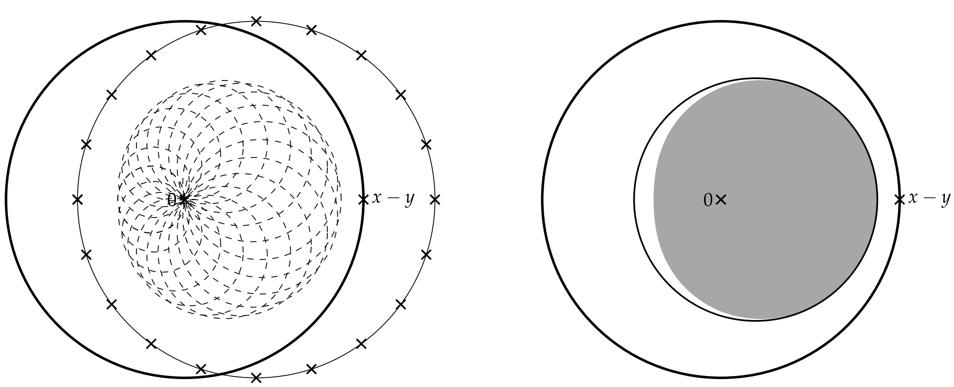

Figure 12 Illustration of composition of 1.4-scaled 0.62-averaged operator with 1.6-cocoercive operator. The composition comes from the forward–backward map $J_{\gamma B}(\mathbf{Id} - \gamma A)$ with $A + 0.2\mathbf{Id}$ 1-cocoercive, B 0.3-monotone, and $\gamma = 2$

The final example in Fig. 13 considers a similar forward–backward setting where the sum is not monotone. We let $A + 0.4\mathrm{Id}$ be 1-cocoercive, B be maximally 0.3-monotone, which implies that the sum is -0.1-monotone, i.e., it is not monotone. We use step-length $\gamma = 2$. The proof of Theorem 6.5 shows that, in our setting, R_1 is 1.8-scaled 0.62-averaged and that R_2 is 1.6-cocoercive. Theorem 3.4 implies that the composition is of the form $0.35\mathrm{Id} + 0.78N$, where N is nonexpansive, i.e., it is 1.12-Lipschitz and not conic, averaged, or contractive.

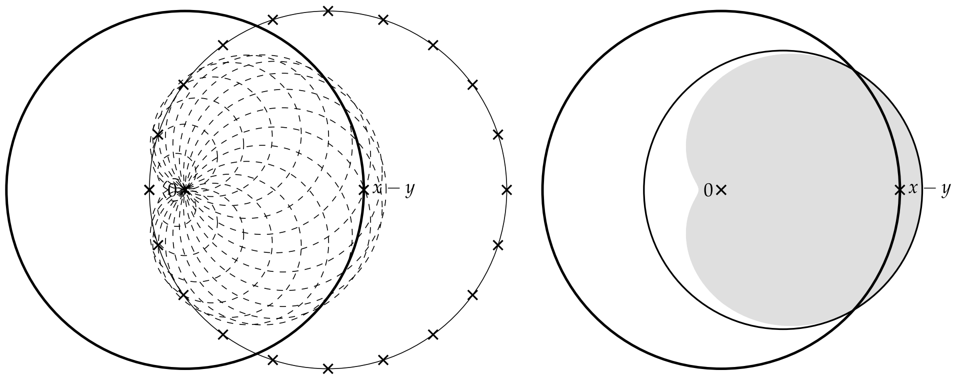

Figure 13 Illustration of composition of 1.8-scaled 0.62-averaged operator with 1.6-cocoercive operator. The composition comes from the forward–backward map $J_{\gamma B}(\mathbf{Id} - \gamma A)$ with $A + 0.4\mathbf{Id}$ 1-cocoercive, B 0.3-monotone, and $\gamma = 2$

Appendix A

Proof of Lemma 2.3 Indeed, observe that

$$R(\lambda) = (1 - 2\lambda)\operatorname{Id} + \lambda(\operatorname{Id} + R_2R_1) \tag{89}$$

and

$$\operatorname{Id} - R(\lambda) = \lambda(\operatorname{Id} - R_2R_1). \tag{90}$$

In view of (89) and (90) we have

$$\begin{aligned}
&\big\langle R(\lambda)x - R(\lambda)y \mid \big(\operatorname{Id} - R(\lambda)\big)x - \big(\operatorname{Id} - R(\lambda)\big)y\big\rangle \\
&\quad = (1 - 2\lambda)\big\langle x - y \mid \operatorname{Id} - R(\lambda))x - \big(\operatorname{Id} - R(\lambda)\big)y\big\rangle \\
&\qquad + \lambda^2\big\langle (x - y) - (R_2R_1x - R_2R_1y) \mid (x - y) + (R_2R_1x - R_2R_1y)\big\rangle \\
&\quad = (1 - 2\lambda)\big\langle x - y \mid \operatorname{Id} - R(\lambda))x - \big(\operatorname{Id} - R(\lambda)\big)y\big\rangle + \lambda^2\big(\|x - y\|^2 - \|R_2R_1x - R_2R_1y\|^2\big) \\
&\quad = (1 - 2\lambda)\big\langle x - y \mid \operatorname{Id} - R(\lambda))x - \big(\operatorname{Id} - R(\lambda)\big)y\big\rangle \\
&\qquad + \lambda^2\big(\|x - y\|^2 - \|R_1x - R_1y\|^2 + \|R_1x - R_1y\|^2 - \|R_2R_1x - R_2R_1y\|^2\big) \\
&\quad = (1 - 2\lambda)\big\langle x - y \mid \big(\operatorname{Id} - R(\lambda)\big)x - \big(\operatorname{Id} - R(\lambda)\big)y\big\rangle \\
&\qquad + \lambda^2\big\langle (\operatorname{Id} + R_1)x - (\operatorname{Id} + R_1)y \mid (\operatorname{Id} - R_1)x - (\operatorname{Id} - R_1)y\big\rangle \\
&\qquad + \lambda^2\big\langle (\operatorname{Id} + R_2)R_1x - (\operatorname{Id} + R_2)R_1y \mid (\operatorname{Id} - R_2)R_1x - (\operatorname{Id} - R_2)R_1y\big\rangle,
\end{aligned}$$

and the conclusion follows. □

Appendix B

Proof of Lemma 2.12 (i): Because $\frac{1}{\beta}A$ is nonexpansive, we learn from [2, Example 20.7] that $\operatorname{Id} + \frac{1}{\beta}A$, as is $\beta\operatorname{Id} + A$, is maximally monotone. The conclusion now follows in view of e.g., [3, Lemma 2.5]. (ii): This is clear by observing that $\frac{1}{2\beta}(\beta\operatorname{Id} + A) = \frac{1}{2}(\operatorname{Id} + \frac{1}{\beta}A)$. □

Appendix C

Proof of Lemma 2.13 Indeed, by assumption, there exist nonexpansive mappings $N_1: X \to X$ and $N_2: X \to X$ such that

$$T_1 = \frac{\beta}{2}\operatorname{Id} + \frac{\beta}{2}N_1, \qquad T_2 = \frac{\delta}{2}\operatorname{Id} + \frac{\delta}{2}N_2. \tag{91}$$

Now,

$$\frac{1}{\beta}(T_1 - T_2) = \frac{1}{\beta}T_1 - \frac{1}{\beta}T_2 = \frac{1}{2}\operatorname{Id} + \frac{1}{2}N_1 - \frac{\delta}{2\beta}\operatorname{Id} - \frac{\delta}{2\beta}N_2 \tag{92a}$$

$$= \frac{\beta-\delta}{2\beta}\operatorname{Id} + \frac{1}{2}N_1 - \frac{\delta}{2\beta}N_2. \tag{92b}$$

Using the triangle inequality, one can directly verify that $\frac{1}{\beta}(T_1 - T_2)$ is Lipschitz continuous with a constant $\frac{\beta-\delta}{2\beta} + \frac{1}{2} + \frac{\delta}{2\beta} = 1$. The proof is complete. □

Appendix D

Proof of Corollary 2.14 (i): It follows from Fact 7.5 that ∇f_1 (respectively ∇f_2) is $\frac{1}{\beta}$-cocoercive (respectively $\frac{1}{\delta}$-cocoercive). Now apply Lemma 2.13 with (T_1, T_2) replaced by $(\nabla f_1, \nabla f_2)$. (ii): Combine (i) with Fact 7.5 applied with f replaced by $f_1 - f_2$. □

Appendix E

Proof of Lemma 2.15 (i): Indeed, we have $\delta T = (1-(1-\delta(1-\alpha)))\operatorname{Id} + \delta\alpha N = (1-(1-\delta(1-\alpha)))\operatorname{Id} + (1-\delta(1-\alpha))\widetilde{N}$, where $\widetilde{N} = ((\delta\alpha)/(1-\delta(1-\alpha)))N$. Note that $(\delta\alpha)/(1-\delta(1-\alpha) \leq 1$, hence $\widetilde{N}$ is nonexpansive and the conclusion follows. (ii): Clear. □

Acknowledgements
Not applicable.

Funding
PG was partially supported by the Swedish Research Council and the Wallenberg AI, Autonomous Systems and Software Program (WASP). WMM was partially supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant (NSERC-DG).

Availability of data and materials
Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Declarations

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
All authors contributed equally in writing this article. All authors read and approved the final manuscript.

Author details
[1]Department of Automatic Control, Lund University, Lund, Sweden. [2]Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada. [3]Mathematics Department, Faculty of Science, Mansoura University, Mansoura 35516, Egypt.

Publisher's Note

Received: 15 April 2021 Accepted: 17 October 2021 Published online: 20 December 2021

References

1. Bartz, S., Dao, M.N., Phan, H.M.: Conical averagedness and convergence analysis of fixed point algorithms. arXiv preprint (2019). arXiv:1910.14185
2. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces, 2nd edn. Springer, New York (2017)
3. Bauschke, H.H., Moursi, W.M., Wang, X.: Generalized monotone operators and their averaged resolvents. Math. Program., Ser. B (2020). https://doi.org/10.1007/s10107-020-01500-6
4. Burachik, R.S., Iusem, A.N.: Set-Valued Mappings and Enlargements of Monotone Operators, vol. 8. Springer, Berlin (2007)
5. Clarke, F.H.: Optimization and Nonsmooth Analysis, vol. 5. SIAM, Philadelphia (1990)
6. Combettes, P.L.: Solving monotone inclusions via compositions of nonexpansive averaged operators. Optimization **53**(5–6), 475–504 (2004)
7. Combettes, P.L., Pennanen, T.: Proximal methods for cohypomonotone operators. SIAM J. Control Optim. **43**(2), 731–742 (2004)
8. Combettes, P.L., Yamada, I.: Compositions and convex combinations of averaged nonexpansive operators. J. Math. Anal. Appl. **425**(1), 55–70 (2015)
9. 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)
10. Eckstein, J.: Splitting methods for monotone operators with applications to parallel optimization. Ph.D. thesis, MIT (1989)
11. Eckstein, J., Bertsekas, D.P.: On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators. Math. Program. **55**(1), 293–318 (1992)
12. Giselsson, P.: Tight global linear convergence rate bounds for Douglas–Rachford splitting. J. Fixed Point Theory Appl. **19**, 2241–2270 (2017). https://doi.org/10.1007/s11784-017-0417-1
13. Giselsson, P.: Nonlinear forward–backward splitting with projection correction (2019). arXiv:1908.07449
14. Giselsson, P.: Lecture notes on large-scale convex optimization (2015). http://control.lth.se/education/doctorate-program/large-scale-convex-optimization/
15. Giselsson, P., Boyd, S.: Linear convergence and metric selection for Douglas–Rachford splitting and ADMM. IEEE Trans. Autom. Control **62**(2), 532–544 (2017)
16. Guo, K., Han, D.: A note on the Douglas–Rachford splitting method for optimization problems involving hypoconvex functions. J. Glob. Optim. **72**(3), 431–441 (2018)
17. Guo, K., Han, D., Yuan, X.: Convergence analysis of Douglas–Rachford splitting method for strongly + weakly convex programming. SIAM J. Numer. Anal. **55**(4), 1549–1577 (2017)
18. Latafat, P., Patrinos, P.: Asymmetric forward–backward-adjoint splitting for solving monotone inclusions involving three operators. Comput. Optim. Appl. **68**(1), 57–93 (2017)
19. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation I: Basic Theory, vol. 330. Springer, Berlin (2006)
20. Mordukhovich, B.S.: Variational Analysis and Applications. Springer Monographs in Mathematics (2018)
21. Ogura, N., Yamada, I.: Non-strictly convex minimization over the fixed point set of an asymptotically shrinking nonexpansive mapping. Numer. Funct. Anal. Optim. **22**(1–2), 113–137 (2002)
22. Reich, S.: On the asymptotic behavior of nonlinear semigroups and the range of accretive operators. J. Math. Anal. Appl. **79**(1), 113–126 (1981)
23. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis, 3rd edn. Springer, Berlin (1998)
24. Ryu, E.K., Hannah, W., Yin, R.: Scaled relative graph: nonexpansive operators via 2D Euclidean geometry (2019). arXiv:1902.09788
25. Ryu, E.K., Taylor, A.B., Bergeling, C., Giselsson, P.: Operator splitting performance estimation: tight contraction factors and optimal parameter selection. SIAM J. Optim. **30**(3), 2251–2271 (2020)
26. Wang, X.: On Chebyshev functions and Klee functions. J. Math. Anal. Appl. **368**(1), 293–310 (2010)
27. Wen, B., Chen, X., Pong, T.K.: Linear convergence of proximal gradient algorithm with extrapolation for a class of nonconvex nonsmooth minimization problems. SIAM J. Optim. **27**(1), 124–145 (2017)