Canad. J. Math. 2025, pp. 1–26
http://dx.doi.org/10.4153/S0008414X25101284

On Bauschke–Bendit–Moursi modulus of averagedness and classifications of averaged nonexpansive operators

Shuang Song and Xianfu Wang

Abstract. Averaged operators are important in Convex Analysis and Optimization Algorithms. In this article, we propose classifications of averaged operators, firmly nonexpansive operators, and proximal operators using the Bauschke–Bendit–Moursi modulus of averagedness. We show that if an operator is averaged with a constant less than 1/2, then it is a bi-Lipschitz homeomorphism. Amazingly the proximal operator of a convex function has its modulus of averagedness less than 1/2 if and only if the function is Lipschitz smooth. Some results on the averagedness of operator compositions are obtained. Explicit formulae for calculating the modulus of averagedness of resolvents and proximal operators in terms of various values associated with the maximally monotone operator or subdifferential are also given. Examples are provided to illustrate our results.

1 Introduction

Throughout, we assume that

$$X \text{ is a real Hilbert space with inner product } \langle\cdot,\cdot\rangle : X \times X \to \mathbb{R},$$

and induced norm $\|\cdot\|$. Let Id denote the identity operator on X. Recall the following well-known definitions [6, 13].

Definition 1.1 Let $T : X \to X$ and $\mu > 0$. Then, T is

(i) nonexpansive[1] if

$$(\forall x \in X)(\forall y \in X) \quad \|Tx - Ty\| \leqslant \|x - y\|;$$

Received by the editors October 28, 2024; revised June 24, 2025; accepted July 30, 2025.
Published online on Cambridge Core August 7, 2025.
S. Song was partially supported by the UBC Graduate Research Scholarships and the Mitacs Globalink Graduate Fellowship.
X. Wang was partially supported by the Natural Sciences and Engineering Research Council of Canada.
AMS subject classification: **47H09**, **47H05**, 47A05, 52A41, 49J53.
Keywords: Averaged operator, Bauschke–Bendit–Moursi modulus of averagedness, firmly nonexpansive operator, maximally monotone operator, proximal operator.

[1]For convenience, we shall assume that T has a full domain throughout the article while one can generalize it to be on a proper subset of X.

(ii) firmly nonexpansive if

$$(\forall x \in X)(\forall y \in X) \quad \|Tx - Ty\|^2 + \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 \leqslant \|x - y\|^2;$$

(iii) μ-cocoercive if μT is firmly nonexpansive.

Definition 1.2 Let $T : X \to X$ be nonexpansive. T is k-averaged[2] if T can be represented as

$$T = (1 - k)\mathrm{Id} + kN,$$

where $N : X \to X$ is nonexpansive, and $k \in [0, 1]$.

Averaged operators are important in optimization (see, e.g., [1, 3, 5, 6, 9, 10, 13–15, 20, 24]). Firmly nonexpansive operators, being 1/2-averaged [6, Proposition 4.4], form a proper subclass of the class of averaged operators. From the definition, we have Id is the only 0-averaged operator. When $k \in (0, 1]$, various characterizations of k-averagedness (see [3, Proposition 2.2], [6, 13]) are available, including

$$(1.1)\quad (\forall x \in X)(\forall y \in X) \quad \|Tx - Ty\|^2 \leqslant \|x - y\|^2 - \frac{1-k}{k}\|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2,$$

and $(\forall x \in X)(\forall y \in X)$

$$(1.2)\quad \|Tx - Ty\|^2 \leqslant \langle x - y, Tx - Ty\rangle + (1 - 2k)(\langle x - y, Tx - Ty\rangle - \|x - y\|^2).$$

When $k = 0$, while the historic Definition 1.2 gives $T = \mathrm{Id}$ (linear), characterization (1.2) gives $T = \mathrm{Id} + v$ (affine) for some $v \in X$, hence they are not equivalent in this case. From (1.1) or (1.2) and the fact that Id is the only 0-averaged operator, we can deduce that if an operator is k_0-averaged, then it is k-averaged for every $k \geqslant k_0$. This motivates the following definition, which was proposed by Bauschke, Bendit, and Moursi [5].

Definition 1.3 (Bauschke–Bendit–Moursi modulus of averagedness) Let $T : X \to X$ be nonexpansive. The Bauschke–Bendit–Moursi modulus of averagedness of T is defined by

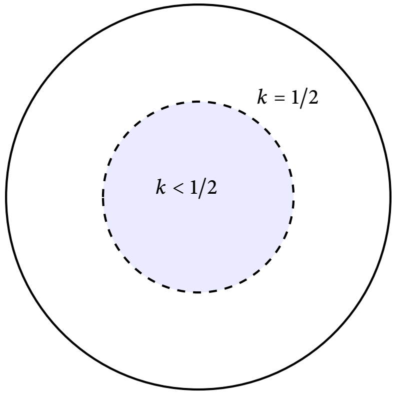
$$k(T) := \inf\{k \in [0, 1] \mid T \text{ is } k\text{-averaged}\}.$$

We call it the BBM modulus of averagedness.

It is natural to ask: How does the modulus of averagedness impact classifications of averaged operators? In view of Definition 1.3, if $T : X \to X$ is firmly nonexpansive then $k(T) \leqslant 1/2$. Based on this, we define the following, which classifies the class of firmly nonexpansive operators using the modulus of averagedness.

Definition 1.4 (Normal and special nonexpansiveness) Let $T : X \to X$. We say that T is normally (firmly) nonexpansive if $k(T) < 1/2$, and T is specially (firmly) nonexpansive if $k(T) = 1/2$.

[2] Usually, one excludes the cases $k = 0$ and $k = 1$ in the study of averaged operators, but it is very convenient in this article to allow these cases.



Let $\Gamma_0(X)$ denote the set of all proper lower semicontinuous convex functions from X to $(-\infty, +\infty]$. Recall that for $f \in \Gamma_0(X)$, its proximal operator is defined by $(\forall x \in X)\ \mathrm{P}_f(x) := \operatorname*{argmin}_{u \in X} \left\{ f(u) + \frac{1}{2}\|u - x\|^2 \right\}$. For a nonempty closed convex subset C of X, its indicator function is defined by $\iota_C(x) := 0$ if $x \in C$, and $+\infty$ otherwise. If $f = \iota_C$, we write $\mathrm{P}_f = P_C$, the projection operator onto C. It is well known that P_f is firmly nonexpansive [6], which implies $k(\mathrm{P}_f) \leqslant 1/2$. Some natural questions arise: Given $f \in \Gamma_0(X)$, when is P_f normally (or specially) nonexpansive? how can we calculate $k(\mathrm{P}_f)$? In [5], these problems are essentially solved in linear cases, or, in smooth case on the real line.

The goal of this article is to classify averaged nonexpansive operators, including firmly nonexpansive operators, via the Bauschke–Bendit–Moursi modulus of averagedness in a general Hilbert space. We provide some fundamental properties of modulus of averagedness of averaged mappings, firmly nonexpansive mappings and proximal mappings. We determine what properties normally (or specially) nonexpansive operators possess by using the monotone operator theory. One amazing result is that a proximal mapping of a convex function has its modulus of averagedness less than $1/2$ *if and only if the function is Lipschitz smooth. Many examples are provided to illustrate our results. Bauschke–Bendit–Moursi modulus of averagedness turns out to be an extremely powerful tool in studying averaged operators and firmly nonexpansive operators!*

The rest of the article is organized as follows. In Section 2, we explore some basic properties of the modulus function and show that a normally nonexpansive operator is a bi-Lipschitz homeomorphism. In Section 3, averagedness of operator compositions and some asymptotic behaviors of averaged operators are examined. In particular, the limiting operator of an averaged operator is a projection if and only if its BBM modulus is $1/2$. In Sections 4 and 5, we investigate both normal and special nonexpansiveness of resolvents and proximal operators. Our surprising results are Theorem 4.17 and Theorem 5.3, characterizing normal and special resolvents and proximal operators. In Section 6, we establish formulae of modulus of averagedness of resolvents in terms of various values of maximally monotone operators. Finally, in Section 7, we extend a modulus of averagedness formula on a composition of two projections by Bauschke, Bendit, and Moursi in $\mathbb{R}^2$ to general Hilbert spaces.

2 Bijective theorem

2.1 Auxiliary results

This section collects preparatory results on modulus of averagedness used in later proofs. For any operator $T : X \to X$ and any $v \in X$, the operator $T + v$ is defined by

$$(\forall x \in X) \quad (T + v)x := Tx + v.$$

Proposition 2.1 *Let $T : X \to X$ be nonexpansive and $v \in X$. Then,*
(i) $k(T + v) = k(T)$.
(ii) $k(T(\cdot + v)) = k(T)$.

Proof (i): The result follows by combining $(T + v)x - (T + v)y = Tx - Ty$ with characterization (1.2).

(ii): The result follows by combining $x - y = (x + v) - (y + v)$ with characterization (1.2). ∎

Proposition 2.2 *Let $T : X \to X$ be nonexpansive. If $k(T) > 0$, then T is $k(T)$-averaged. Moreover, T is β-averaged for every $\beta \in [k(T), 1]$.*

Proof Due to $k(T) > 0$, we can use characterization either (1.1) or (1.2). The right hand side of (1.1) or (1.2) is a continuous and increasing function in term of k, thus the result follows. ∎

Let $\mathrm{Fix}T := \{x \in X \mid Tx = x\}$ denote the set of fixed points of $T : X \to X$. Our following result characterizes $k(T) = 0$.

Proposition 2.3 *Let $T : X \to X$ be nonexpansive. Then,*

$$k(T) = 0 \Leftrightarrow \exists v \in X : T = \mathrm{Id} + v. \tag{2.1}$$

If, in addition, $\mathrm{Fix}T \neq \varnothing$, *then*

$$k(T) = 0 \Leftrightarrow T = \mathrm{Id}. \tag{2.2}$$

Proof Suppose $\exists v \in X : T = \mathrm{Id} + v$. Obviously $k(\mathrm{Id}) = 0$. Thus, by Proposition 2.1, $k(T) = k(\mathrm{Id} + v) = 0$.

Suppose $k(T) = 0$. Assume that for any $v \in X$: $T \neq \mathrm{Id} + v$. Then, there exist $x_0, y_0 \in X$ such that $(T - \mathrm{Id})x_0 \neq (T - \mathrm{Id})y_0$, whence $\|(T - \mathrm{Id})x_0 - (T - \mathrm{Id})y_0\|^2 > 0$. Our assumption implies $T \neq \mathrm{Id}$, and Id is the only 0-averaged operator, thus there exists a sequence $(k_n)_{n\in\mathbb{N}}$ in $(0, 1]$ such that T is k_n-averaged and $k_n \to 0$. Now characterization (1.1) implies that for any $n \in \mathbb{N}$:

$$\|Tx_0 - Ty_0\|^2 \leqslant \|x_0 - y_0\|^2 - \frac{1 - k_n}{k_n}\|(\mathrm{Id} - T)x_0 - (\mathrm{Id} - T)y_0\|^2,$$

i.e.,

$$0 \leqslant \|x_0 - y_0\|^2 - \|Tx_0 - Ty_0\|^2 + \left(1 - \frac{1}{k_n}\right)\|(T - \mathrm{Id})x_0 - (T - \mathrm{Id})y_0\|^2.$$

Note that $\|(T - \mathrm{Id})x_0 - (T - \mathrm{Id})y_0\|^2 > 0$. Now letting $n \to \infty$ yields $0 \leqslant -\infty$, which is a contradiction.

When $\mathrm{Fix}T \neq \varnothing$, (2.2) follows from (2.1). ∎

Proposition 2.4 *Let $T : X \to X$ be nonexpansive. Then, T is firmly nonexpansive if and only if $k(T) \leqslant 1/2$.*

Proof "$\Leftarrow$": When $0 < k(T) < 1/2$, apply Proposition 2.2. When $k(T) = 0$, apply Proposition 2.3.

"$\Rightarrow$": The assumption implies that T is 1/2-averaged. Hence, $k(T) \leqslant 1/2$. ■

Example 2.5 If $T : X \to X$ is a constant mapping, i.e., $(\exists v \in X)(\forall x \in X)\ Tx = v$, then $k(T) = 1/2$.

Proof Because T is firmly nonexpansive, $k(T) \leqslant 1/2$. By (1.2), if T is k-averaged, then $2k \geqslant 1$, so $k(T) \geqslant 1/2$. Altogether, $k(T) = 1/2$. ■

We end up this section with a fact on convexity.

Fact 2.6 [5, Fact 1.3] *Let $T_1, T_2 : X \to X$ be nonexpansive and $\lambda \in [0,1]$. Then, $k(\lambda T_1 + (1-\lambda)T_2) \leqslant \lambda k(T_1) + (1-\lambda)k(T_2)$. Consequently, $T \mapsto k(T)$ is a convex function on the set of averaged mappings, as well as on the set of firmly nonexpansive mappings.*

Corollary 2.7 *Let $T : X \to X$ be nonexpansive and $\lambda \in [0,1]$. Then, $k(\lambda T) \leqslant \lambda k(T) + (1-\lambda)/2$.*

Proof Let T_2 be zero mapping in Fact 2.6 and apply Example 2.5. ■

2.2 Bijective theorem

In this section, we will show that normally nonexpansive operator must be bijective and bi-Lipschitz. First, we prove that normally nonexpansive operators must be bi-Lipschitz and injective.

Lemma 2.8 *Let $T : X \to X$ be normally nonexpansive. Then, T is a bi-Lipschitz homeomorphism from X to* ran T. *In particular, T is injective.*

Proof In view of Proposition 2.3, we may assume $k(T) > 0$. Then, T is $k(T)$-averaged by Proposition 2.2, i.e.,

$$(\forall x \in X)(\forall y \in X) \quad \|Tx - Ty\|^2 + (1 - 2k(T))\|x - y\|^2 \leqslant 2(1 - k(T))\langle x - y, Tx - Ty\rangle.$$

Since $k(T) < \frac{1}{2}$, there exists $\alpha \in (0, \frac{1}{2})$ such that $k(T) = \frac{1}{2} - \alpha$. Substituting $k(T)$ in above inequality and using the Cauchy–Schwarz inequality, we have

$$\|Tx - Ty\|^2 + 2\alpha\|x - y\|^2 \leqslant (1 + 2\alpha)\langle x - y, Tx - Ty\rangle,$$
$$\|Tx - Ty\|^2 + 2\alpha\|x - y\|^2 \leqslant (1 + 2\alpha)\left(\|x - y\|\,\|Tx - Ty\|\right),$$
$$2\alpha\left(\|x - y\|^2 - \|x - y\|\,\|Tx - Ty\|\right) \leqslant \|x - y\|\,\|Tx - Ty\| - \|Tx - Ty\|^2,$$
$$2\alpha\|x - y\|\left(\|x - y\| - \|Tx - Ty\|\right) \leqslant \|Tx - Ty\|\left(\|x - y\| - \|Tx - Ty\|\right).$$

Now if $\|x - y\| - \|Tx - Ty\| = 0$, then $\|Tx - Ty\| = \|x - y\| \geqslant 2\alpha\|x - y\|$ since $2\alpha < 1$. If $\|x - y\| - \|Tx - Ty\| \neq 0$, then $2\alpha\|x - y\| \leqslant \|Tx - Ty\|$. Thus in both cases we have $2\alpha\|x - y\| \leqslant \|Tx - Ty\|$. Combining it with $\|Tx - Ty\| \leqslant \|x - y\|$, we have

$$2\alpha\|x - y\| \leqslant \|Tx - Ty\| \leqslant \|x - y\|.$$

i.e., T is a bi-Lipschitz homeomorphism from X to ran T. ■

Next, we make use of monotone operator theory to prove that normally nonexpansive operators must also be surjective.

Fact 2.9 [6, Example 20.30] *Let $T : X \to X$ be firmly nonexpansive. Then, T is maximally monotone.*

Fact 2.10 ((Rockafellar–Vesely) [6, Corollary 21.24]) *Let $A : X \rightrightarrows X$ be a maximally monotone operator such that $\lim_{\|x\| \to +\infty} \inf \|Ax\| = +\infty$. Then, A is surjective.*

Lemma 2.11 *Let $T : X \to X$ be normally nonexpansive. Then, T is surjective.*

Proof By Lemma 2.8, T is bi-Lipschitz since T is normally nonexpansive. Thus, there exists $\varepsilon > 0$, such that $\varepsilon\|x - y\| \leqslant \|Tx - Ty\|$. Let $y = 0$, then $\varepsilon\|x\| \leqslant \|Tx - T0\|$. Using the triangle inequality, we have

$$\|Tx\| \geqslant \varepsilon\|x\| - \|T0\|.$$

Thus, $\lim_{\|x\| \to \infty} \|Tx\| = \infty$. Combining Fact 2.9 with Fact 2.10 we complete the proof. ■

Theorem 2.12 (bi-Lipschitz homeomorphism) *Let $T : X \to X$ be normally nonexpansive. Then, T is a bi-Lipschitz homeomorphism of X. In particular, T is bijective.*

Proof Combine Lemmas 2.8 and 2.11. ■

Taking the contrapositive of Theorem 2.12, we obtain a lower bound for modulus of averagedness.

Corollary 2.13 *Let $T : X \to X$ be nonexpansive. If T is not bijective, then $k(T) \geqslant 1/2$.*

Remark 2.14 In terms of compact operators (see, e.g., [23]), Theorem 2.12 implies that X is finite-dimensional if and only if there exists a normally nonexpansive compact operator on X.

Example 2.15 (Averagedness of projection) Let C be a nonempty closed convex set in X and $C \neq X$. Then, P_C is specially nonexpansive.

Proof We have $k(P_C) \leqslant 1/2$ since P_C is firmly nonexpansive. Now since $C \neq X$, let $x_0 \in X \backslash C$. Because $P_C(x_0) \in C$ and $x_0 \in X \backslash C$, we have $P_C(x_0) \neq x_0$. However, $P_C(x_0) = P_C(P_C(x_0))$. Thus, P_C is not injective. Therefore, P_C is specially nonexpansive by Corollary 2.13. Another way is to observe that P_C is not surjective. ■

Corollary 2.16 *Let $M \in \mathbb{R}^{n \times n}$ be nonexpansive. If* $\det(M) = 0$*, then* $k(M) \geqslant 1/2$.

Remark 2.17 Consider the matrix

$$A = \frac{1}{2}\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}.$$

Then, one can verify that $k(A) = 3/4 > 1/2$. However, $\det(A) \neq 0$ and thus A is a bi-Lipschitz homeomorphism of $\mathbb{R}^2$. Hence, the converse of Theorem 2.12 fails. We will show later that the converse of Theorem 2.12 does hold when T is a proximal operator (see Theorem 5.3).

3 Operator compositions and limiting operator

In this section, we examine the modulus of averagedness of operator compositions and explore its asymptotic properties.

3.1 Composition

Proposition 3.1 *Let T_1 and T_2 be nonexpansive operators from X to X. Suppose one of the following holds:*

(i) *T_1 is not surjective.*
(ii) *T_2 is not injective.*
(iii) *T_1 is bijective and T_2 is not surjective.*
(iv) *T_2 is bijective and T_1 is not injective.*

Then, $k(T_1T_2) \geqslant 1/2$.

Proof Since T_1 and T_2 are nonexpansive operators, we have T_1T_2 is nonexpansive as well. Each one of the four conditions implies that T_1T_2 is not bijective. Now, use Corollary 2.13. ■

Ogura and Yamada [20] obtained the following result about the averagedness of operator compositions.

Fact 3.2 ([20, Theorem 3] (see also [15, Proposition 2.4])) *Let $T_1 : X \to X$ be α_1-averaged, and let $T_2 : X \to X$ be α_2-averaged, where $\alpha_1, \alpha_2 \in (0,1)$. Set*

$$T = T_1T_2 \quad \text{and} \quad \alpha = \frac{\alpha_1 + \alpha_2 - 2\alpha_1\alpha_2}{1 - \alpha_1\alpha_2}.$$

Then, $\alpha \in (0,1)$ and T is α-averaged.

Formulating this result here using the modulus of averagedness, we have the following result.

Proposition 3.3 *Let $T_1 : X \to X$ and $T_2 : X \to X$ be nonexpansive. Suppose $k(T_1)k(T_2) \neq 1$. Then,*

$$k(T_1T_2) \leqslant \frac{k(T_1) + k(T_2) - 2k(T_1)k(T_2)}{1 - k(T_1)k(T_2)}.$$

Proof Let $\varphi(T_1, T_2) := \frac{k(T_1)+k(T_2)-2k(T_1)k(T_2)}{1-k(T_1)k(T_2)}$. We consider five cases.

Case 1: $k(T_i) = 1$ for some $i \in \{1,2\}$. Then, $\varphi(T_1, T_2) = 1$. Since T_1 and T_2 are nonexpansive, we have T_1T_2 is nonexpansive, i.e., $k(T_1T_2) \leqslant 1 = \varphi(T_1, T_2)$.

Case 2: $k(T_i) \in (0,1)$ for any $i \in \{1,2\}$. Then, combining Proposition 2.2 and Fact 3.2, we have T_1T_2 is $\varphi(T_1, T_2)$-averaged. Thus, $k(T_1T_2) \leqslant \varphi(T_1, T_2)$.

Case 3: $k(T_1) = 0$ and $k(T_2) \in (0,1)$. Then, there exists $v_1 \in X$ such that $T_1 = \mathrm{Id} + v_1$ by Proposition 2.3. Thus, $T_1T_2 = T_2 + v_1$ and $k(T_1T_2) = k(T_2 + v_1) = k(T_2)$ by Proposition 2.1. While $\varphi(T_1, T_2) = k(T_2)$ in this case, we have $k(T_1T_2) = \varphi(T_1, T_2)$.

Case 4: $k(T_1) \in (0,1)$ and $k(T_2) = 0$. Then, there exists $v_2 \in X$ such that $T_2 = \mathrm{Id} + v_2$ by Proposition 2.3. Thus, $T_1T_2 = T_1(\cdot + v_2)$ and $k(T_1T_2) = k(T_1(\cdot + v_2)) = k(T_1)$ by Proposition 2.1. While $\varphi(T_1, T_2) = k(T_1)$ in this case, we have $k(T_1T_2) = \varphi(T_1, T_2)$.

Case 5: $k(T_1) = k(T_2) = 0$. Then, there exist $v_1 \in X$ and $v_2 \in X$ such that $T_1 = \mathrm{Id} + v_1$ and $T_2 = \mathrm{Id} + v_2$ by Proposition 2.3. Thus, $T_1T_2 = \mathrm{Id} + v_2 + v_1$ and $k(T_1T_2) = k(\mathrm{Id} + v_2 + v_1) = 0$. While $\varphi(T_1, T_2) = 0$ in this case, we have $k(T_1T_2) = \varphi(T_1, T_2)$.

Altogether, we complete the proof. ■

Proposition 3.4 *Let C be a nonempty closed convex set in X and $C \neq X$. Then, for any nonexpansive operator $T : X \to X$:*

$$\frac{1}{2} \leqslant k(T \circ P_C) \leqslant \frac{1}{2 - k(T)}$$

and

$$\frac{1}{2} \leqslant k(P_C \circ T) \leqslant \frac{1}{2 - k(T)}.$$

Proof Observe that P_C is neither surjective nor injective in this case. Thus, by Proposition 3.1, we have $k(T \circ P_C) \geqslant 1/2$ and $k(P_C \circ T) \geqslant 1/2$. Now by Example 2.15,

$$\frac{k(T) + k(P_C) - 2k(T)k(P_C)}{1 - k(T)k(P_C)} = \frac{1}{2 - k(T)}.$$

Thus, by Proposition 3.3, we have $k(T \circ P_C) \leqslant \frac{1}{2-k(T)}$ and $k(P_C \circ T) \leqslant \frac{1}{2-k(T)}$, which complete the proof. ■

Remark 3.5 Particularly, if we let $T = P_V$ and $C = U$, where U and V are both closed linear subspaces, then $k(P_V P_U) = \frac{1+c_F}{2+c_F} \in \left[\frac{1}{2}, \frac{2}{3}\right]$, where $c_F \in [0,1]$ (see [5, Corollary 3.3]). This coincides with the bounds we obtained as $\frac{1}{2-k(P_U)} = \frac{2}{3}$ by Example 2.15.

We can generalize the results of two operator compositions to finite operator compositions.

Proposition 3.6 *Let $m \geqslant 2$ be an integer and let $I = \{1, \ldots, m\}$. For any $i \in I$, let T_i be nonexpansive from X to X. Suppose one of the following holds:*

(i) *T_1 is not surjective.*
(ii) *T_m is not injective.*
(iii) *T_1 is bijective and $T_2 \cdots T_m$ is not surjective.*
(iv) *T_m is bijective and $T_1 \cdots T_{m-1}$ is not injective.*

Then, $k(T_1 \cdots T_m) \geqslant 1/2$.

Proof Apply Proposition 3.1. ■

Corollary 3.7 *Let $C_1, \ldots, C_m$ be nonempty closed convex sets in X. If $C_1 \neq X$ or $C_m \neq X$, then $k(P_{C_1} \cdots P_{C_m}) \geqslant 1/2$.*

The following result is about modulus of averagedness of isometries.

Proposition 3.8 *Let A be a $n \times n$ orthogonal matrix and $A \neq \mathrm{Id}$. Then, $k(A) = 1$.*

Proof Since A is orthogonal, we have $\|Ax - Ay\| = \|x - y\|$. On the other hand, $\mathrm{ran}(\mathrm{Id} - A)$ is not a singleton. Hence, $k(A) = 1$ by using (1.1). ■

Corollary 3.9 *Let $m \geqslant 1$ be an integer and let $I = \{1, \ldots, m\}$. For any $i \in I$, let A_i be a $n \times n$ orthogonal matrix. Suppose that $A_1 \cdots A_m \neq \mathrm{Id}$. Then, $k(A_1 \cdots A_m) = 1$.*

3.2 Limiting operator

In this section, we discuss the asymptotic behavior of modulus of averagedness. Recall that a sequence $(x_n)_{n\in\mathbb{N}}$ in a Hilbert space X is said to converge weakly to a point x in X if $(\forall y \in X)\ \lim_{n\to\infty} \langle x_n, y\rangle = \langle x, y\rangle$. We use the notation $\lim^{w}_{n\to\infty} x_n$ for the weak limit of $(x_n)_{n\in\mathbb{N}}$. Recall for a nonexpansive operator $T : X \to X$, $\operatorname{Fix} T$ is closed and convex (see, e.g., [9, Proposition 22.9]).

Fact 3.10 [6, Proposition 5.16] *Let $\alpha \in (0,1)$ and let $T : X \to X$ be α-averaged such that* $\operatorname{Fix} T \neq \varnothing$. *Then, for any $x \in X$, $(T^n x)_{n\in\mathbb{N}}$ converges weakly to a point in* $\operatorname{Fix} T$.

In view of the above fact, we propose the following type of operator.

Definition 3.1 (Limiting operator) Let $\alpha \in (0,1)$ and let $T : X \to X$ be α-averaged such that $\operatorname{Fix} T \neq \varnothing$. Define its limiting operator $T_\infty : X \to X$ by $x \mapsto \lim^{w}_{n\to\infty} T^n x$.

Remark 3.11 The full domain and single-valuedness of T_∞ are guaranteed by Fact 3.10. Hence, $T_\infty : X \to X$ is well defined.

Example 3.12

(i) [6, Example 5.29] Let $\alpha \in (0,1)$ and let $T : X \to X$ be α-averaged such that $\operatorname{Fix} T \neq \varnothing$. Suppose T is linear. Then, $T_\infty = P_{\operatorname{Fix} T}$.

(ii) [6, Proposition 5.9] Let $\alpha \in (0,1)$ and let $T : X \to X$ be α-averaged such that $\operatorname{Fix} T \neq \varnothing$. Suppose $\operatorname{Fix} T$ is a closed affine subspace of X. Then, $T_\infty = P_{\operatorname{Fix} T}$.

The limiting operator of an averaged mapping enjoys the following pleasing properties.

Proposition 3.13 *Let $\alpha \in (0,1)$ and let $T : X \to X$ be α-averaged such that* $\operatorname{Fix} T \neq \varnothing, X$. *Then, the following hold:*

(i) $\operatorname{Fix} T = \operatorname{Fix} T_\infty = \operatorname{ran} T_\infty$.

(ii) $(T_\infty)^2 = T_\infty$.

(iii) $k(T_\infty) \in \left[\frac{1}{2}, 1\right]$.

Proof (i): If $x \notin \operatorname{Fix} T$, then $T_\infty x \neq x$ since $T_\infty x \in \operatorname{Fix} T$ by Fact 3.10. If $x \in \operatorname{Fix} T$, then $T_\infty x = \lim^{w}_{n\to\infty} T^n x = \lim^{w}_{n\to\infty} x = x$. Thus, $\operatorname{Fix} T = \operatorname{Fix} T_\infty$. The equality $\operatorname{Fix} T = \operatorname{ran} T_\infty$ follows by using Fact 3.10 again.

(ii): For any $x \in X$, $T_\infty x \in \operatorname{ran} T_\infty$, thus $T_\infty x \in \operatorname{Fix} T_\infty$ by (i). Therefore, $(T_\infty)^2 x = T_\infty (T_\infty x) = T_\infty x$, which implies that $(T_\infty)^2 = T_\infty$.

(iii): Since the norm is weakly lower-semicontinuous, we have

$$(\forall x \in X)(\forall y \in X) \quad \|T_\infty x - T_\infty y\| \leqslant \liminf_{n\to\infty} \|T^n x - T^n y\|.$$

As $T : X \to X$ is nonexpansive, by induction we have for any $n \in \mathbb{N}$, $\|T^n x - T^n y\| \leqslant \|x - y\|$. Altogether, T_∞ is nonexpansive, which implies that $k(T_\infty) \leqslant 1$. On the other hand, we have $\operatorname{ran} T_\infty = \operatorname{Fix} T \neq X$ by (i) and the assumption, so T_∞ is not surjective. Thus, $k(T_\infty) \geqslant 1/2$ by Corollary 2.13. ■

The modulus of averagedness provides further insights into the limiting operator.

Theorem 3.14 *Let $\alpha \in (0,1)$ and let $T : X \to X$ be α-averaged such that* $\operatorname{Fix} T \neq \varnothing, X$. *Then, the following are equivalent:*

(i) $T_\infty = P_{\operatorname{Fix} T}$.
(ii) $k(T_\infty) \leqslant 1/2$.
(iii) $k(T_\infty) = 1/2$.

Proof (i) $\Rightarrow$ (ii): Obvious.

(ii) $\Rightarrow$ (i): The result follows by combining Proposition 3.13(i)&(ii) and the fact that if $T : X \to X$ is firmly nonexpansive and $T \circ T = T$, then $T = P_{\operatorname{ran} T}$ (see [9, Exercise 22.5] or [8, Theorem 2.1(xx)]).

(ii) $\Leftrightarrow$ (iii): Apply Proposition 3.13(iii). ∎

In the following, we discuss limiting operator on $\mathbb{R}$. The following extends [5, Proposition 2.8] from differentiable functions to locally Lipschitz functions. Below $\partial_L g$ denotes the Mordukhovich limiting subdifferential [12, 19, 22].

Lemma 3.15 *Let $g : \mathbb{R} \to \mathbb{R}$ be a locally Lipschitz function. Then, g is nonexpansive if and only if $(\forall x \in \mathbb{R})\ \partial_L g(x) \subset [-1, 1]$ in which case $k(g) = (1 - \inf \partial_L g(\mathbb{R}))/2$.*

Proof The nonexpansiveness characterization of g follows from [12, Theorem 3.4.8]. Write $g = (1-\alpha)\mathrm{Id} + \alpha N$, where $\alpha \in [0, 1]$ and $N : \mathbb{R} \to \mathbb{R}$ is nonexpansive. If $\alpha = 0$, the result clearly holds. Let us assume $\alpha > 0$. Then, $N(x) = (g(x) - (1-\alpha)x)/\alpha$ and $\partial_L N(x) = (\partial_L g(x) - (1-\alpha))/\alpha$. N is nonexpansive is equivalent to

$$(\forall x \in \mathbb{R})\ (\partial_L g(x) - (1-\alpha))/\alpha \subseteq [-1, 1] \quad \Leftrightarrow \quad (\forall x \in \mathbb{R})\ \partial_L g(x) \subseteq [1 - 2\alpha, 1],$$

from which

$$\alpha \geqslant \frac{1 - \inf \partial_L g(\mathbb{R})}{2}$$

and the result follows. ∎

In Example 3.12 we see that if $T : X \to X$ is α-averaged and linear with $\operatorname{Fix} T \neq \varnothing, X$, then $k(T_\infty) = 1/2$. The following example shows that it is not true in nonlinear case.

Example 3.16 Let

$$f(x) := \begin{cases} 0 & \text{if } x \leqslant 0, \\ x & \text{if } 0 \leqslant x \leqslant 1, \\ -\frac{1}{2}x + \frac{3}{2} & \text{if } x \geqslant 1. \end{cases}$$

Then, f is $(3/4)$-averaged and $\operatorname{Fix} T = [0, 1]$. However,

$$f_\infty(x) = \begin{cases} 0 & \text{if } x \leqslant 0 \text{ or } x \geqslant 3, \\ x & \text{if } 0 \leqslant x \leqslant 1, \\ -\frac{1}{2}x + \frac{3}{2} & \text{if } 1 \leqslant x \leqslant 3, \end{cases}$$

and $k(f_\infty) = 3/4$.

Proof By computation, we have

$$\partial_L f(x) = \begin{cases} \{0\} & \text{if } x < 0, \\ [0,1] & \text{if } x = 0, \\ \{1\} & \text{if } 0 < x < 1, \\ \{-1/2, 1\} & \text{if } x = 1, \\ \{-1/2\} & \text{if } x > 1, \end{cases} \text{ and } \partial_L f_\infty(x) = \begin{cases} \{0\} & \text{if } x < 0 \text{ or } x > 3, \\ [0,1] & \text{if } x = 0, \\ \{1\} & \text{if } 0 < x < 1, \\ \{-1/2, 1\} & \text{if } x = 1, \\ \{-1/2\} & \text{if } 1 < x < 3, \\ [-1/2, 0] & \text{if } x = 3. \end{cases}$$

Applying Lemma 3.15, we obtain $k(f)$ and $k(f_\infty)$. ■

Next, we show that if $T : X \to X$ is firmly nonexpansive, a stronger condition than averagedness, then on the real line it is true that $k(T_\infty) = 1/2$.

Proposition 3.17 *Let $f : \mathbb{R} \to \mathbb{R}$ be firmly nonexpansive such that* $\operatorname{Fix} f \neq \varnothing, \mathbb{R}$. *Then, $f_\infty = P_{\operatorname{Fix} f}$. Consequently, $k(f_\infty) = 1/2$.*

Proof Since f is firmly nonexpansive, we have f is nondecreasing and nonexpansive. Now as $\operatorname{Fix} f \subseteq \mathbb{R}$ is closed and convex, it must be one of the form $[a, +\infty)$, $(-\infty, b]$ or $[a, b]$ with $a, b \in \mathbb{R}$ because $\operatorname{Fix} f \neq \varnothing, \mathbb{R}$. Since the proofs for all cases are similar, let us assume that $\operatorname{Fix} f = [a, b]$. When $x \geqslant b$, because f is nondecreasing, we have $f(x) \geqslant f(b) = b$, $f^2(x) \geqslant f(b) = b$, and an induction leads $f^n(x) \geqslant b$. Then, $f_\infty(x) \geqslant b$ by Fact 3.10. Since $f_\infty(x) \in [a, b]$ by Fact 3.10 again, we derive that $f_\infty(x) = b$. Similar arguments give $f_\infty(x) = a$ when $x \leqslant a$. Clearly, when $x \in [a, b]$, $(\forall n \in \mathbb{N})\ f^n(x) = x$, so $f_\infty(x) = x$. Altogether $f_\infty = P_{\operatorname{Fix} f}$. ■

Motivated by Example 3.12 and Proposition 3.17, one might conjecture that $k(T_\infty) = 1/2$ whenever $k(T) \leqslant 1/2$. However, this is not true in general. To find a counter example, by Theorem 3.14, it suffices to find a firmly nonexpansive operator such that its limiting operator is not a projection. We conclude this section with the following example from [7, Example 4.2].

Example 3.18 Suppose that $X = \mathbb{R}^2$. Let $A = \mathbb{R}(1,1)$ and $B = \{(x, y) \in \mathbb{R}^2 \mid -y \leqslant x \leqslant 2\}$. For $z = (x, y) \in \mathbb{R}^2$, we have $P_A(z) = \left(\frac{x+y}{2}, \frac{x+y}{2}\right)$ and

$$P_B(z) = \begin{cases} (2, y) & \text{if } x \geq 2, y \geq -2, \\ (2, -2) & \text{if } y \leq \min\{x-4, -2\}, \\ ((x-y)/2, -(x-y)/2) & \text{if } x - 4 < y \leq -x, \\ (x, y) & \text{if } (x, y) \in B. \end{cases}$$

Then, the *Douglas–Rachford* operator $T = \operatorname{Id} - P_A + P_B(2P_A - \operatorname{Id})$ is firmly nonexpansive and has $k(T_\infty) > 1/2$. By Theorem 3.14, it suffices to show $T_\infty \neq P_{\operatorname{Fix} T}$. Indeed, by [7, Fact 3.1] we have $\operatorname{Fix} T = \{s(1,1) \mid s \in [0,2]\}$ because of $A \cap \operatorname{int} B \neq \varnothing$. Let $z_0 = (4, 10)$, and $(\forall n \in \mathbb{N})\ z_{n+1} = Tz_n$. Direct computations give

$$z_1 = (-1, 7), z_2 = (-2, 3), z_3 = (-1/2, 1/2), \text{ and } z_4 = (0, 0).$$

On the other hand, let $z^* = (2, 2)$, then $\{z_4, z^*\} \subset \operatorname{Fix} T$. Thus, $T_\infty z_0 = z_4$ while $P_{\operatorname{Fix} T} z_0 \neq z_4$ as $\|z_0 - z^*\| = 2\sqrt{17} < \|z_0 - z_4\| = 2\sqrt{29}$.

4 Resolvent

Let $A : X \rightrightarrows X$ be a set-valued operator, i.e., a mapping from X to its power set. Recall that the resolvent of A is $J_A := (\mathrm{Id} + A)^{-1}$ and the reflected resolvent of A is $R_A := 2J_A - \mathrm{Id}$. The graph of A is $\operatorname{gra} A := \{(x,u) \in X \times X \mid u \in Ax\}$ and the inverse of A, denoted by A^{-1}, is the operator with graph $\operatorname{gra} A^{-1} := \{(u,x) \in X \times X \mid u \in Ax\}$. The domain of A is $\operatorname{dom} A := \{x \in X \mid Ax \neq \varnothing\}$. A is monotone, if

$$\forall (x,u), (y,v) \in \operatorname{gra} A, \quad \langle x - y, u - v \rangle \geqslant 0.$$

A is maximally monotone, if it is monotone and there is no monotone operator $B : X \rightrightarrows X$ such that $\operatorname{gra} A$ is properly contained in $\operatorname{gra} B$. Unless stated otherwise, we assume from now on that

$$A : X \rightrightarrows X \text{ and } B : X \rightrightarrows X \text{ are maximally monotone operators.}$$

Fact 4.1 ((Minty's theorem) [6, Proposition 23.8]) *Let $T : X \to X$. Then, T is firmly nonexpansive if and only if T is the resolvent of a maximally monotone operator.*

The goal of this section is to give characterizations of normal and special nonexpansiveness by using the monotone operator theory.

4.1 Auxiliary results

We first provide a nice formula for the modulus of averagedness of $(1-\lambda)\mathrm{Id} + \lambda T$ in terms of the modulus of averagedness of T. The following is an adaption of [6, Proposition 4.40]. For completeness, we include a simple proof.

Fact 4.2 *Let $T : X \to X$ be nonexpansive and let $\lambda \in (0,1]$. For $\alpha \in [0,1]$, T is α-averaged if and only if $(1-\lambda)\mathrm{Id} + \lambda T$ is $\lambda\alpha$-averaged.*

Proof Suppose T is α-averaged. Then, $T = (1-\alpha)\mathrm{Id} + \alpha R$ with R being nonexpansive. It follows that

$$(1-\lambda)\mathrm{Id} + \lambda T = (1-\lambda)\mathrm{Id} + \lambda(1-\alpha)\mathrm{Id} + \lambda\alpha R \tag{4.1}$$

$$= (1-\lambda\alpha)\mathrm{Id} + \lambda\alpha R, \tag{4.2}$$

so that $(1-\lambda)\mathrm{Id} + \lambda T$ is $\lambda\alpha$-averaged. Because $\lambda \in (0,1]$, the reverse direction also holds. ■

Lemma 4.3 *Let $T : X \to X$ be nonexpansive. Then, for every $\lambda \in [0,1]$, we have*

$$k((1-\lambda)\mathrm{Id} + \lambda T) = \lambda k(T). \tag{4.3}$$

Proof We split the proof into the following cases.

Case 1: $\lambda = 0$. Clearly, (4.3) holds because $k(\mathrm{Id}) = 0$.

Case 2: $\lambda > 0$. We show (4.3) by two subcases.

Case 2.1: $k((1-\lambda)\mathrm{Id} + \lambda T) = 0$. By Proposition 2.3, there exists $v \in X$: $(1-\lambda)\mathrm{Id} + \lambda T = \mathrm{Id} + v$ such that $T = \mathrm{Id} + v/\lambda$. Then, $k((1-\lambda)\mathrm{Id} + \lambda T) = 0 = k(T)$ by Proposition 2.3 again.

Case 2.2: $k((1-\lambda)\mathrm{Id}+\lambda T)>0$. On one hand, we derive $k((1-\lambda)\mathrm{Id}+\lambda T)\leqslant \lambda k(T)$ by Fact 4.2. On the other hand, since $(1-\lambda)\mathrm{Id}+\lambda T$ is λ-averaged, we have $0<k((1-\lambda)\mathrm{Id}+\lambda T)\leqslant\lambda$. For every $\beta\in[k((1-\lambda)\mathrm{Id}+\lambda T),\lambda]$, the mapping $(1-\lambda)\mathrm{Id}+\lambda T$ is β-averaged. Write $\beta=\lambda\alpha$ with $\alpha=\beta/\lambda\in(0,1]$. Fact 4.2 implies that T is α-averaged, thus $k(T)\leqslant\beta/\lambda$. Taking infimum over β gives $k(T)\leqslant k((1-\lambda)\mathrm{Id}+\lambda T)/\lambda$, i.e., $\lambda k(T)\leqslant k((1-\lambda)\mathrm{Id}+\lambda T)$. Therefore, $k((1-\lambda)\mathrm{Id}+\lambda T)=\lambda k(T)$.

Altogether, (4.3) holds. ■

Example 4.4 Let C be a nonempty closed convex set in X and $C\neq X$. Consider the reflector to C defined by $R_C:=2P_C-\mathrm{Id}$. Then, the following hold:

(i) $k(R_C)=1$.
(ii) For $\lambda\in[0,1]$, $k((1-\lambda)\mathrm{Id}+\lambda R_C))=\lambda$.
(iii) For $\lambda\in[0,1]$, $k((1-\lambda)\mathrm{Id}+\lambda P_C))=\lambda/2$.

Proof Apply Example 2.15 and Lemma 4.3. ■

Remark 4.5 This recovers [5, Example 2.3] for $C=V$, a closed subspace of X.

Example 4.6 Let $A: X\rightrightarrows X$ be maximally monotone. Consider the reflected resolvent of A defined by $R_A:=2J_A-\mathrm{Id}$. Then, $k(R_A)=2k(J_A)$ by Lemma 4.3. Consequently, $k(R_A)<1$ (that is, R_A is α-averaged for some $\alpha\in[0,1)$) if and only if J_A is normally nonexpansive. Likewise, $k(R_A)=1$ if and only if J_A is specially nonexpansive.

The following result concerning the Douglas–Rachford operator (see, e.g., [6, 9]) is of independent interest.

Theorem 4.7 *Let U, V be two closed subspaces of X, and $U\neq V$. Consider the Douglas–Rachford operator*

$$T_{U,V}:=\frac{\mathrm{Id}+R_UR_V}{2}.$$

Then, $k(T_{U,V})=1/2$.

Proof We have $R_UR_V\neq\mathrm{Id}$ since $U\neq V$. Note both R_U and R_V are orthogonal. Thus, $k(R_UR_V)=1$ by Corollary 3.9. Therefore, by Lemma 4.3, we have $k(T_{U,V})=k(R_UR_V)/2=1/2$. ■

Remark 4.8 Let $A,B: X\rightrightarrows X$ be two maximally monotone operators. The Douglas–Rachford operator related to (A,B) is

$$T_{A,B}=\frac{\mathrm{Id}+R_AR_B}{2}.$$

It is interesting to know $k(T_{A,B})$ in general.

Next, we recall Yosida regularizations of monotone operators. They are essential for our proofs in Section 4.2.

Definition 4.1 (Yosida regularization) For $\mu>0$, the Yosida μ-regularization of A is the operator

$$Y_\mu(A):=\left(\mu\mathrm{Id}+A^{-1}\right)^{-1}.$$

For Yosida regularization, we have the classic identity: $Y_\mu(A) = \mu^{-1}\left(\mathrm{Id} - J_{\mu A}\right)$; see [22, Lemma 12.14]. The following result is [6, Theorem 23.7(iv)]. Here, we take the opportunity to give a detailed proof.

Proposition 4.9 *For $\alpha, \mu > 0$, the following formula holds*

$$J_{\alpha Y_\mu(A)} = \frac{\mu}{\mu+\alpha}\mathrm{Id} + \frac{\alpha}{\mu+\alpha}J_{(\mu+\alpha)A}.$$

Proof First,

$$\begin{aligned}\alpha Y_\mu(A) &= \alpha\left(\mu\mathrm{Id} + A^{-1}\right)^{-1} = \left[\left(\mu\mathrm{Id} + A^{-1}\right)\left(\alpha^{-1}\mathrm{Id}\right)\right]^{-1}\\ &= \left(\alpha^{-1}\mu\mathrm{Id} + A^{-1}(\alpha^{-1}\mathrm{Id})\right)^{-1} = \left(\alpha^{-1}\mu\mathrm{Id} + (\alpha A)^{-1}\right)^{-1}\\ &= Y_{\alpha^{-1}\mu}(\alpha A).\end{aligned}$$

Thus, we only need to prove the formula holds for $\alpha = 1$.

Let $y \in X$, $z = J_{(\mu+1)A}(y)$ and $x = \frac{\mu}{\mu+1}y + \frac{1}{\mu+1}z$. We will prove $x = J_{Y_\mu(A)}(y)$. We have $z = (\mu+1)x - \mu y$, $y - z = \frac{\mu+1}{\mu}(x-z)$ and $Y_\mu(A) = \frac{1}{\mu}\left(\mathrm{Id} - J_{\mu A}\right)$. Thus,

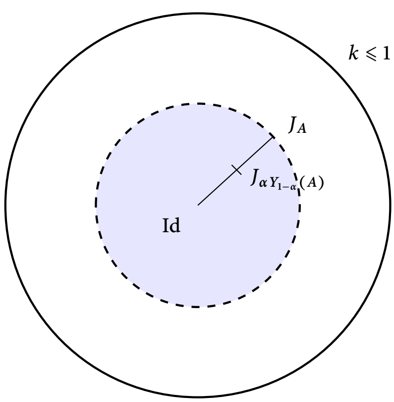
$$\begin{aligned}z = J_{(\mu+1)A}(y) &\Leftrightarrow y - z \in (\mu+1)Az \Leftrightarrow \frac{\mu+1}{\mu}(x-z) \in (\mu+1)Az\\ &\Leftrightarrow x - z \in (\mu A)z \Leftrightarrow z = J_{\mu A}x \Leftrightarrow (\mu+1)x - \mu y = J_{\mu A}x\\ &\Leftrightarrow y - x = \frac{x - J_{\mu A}x}{\mu} = Y_\mu(A)(x) \Leftrightarrow x = J_{Y_\mu(A)}(y).\end{aligned}$$

■

Combining Proposition 4.9 and Lemma 4.3, we have the following.

Corollary 4.10 *For any $\alpha \in [0,1)$, the following hold:*

(i) $J_{\alpha Y_{1-\alpha}(A)} = (1-\alpha)\mathrm{Id} + \alpha J_A$.
(ii) $k(J_{\alpha Y_{1-\alpha}}(A)) = \alpha k(J_A)$.



Corollary 4.11 *For $\mu > 0$, the following hold:*

(i) $J_{Y_\mu(A)} = \frac{\mu}{\mu+1}\mathrm{Id} + \frac{1}{\mu+1}J_{(\mu+1)A}$.
(ii) $k(J_{Y_\mu(A)}) = \frac{1}{\mu+1}k(J_{(\mu+1)A})$.

Example 4.12 Let C be a nonempty closed convex set in X and $C \neq X$. Consider the normal cone to C defined by $N_C(x) := \{u \in X \mid \sup_{c\in C}\langle c - x, u\rangle \leqslant 0\}$ if $x \in C$, and $\varnothing$ otherwise. Then,

$$(\forall \mu > 0)\ k(J_{Y_\mu(N_C)}) = k(J_{\mu^{-1}(\mathrm{Id}-P_C)}) = \frac{1}{2(\mu+1)}. \tag{4.4}$$

In particular,

$$(\forall \alpha \in (0,1))\ k(J_{\alpha Y_{1-\alpha}(N_C)}) = k(J_{\alpha(1-\alpha)^{-1}(\mathrm{Id}-P_C)}) = \frac{\alpha}{2}. \tag{4.5}$$

Proof Apply Corollary 4.10 with $A = N_C$ to obtain

$$J_{Y_\mu(N_C)} = J_{\mu^{-1}(\mathrm{Id}-J_{\mu N_C})} = J_{\mu^{-1}(\mathrm{Id}-P_C)} \tag{4.6}$$

$$= \frac{\mu}{\mu+1}\mathrm{Id} + \frac{1}{\mu+1}J_{(\mu+1)N_C} = \frac{\mu}{\mu+1}\mathrm{Id} + \frac{1}{\mu+1}J_{N_C} \tag{4.7}$$

$$= \frac{\mu}{\mu+1}\mathrm{Id} + \frac{1}{\mu+1}P_C. \tag{4.8}$$

Using Lemma 4.3 and $k(P_C) = 1/2$ because $C \neq X$, we have

$$k(J_{\mu^{-1}(\mathrm{Id}-P_C)}) = \frac{1}{\mu+1}k(P_C) = \frac{1}{2(\mu+1)}.$$

Finally, (4.5) follows from (4.4) by using $\mu = (1-\alpha)/\alpha$. ■

Remark 4.13 Observe that Corollary 4.11(i) shows that $Y_\mu(A)$ is the resolvent average of monotone operators 0 and $(\mu+1)A$ (see, e.g., [4].

4.2 Characterization of normally averaged mappings

The Yosida regularization of monotone operators provides the key. Recall that $T : X \to X$ is μ-cocoercive with $\mu > 0$ if μT is firmly nonexpansive, i.e.,

$$(\forall x \in X)(\forall y \in X) \quad \langle x - y, Tx - Ty\rangle \geqslant \mu\|Tx - Ty\|^2.$$

Fact 4.14 [6, Proposition 23.21 (ii)] *$T : X \to X$ is μ-cocoercive if and only if there exists a maximally monotone operator $A : X \rightrightarrows X$ such that $T = Y_\mu(A)$.*

Lemma 4.15 *Let $A : X \rightrightarrows X$ be maximally monotone. Suppose that J_A is normally nonexpansive. Then, A is single-valued with full domain, and cocoercive.*

Proof If $k(J_A) = 0$, Proposition 2.3 shows that $J_A = \mathrm{Id} + v$ for some $v \in X$. Then, $A := -v$, which is clearly single-valued with full domain, and cocoercive. Hence, we shall assume $0 < k(J_A) < 1/2$. Set

$$N = \frac{J_A - (1 - 2k(J_A))\,\mathrm{Id}}{2k(J_A)}.$$

Then, $J_A = (1 - 2k(J_A))\,\mathrm{Id} + 2k(J_A)\,N$ and N is nonexpansive with $k(N) = 1/2$ by Lemma 4.3. It follows from Fact 4.1 that N is firmly nonexpansive, i.e., there exists a maximally monotone operator $B : X \rightrightarrows X$ such that $N = J_B$. Thus, by Corollary 4.10, we have

$$\begin{aligned} J_A &= (1 - 2k(J_A))\,\mathrm{Id} + 2k(J_A)\,N = (1 - 2k(J_A))\,\mathrm{Id} + 2k(J_A)\,J_B \\ &= J_{2k(J_A)\,Y_{1-2k(J_A)}(B)}. \end{aligned}$$

Therefore, $A = 2k(J_A)\,Y_{1-2k(J_A)}(B)$. Since J_A is normally nonexpansive, we have $2k(J_A) \in (0,1)$. Thus, $2k(J_A)\,Y_{1-2k(J_A)}(B)$, being a Yosida regularization, is a single-valued, full domain, and cocoercive operator due to Fact 4.14. Hence, A is single-valued with full domain, and cocoercive. ∎

Lemma 4.16 *Suppose $A : X \rightrightarrows X$ is single-valued with full domain, and cocoercive. Then, J_A is normally nonexpansive.*

Proof Since A is single-valued with full domain, and cocoercive, by Fact 4.14, there exist a maximally monotone operator $B : X \rightrightarrows X$ and $\mu > 0$ such that $A = Y_\mu(B)$. Since B is maximally monotone, by Corollary 4.11, we have

$$J_{Y_\mu(B)} = \frac{\mu}{\mu+1}\mathrm{Id} + \frac{1}{\mu+1}J_{(\mu+1)B} = J_A.$$

Since B is maximally monotone and $\mu + 1 > 1$, we have $(\mu+1)B$ is maximally monotone as well. Thus, $k(J_{(\mu+1)B}) \leqslant 1/2$ by Fact 4.1. Now, Lemma 4.3 gives

$$k(J_A) = \frac{1}{\mu+1}k\left(J_{(\mu+1)B}\right) \leqslant \frac{1}{\mu+1}\cdot\frac{1}{2} < \frac{1}{2}.$$

∎

The main result of this section comes as follows.

Theorem 4.17 (Characterization of normally averaged mapping) *Let $A : X \rightrightarrows X$ be maximally monotone. Then, J_A is normally nonexpansive if and only if A is single-valued with full domain, and cocoercive.*

Proof Combine Lemmas 4.15 and 4.16. ∎

In view of Fact 4.1, the characterization of special nonexpansiveness follows immediately as well.

Example 4.18 Let $A \in \mathbb{S}^n_{++}$, the set of $n \times n$ positive definite symmetric matrices. Then, $k(J_A) < 1/2$ and $k(J_{A^{-1}}) < 1/2$ by Theorem 4.17.

The following fact follows from [8, Theorem 2.1(i)&(iv)].

Fact 4.19 *Let $T : X \to X$ be firmly nonexpansive. Then, the following hold:*

(i) *$T = J_A$ for a maximally monotone operator $A : X \rightrightarrows X$.*
(ii) *T is injective if and only if A is at most single-valued, i.e.,*

$$(\forall x \in \operatorname{dom} A) \ Ax \text{ is empty or a singleton.}$$

(iii) *T is surjective if and only if* $\operatorname{dom} A = X$.

Remark 4.20 Combining Theorem 4.17 and Facts 4.19 and 4.1, we recover Theorem 2.12, since A being cocoercive implies $\mathrm{Id} + A$ being Lipschitz.

5 Proximal operator

Let $f \in \Gamma_0(X)$. Recall that the proximal operator of f is given by

$$\mathrm{P}_f(x) := \operatorname*{argmin}_{u \in X} \left\{ f(u) + \frac{1}{2}\|u - x\|^2 \right\},$$

that the Moreau envelope of f with parameter $\mu > 0$ is defined by $e_\mu f(x) := \min_{u \in X}(f(u) + \frac{1}{2\mu}\|u - x\|^2)$, and that the Fenchel conjugate of f is defined by $f^*(y) := \sup_{x \in X}(\langle x, y\rangle - f(x))$ for $y \in X$. It is well known that $\mathrm{P}_f = (\mathrm{Id} + \partial f)^{-1}$, where ∂f is the subdifferential of f given by $\partial f(x) := \{u \in X \mid (\forall y \in X) \ f(y) \geqslant f(x) + \langle u, y - x\rangle\}$ if $x \in \operatorname{dom} f$, and $\varnothing$ if $x \notin \operatorname{dom} f$. Also, P_f is firmly nonexpansive, i.e., $k\left(\mathrm{P}_f\right) \leqslant 1/2$ (see, e.g., [6, 22]).

In this section, we will characterize the normal and special nonexpansiveness of P_f. We begin with the following definition.

Definition 5.1 (L-smoothness) Let $L \in [0, +\infty)$. Then, f is L-smooth on X if f is Fréchet differentiable on X and ∇f is L-Lipschitz, i.e.,

$$(\forall x \in X)(\forall y \in X) \quad \|\nabla f(x) - \nabla f(y)\| \leqslant L\|x - y\|.$$

Fact 5.1 ((Baillon-Haddad) [2] (see also [6, Corollary 18.17])) *Let $f \in \Gamma_0(X)$. Suppose f is Fréchet differentiable on X. Then, ∇f is μ-cocoercive if and only if ∇f is μ^{-1}-Lipschitz continuous.*

For further properties of L-smooth functions, see [6, 11, 22]. We also need

Fact 5.2 ((Moreau) [6, Theorem 20.25]) *Let $f \in \Gamma_0(X)$. Then, ∂f is maximally monotone.*

The following interesting result characterizes a L-smooth function f via the modulus of averagedness of P_f. It shows that for proximal operators not only can Theorem 4.17 be significantly improved but also the converse of Theorem 2.12 holds.

Theorem 5.3 (Characterization of normal proximal operator) *Let $f \in \Gamma_0(X)$. Then, the following are equivalent:*

(i) *P_f is normally nonexpansive.*
(ii) *There exists $L > 0$ such that f is L-smooth on X.*

(iii) f^* *is* $1/L$*-strongly convex for some* $L > 0$.
(iv) P_{f^*} *is a Banach contraction.*
(v) P_f *is a bi-Lipschitz homeomorphism of* X.

Proof "(i)⇔(ii)": By Fact 5.2, ∂f is maximally monotone. Let $A = \partial f$ in Theorem 4.17 and combine it with Fact 5.1.

"(ii)⇔(iii)": Apply [6, Theorem 18.15].

"(iii)⇔(iv)": Apply [18, Corollary 3.6].

"(i)⇒(v)": Apply Theorem 2.12.

"(v)⇒(i)": The assumption implies that $(\mathrm{P}_f)^{-1} = \mathrm{Id} + \partial f$ is full domain, single-valued and Lipschitz, so is $\partial f = \nabla f$. By Fact 5.1, ∇f is co-coercive. It remains to apply Theorem 4.17. ■

Remark 5.4 (1) Bi-Lipschitz homeomorphisms of a Euclidean space form an important class of operators. For instance, Hausdorff dimension, which plays a central role in fractal geometry and harmonic analysis, is bi-Lipschitz invariant (see [17]). Theorem 5.3(i) ⇔ (v) thus provides a large class of such nonlinear operators.

(2) By endowing $\Gamma_0(X)$ with the topology of epi-convergence (see, e.g., [21, Proposition 3.5, Corollary 4.18]), Theorem 5.3 (i) ⇔ (ii) implies that *most* convex functions have their proximal mappings with modulus of averagedness exactly 1/2, in the sense of co-meagerness (the complement of a meager set).

The characterization of special proximal operator follows immediately as well. The following example shows that P_f being only bijective does not imply that P_f is normally nonexpansive.

Example 5.5 Let $X = \mathbb{R}$. Define

$$\varphi(x) := \begin{cases} \ln x & \text{if } x \geqslant e, \\ \frac{1}{e}x & \text{if } -e < x < e, \\ -\ln(-x) & \text{if } x \leqslant -e. \end{cases}$$

Then, the following hold:

(i) φ is a proximal operator of a function in $\Gamma_0(\mathbb{R})$.
(ii) φ is a bijection.
(iii) φ is specially nonexpansive.
(iv) The inverse mapping of φ:

$$(\varphi)^{-1}(y) = \begin{cases} e^y & \text{if } y \geqslant 1, \\ ey & \text{if } -1 \leqslant y \leqslant 1, \\ -e^{-y} & \text{if } y \leqslant -1, \end{cases}$$

is not Lipschitz.

Proof (i): φ is a proximal operator because it is nonexpansive and increasing (see [6, Proposition 24.31]). (ii): Obvious. (iii): We have that φ is differentiable with

$$\varphi'(x) = \begin{cases} \frac{1}{x} & \text{if } x \geqslant e, \\ \frac{1}{e} & \text{if } -e < x < e, \\ -\frac{1}{x} & \text{if } x \leqslant -e. \end{cases}$$

Thus, $\inf_{x\in\mathbb{R}} \varphi'(x) = 0$. By Lemma 3.15 or [5, Proposition 2.8], $k(\varphi) = (1 - \inf_{x\in\mathbb{R}} \varphi'(x))/2 = 1/2$. (iv): Direct calculations. ■

Corollary 5.6 *Let $f \in \Gamma_0(X)$. Suppose* $\operatorname{dom} f \neq X$. *Then,* P_f *is specially nonexpansive.*

Proof Observe that $\operatorname{dom} f \neq X$ implies $\operatorname{dom} \partial f \neq X$. Thus, f is not L-smooth for any $L > 0$ and the result follows by Theorem 5.3. ■

Remark 5.7 When C is a nonempty closed convex subset of X and $C \neq X$, obviously $\iota_C \in \Gamma_0(X)$ and $\operatorname{dom} \iota_C = C$. By Corollary 5.6, P_C is specially nonexpansive, which recovers Example 2.15.

For the Moreau envelope, we have the following result.

Proposition 5.8 *Let $f \in \Gamma_0(X)$ and let $\mu, \alpha > 0$. Then,*

$$k(\mathrm{P}_{\alpha e_\mu f}) = \frac{\alpha}{\mu + \alpha} k(\mathrm{P}_{(\mu+\alpha) f}). \tag{5.1}$$

If, in addition, f is not Lipschitz smooth, then

$$k(\mathrm{P}_{\alpha e_\mu f}) = \frac{1}{2} \frac{\alpha}{\mu + \alpha}.$$

Proof By [9, Theorem 27.9], we have

$$\mathrm{P}_{\alpha e_\mu f} = \frac{\mu}{\mu + \alpha} \mathrm{Id} + \frac{\alpha}{\mu + \alpha} \mathrm{P}_{(\mu+\alpha) f}.$$

It suffices to apply Lemma 4.3.

If, in addition, f is not Lipschitz smooth, then $(\mu + \alpha) f$ is not Lipschitz smooth so that $k(\mathrm{P}_{(\mu+\alpha) f}) = 1/2$ by Theorem 5.3. Use (5.1) to complete the proof. ■

Example 5.9 Let $\mu, \alpha > 0$. Consider the Huber function defined by

$$H_\mu : X \to \mathbb{R} : x \mapsto \begin{cases} \frac{1}{2\mu} \|x\|^2 & \text{if } \|x\| \leqslant \mu, \\ \|x\| - \frac{\mu}{2} & \text{if } \|x\| > \mu. \end{cases}$$

It is well-known that $H_\mu = e_\mu \| \cdot \|$ and that $\| \cdot \|$ is not Lipschitz smooth, Therefore, by Proposition 5.8,

$$k(\mathrm{P}_{\alpha H_\mu}) = \frac{1}{2} \frac{\alpha}{\mu + \alpha}.$$

Example 5.10 Let C be a nonempty closed convex subset of X and $C \neq X$. Consider the support function of C defined by $\sigma_C : X \to [-\infty, +\infty] : x \mapsto \sup_{c\in C} \langle c, x\rangle$. Then, the following hold:

(i) If C is a singleton, then $k(P_C) = 1/2$ and $(\forall \lambda > 0)\ k(\mathrm{P}_{\lambda\sigma_C}) = 0$.
(ii) If C contains more than one point, then $k(P_C) = 1/2$ and $(\forall \lambda > 0)\ k(\mathrm{P}_{\lambda\sigma_C}) = 1/2$.

Proof The fact that $k(P_C) = 1/2$ has been given by Example 2.15. Now observe that the support function σ_C has $\mathrm{P}_{\lambda\sigma_C} = \mathrm{Id} - \lambda P_C(\cdot/\lambda)$.

(i): We have $\mathrm{P}_{\lambda\sigma_C} = \mathrm{Id} + v$ for some $v \in X$. Then, apply Proposition 2.3.

(ii): The function $\lambda\sigma_C(x)$ is not Lipschitz smooth, since it is not differentiable at 0. Apply Theorem 5.3 to derive $k(\mathrm{P}_{\lambda\sigma_C}) = 1/2$. ■

6 Compute modulus of averagedness via other constants or values

In this section, introducing monotone value for monotone operators, cocercive value for cocoercive mappings and Lipschitz value for Lipschitz mappings, we provide various formulae to quantify the modulus of averagedness for resolvents and proximal operators.

6.1 Monotone value and cocoercive value

Recall that we assume $A : X \rightrightarrows X$ is a maximally monotone operator. For $\mu > 0$, we say that A is μ-strongly monotone if $A - \mu\mathrm{Id}$ is monotone, i.e.,

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

It is clear that if an operator is μ_0-strongly monotone (or cocoercive), then it is μ-strongly monotone (or cocoercive) for $\mu \leqslant \mu_0$. Observing this property, we define the following functions for a maximally monotone operator.

Definition 6.1 (Monotone value) Suppose that A is strongly monotone. The monotone value (or best strong monotonicity constant) of A is defined by

$$m(A) := \sup\{\mu > 0 \mid A \text{ is } \mu\text{-strongly monotone}\}.$$

Otherwise, we define $m(A) = 0$.

Definition 6.2 (Cocoercive value) Suppose A is single-valued with full domain, and cocoercive. The cocoercive value (or best cocoercivity constant) of A is defined by

$$c(A) := \sup\{\mu > 0 \mid A \text{ is } \mu\text{-cocoercive}\}.$$

Otherwise, we define $c(A) = 0$.

We present basic properties of monotone value and cocoercive value. Note an operator is μ-cocoercive if and only if its inverse is μ-strongly monotone.

Proposition 6.1 *Let $\mu > 0$. The following hold:*

(i) *(duality) $m(A) = c\left(A^{-1}\right)$ and $m\left(A^{-1}\right) = c(A)$.*
(ii) *$m(\mu A) = \mu m(A)$ and $c(\mu A) = \mu^{-1}c(A)$.*
(iii) *$c(A) = +\infty$ if and only if A is a constant operator on X.*
(iv) *$m(A+B) \geqslant m(A) + m(B)$ and*

$$c\left(\left(A^{-1}+B^{-1}\right)^{-1}\right) \geqslant c(A) + c(B).$$

(v) *(Yosida regularization) $m(A+\mu\mathrm{Id}) = m(A) + \mu$ and $c\left(Y_\mu(A)\right) = c(A) + \mu$.*

Proof (i), (ii), (iii) and (iv) can be directly verified. (v): Since $Y_\mu(A) = \left(\mu\mathrm{Id} + A^{-1}\right)^{-1}$, we have

$$\begin{aligned} c\left(Y_\mu(A)\right) &= c\left(\left(\mu\mathrm{Id}+A^{-1}\right)^{-1}\right) = m\left(\mu\mathrm{Id}+A^{-1}\right) \\ &= \mu + m\left(A^{-1}\right) = \mu + c(A). \end{aligned}$$

■

The following fact connects averaged operators with cocoercive mappings, and can be directly verified.

Fact 6.2 [24, Proposition 3.4(iii)] *Let $T : X \to X$ be nonexpansive and $\alpha \in (0, 1)$. Then, T is α-averaged if and only if* $\mathrm{Id} - T$ *is* $1/(2\alpha)$*-cocoercive.*

Proposition 6.3 *Let $T : X \to X$ be nonexpansive. Then,*

$$k(T) = \frac{1}{2c(\mathrm{Id} - T)}.$$

Proof Combine Proposition 2.3 and Fact 6.2. ∎

Corollary 6.4 *Let $T : X \to X$ be normally nonexpansive. Then,* $\mathrm{Id} - T$ *is a Banach contraction with constant* $2k(T)$.

Proof By Proposition 6.3, $\mathrm{Id} - T$ is cocoercive with constant $1/(2k(T))$. Using the Cauchy–Schwarz inequality, we have that $\mathrm{Id} - T$ is Lipschitz with constant $2k(T)$. The contraction property follows by $2k(T) < 1$ since T is normally nonexpansive. ∎

Remark 6.5 Lemma 2.11 can also be proved by using Corollary 6.4 and the Banach fixed-point theorem. Indeed, given a normally nonexpansive T and for any $v \in X$, the mapping $x \mapsto x - Tx + v$ is a Banach contraction and, therefore, has a fixed point x_0. Then, $x_0 = x_0 - Tx_0 + v$ which implies that $Tx_0 = v$, therefore, T is surjective.

The following result connects the modulus of averagedness of a resolvent to the co-coercivity of associated maximally monotone operator.

Proposition 6.6 (Modulus of averagedness via cocoercive value) *Let $A : X \rightrightarrows X$ be maximally monotone and $\alpha > 0$. Then,*

$$k(J_{\alpha A}) = \frac{1}{2}\frac{\alpha}{\alpha + c(A)}.$$

Proof In view of Proposition 6.1(ii), it suffices to prove the case when $\alpha = 1$. Note that $Y_1(A) = \mathrm{Id} - J_A$. By Proposition 6.1(v), $c(\mathrm{Id} - J_A) = c(A) + 1$. Now apply Proposition 6.3. ∎

We have the following corollary in view of Proposition 6.1(i).

Corollary 6.7 (Modulus of averagedness via monotone value) *Let $A : X \rightrightarrows X$ be maximally monotone and $\alpha > 0$. Then,*

$$k(J_{\alpha A}) = \frac{1}{2}\frac{\alpha}{\alpha + m(A^{-1})}.$$

The following example illustrates our formulae in this section.

Example 6.8 Suppose that $A : X \to X$ is a bounded linear operator and that A is skew, i.e., $(\forall x \in X)\langle x, Ax\rangle = 0$. Then, A is maximally monotone, and the following hold:

(i) If $A \equiv 0$, then $c(A) = +\infty$. Clearly,

$$k(J_A) = k(\mathrm{Id}) = 0 = \frac{1}{2}\frac{1}{1 + \infty}.$$

(ii) If A is not a zero operator, then $c(A) = m(A) = m(A^{-1}) = c(A^{-1}) = 0$. Therefore, the formulae give $k(J_A) = k(J_{A^{-1}}) = 1/2$, which coincides with Theorem 4.17 because A and A^{-1} is not cocoercive.

6.2 Lipschitz value

Definition 6.3 (Lipschitz value) Let $T : X \to X$. The Lipschitz value (or best Lipschitz constant) of T is defined by

$$\ell(T) := \inf \{L \geqslant 0 \mid \forall x, y \in X, \|Tx - Ty\| \leqslant L\|x - y\|\}.$$

Moreover, for a maximally monotone operator $A : X \rightrightarrows X$, define $\ell(A) = +\infty$ if A is not single-valued with full domain.

The following formula connects Lipschitz value with cocoercive value. Note that we follow the convention that $\inf \varnothing = +\infty$, $(+\infty)^{-1} = 0$ and $0^{-1} = +\infty$.

Lemma 6.9 $\ell(A) \leqslant [c(A)]^{-1}$.

Proof Suppose $c(A) \in (0, +\infty)$. Then, A is $c(A)$-cocoercive and, therefore, $[c(A)]^{-1}$-Lipschitz on X by the Cauchy–Schwarz inequality. Thus, $\ell(A) \leqslant [c(A)]^{-1}$.

Suppose $c(A) = +\infty$. It follows from Proposition 6.1(iii) that A is a constant operator. Thus, $\ell(A) = 0 = [c(A)]^{-1}$.

Suppose $c(A) = 0$. Then, $\ell(A) \leqslant +\infty = [c(A)]^{-1}$. ∎

Fact 6.10 [6, Proposition 17.31] *Let f be convex and proper on X, and suppose that $x \in \operatorname{int}\operatorname{dom} f$. Then, f is Gâteaux differentiable at $x \Leftrightarrow \partial f(x)$ is a singleton in which case $\partial f(x) = \{\nabla f(x)\}$.*

Proposition 6.11 $\ell(\partial f) = [c(\partial f)]^{-1}$.

Proof Suppose $c(\partial f) \in (0, +\infty)$. Then, ∂f is singe-valued with full domain. Thus, $\partial f = \nabla f$ by Fact 6.10. While ∇f is $c(\partial f)$-cocoercive, by applying Fact 5.1, we have $\ell(\nabla f) = [c(\nabla f)]^{-1}$.

Suppose $c(\partial f) = +\infty$. Then, ∂f is a constant operator by Proposition 6.1. Thus, $\ell(\partial f) = 0 = [c(\partial f)]^{-1}$.

Suppose $c(\partial f) = 0$. If ∂f is singe-valued with full domain, then again by applying Facts 6.10 and 5.1, we have $\partial f = \nabla f$ is not Lipschitz, thus $\ell(\partial f) = +\infty = [c(\partial f)]^{-1}$. If ∂f is not singe-valued, or not with full domain, then $\ell(\partial f) = +\infty$ by the definition of Lipschitz value. Thus, $\ell(\partial f) = +\infty = [c(\partial f)]^{-1}$. ∎

Now we are able to propose the following interesting formula for proximal operators.

Theorem 6.12 (Modulus of averagedness via Lipschitz value) *Let $f \in \Gamma_0(X)$. Then,*

$$k\left(\mathrm{P}_f\right) = \frac{1}{2}\frac{1}{1 + [\ell(\partial f)]^{-1}}.$$

Proof By Fact 5.2, ∂f is maximally monotone. The result follows by letting $A = \partial f$ in Proposition 6.6 and combining it with Proposition 6.11. ∎

Using $\ell(\alpha T) = \alpha \ell(T)$ for $\alpha > 0$, we obtain the following result.

Corollary 6.13 *Let $f \in \Gamma_0(X)$ be L-smooth on X for some $L > 0$ and let $\alpha > 0$. Then,*

$$k\left(\mathrm{P}_{\alpha f}\right) = \frac{1}{2}\frac{\alpha\ell(\nabla f)}{1+\alpha\ell(\nabla f)}.$$

The following example illustrates our formulae in this section.

Example 6.14 Let C be a nonempty closed convex set in X and $C \neq X$. Consider the distance function of C defined by $d_C(x) : X \to [-\infty, +\infty] : x \mapsto \inf_{c\in C} \|x - c\|$. Then, for any $\alpha > 0$ the following hold:

(i) $k\left(\mathrm{P}_{\frac{\alpha}{2}d_C^2}\right) = \frac{1}{2}\frac{\alpha}{1+\alpha}$.
(ii) $c(\mathrm{Id} - P_C) = \ell(\mathrm{Id} - P_C) = 1$.
(iii) $\mathrm{P}_{\frac{\alpha}{2}d_C^2}$ is a bi-Lipschitz homeomorphism of X.

Proof (i): By [11, Example 6.65], $\mathrm{P}_{\frac{\alpha}{2}d_C^2} = \frac{1}{1+\alpha}\mathrm{Id} + \frac{\alpha}{1+\alpha}P_C$. Thus, we have $k\left(\mathrm{P}_{\frac{\alpha}{2}d_C^2}\right) = \frac{\alpha}{1+\alpha}k(P_C) = \frac{1}{2}\frac{\alpha}{1+\alpha}$ by Lemma 4.3 and Example 2.15.

(ii): We have $c(\mathrm{Id} - P_C) = 1$ by Proposition 6.3 and $k(P_C) = 1/2$. On the other hand, since $\frac{1}{2}d_C^2 \in \Gamma_0(X)$ and $\nabla\frac{1}{2}d_C^2 = \mathrm{Id} - P_C$ (see, e.g., [6, Corollary 12.31], we have $c(\mathrm{Id} - P_C) = \ell(\mathrm{Id} - P_C) = 1$ by Proposition 6.11.

Consequently, Corollary 6.13 is verified by the results of (i) and (ii):

$$k\left(\mathrm{P}_{\frac{\alpha}{2}d_C^2}\right) = \frac{1}{2}\frac{\alpha\ell(\nabla\frac{1}{2}d_C^2)}{1+\alpha\ell(\nabla\frac{1}{2}d_C^2)} = \frac{1}{2}\frac{\alpha\ell(\mathrm{Id}-P_C)}{1+\alpha\ell(\mathrm{Id}-P_C)} = \frac{1}{2}\frac{\alpha}{1+\alpha}.$$

(iii): By (i), $k\left(\mathrm{P}_{\frac{\alpha}{2}d_C^2}\right) = \frac{1}{2}\frac{\alpha}{1+\alpha} < \frac{1}{2}$, i.e., $\mathrm{P}_{\frac{\alpha}{2}d_C^2}$ is normally nonexpansive. The result follows by Theorem 2.12. ■

7 Bauschke, Bendit, & Moursi's example generalized

The following example on the modulus of averagedness of $P_V P_U$ extends [5, Example 3.5] in $\mathbb{R}^2$ to a Hilbert space. Instead of using [5, Theorem 3.2], we provide a much simpler proof.

Example 7.1 Let $\theta \in (0, \pi/2)$. In the product Hilbert space $\mathcal{H} = X \times X$, define

$$U = X \times \{0\}, \quad V = \left\{(y, (\tan\theta)y) \mid y \in X\right\}.$$

Then,

$$k(P_V P_U) = \frac{1+\cos\theta}{2+\cos\theta}. \tag{7.1}$$

Proof We have

$$P_U = \begin{bmatrix} \mathrm{Id} & 0 \\ 0 & 0 \end{bmatrix}, \text{ and } P_V = \begin{bmatrix} \frac{1}{1+\tan^2\theta}\mathrm{Id} & \frac{\tan\theta}{1+\tan^2\theta}\mathrm{Id} \\ \frac{\tan\theta}{1+\tan^2\theta}\mathrm{Id} & \frac{\tan^2\theta}{1+\tan^2\theta}\mathrm{Id} \end{bmatrix}$$

so that

$$P_V P_U = \begin{bmatrix} \frac{1}{1+\tan^2\theta}\mathrm{Id} & 0 \\ \frac{\tan\theta}{1+\tan^2\theta}\mathrm{Id} & 0 \end{bmatrix}.$$

Put $T = P_V P_U$. Then, T is k-averaged if and only if

$$(\forall x \in \mathcal{H}) \ \|Tx\|^2 + (1-2k)\|x\|^2 \leqslant 2(1-k)\langle x, Tx\rangle. \tag{7.2}$$

For $x = (x_1, x_2)$ with $x_i \in X$, we have

$$Tx = \left(\frac{1}{1+\tan^2\theta}x_1, \frac{\tan\theta}{1+\tan^2\theta}x_1\right), \quad \langle Tx, x\rangle = \frac{\|x_1\|^2}{1+\tan^2\theta} + \frac{\tan\theta}{1+\tan^2\theta}\langle x_1, x_2\rangle.$$

Substitute above into (7.2) to obtain

$$\frac{\|x_1\|^2}{1+\tan^2\theta} + (1-2k)(\|x_1\|^2 + \|x_2\|^2) \leqslant 2(1-k)\left(\frac{\|x_1\|^2}{1+\tan^2\theta} + \frac{\tan\theta}{1+\tan^2\theta}\langle x_1, x_2\rangle\right),$$

which can be simplified to

$$(2k-1)\frac{-\tan^2\theta}{1+\tan^2\theta}\|x_1\|^2 + (1-2k)\|x_2\|^2 - 2(1-k)\frac{\tan\theta}{1+\tan^2\theta}\langle x_1, x_2\rangle \leqslant 0. \tag{7.3}$$

When $x_2 = 0$, this gives $(2k-1)(-\tan^2\theta) \leqslant 0$, so $k \geqslant 1/2$. If $k = 1/2$, this gives

$$(\forall x_1, x_2 \in X) \ -\frac{\tan\theta}{1+\tan^2\theta}\langle x_1, x_2\rangle \leqslant 0,$$

which is impossible. Thus, $k > 1/2$. Dividing (7.3) by $\|x_2\|^2$ and applying the Cauchy–Schwarz inequality, we have

$$(2k-1)\frac{\tan^2\theta}{1+\tan^2\theta}\left(\frac{\|x_1\|}{\|x_2\|}\right)^2 + 2(1-k)\frac{\tan\theta}{1+\tan^2\theta}\left(\pm\frac{\|x_1\|}{\|x_2\|}\right) + (2k-1) \geqslant 0. \tag{7.4}$$

Substituting $t = \|x_1\|/\|x_2\|$ into (7.4) yields

$$(2k-1)\frac{\tan^2\theta}{1+\tan^2\theta}t^2 + 2(1-k)\frac{\tan\theta}{1+\tan^2\theta}(\pm t) + (2k-1) \geqslant 0$$

which happens if and only if

$$\left(2(1-k)\frac{\tan\theta}{1+\tan^2\theta}\right)^2 \leqslant 4(2k-1)^2\frac{\tan^2\theta}{1+\tan^2\theta}$$

i.e., $(1-k)^2 \leqslant (2k-1)^2(1+\tan^2\theta)$. Taking square root both sides, we have $1-k \leqslant (2k-1)/\cos\theta$, so that $k \geqslant (1+\cos\theta)/(2+\cos\theta)$. Hence,

$$k(T) = \frac{1+\cos\theta}{2+\cos\theta}.$$

■

Remark 7.2 Let U, V be two closed subspaces of $\mathcal{H}$. Recall that while the cosine of Dixmier angle between U, V is defined by

$$c_D(U, V) = \sup\left\{\langle u, v\rangle \mid u \in U, v \in V, \|u\| \leqslant 1, \|v\| \leqslant 1\right\}, \tag{7.5}$$

the cosine of the Friedrich angle between U, V is defined by

$$c_F(U, V) = \sup\left\{\langle u, v\rangle \mid u \in U \cap (U\cap V)^\perp, v \in V \cap (U \cap V)^\perp, \|u\| \leqslant 1, \|v\| \leqslant 1\right\}. \tag{7.6}$$

For more details on the angle between subspaces, see [5, 16]. With $U = X \times \{0\}$, $V = \{(y, (\tan\theta)y) \mid y \in X\}$ given in Example 7.1, for $\theta \in (0, \pi/2)$, we have $U \cap V = 0$ so that $(U \cap V)^{\perp} = \mathcal{H} = X \times X$. Then,

$$c_D(U, V) = c_F(U, V) \tag{7.7}$$

$$= \{\langle (x, 0), (y, (\tan\theta)y)\rangle \mid x \in X, y \in X, \|x\| \leqslant 1, \|(y, (\tan\theta)y)\| \leqslant 1\} \tag{7.8}$$

$$= \{\langle x, y\rangle \mid x \in X, y \in X, \|x\| \leqslant 1, \|y\| \leqslant \cos\theta\} = \cos\theta. \tag{7.9}$$

Hence, both the Dixmier and Friedrich angles between U and V are exactly θ.

Acknowledgements The authors thank the editor and an anonymous referee for careful reading and constructive comments, especially on Example 3.18 and Remark 6.5. Inspiring discussions with Dr. H.H. Bauschke benefited the article.

References

[1] J. B. Baillon, R. E. Bruck, and S. Reich, *On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces*. Houston J. Math. **4**(1978), 1–9.
[2] J. B. Baillon and G. Haddad, *Quelques propriétés des opérateurs angle-bornés et n-cycliquement monotones*. Israel J. Math. **26**(1977), 137–150.
[3] S. Bartz, M. N. Dao, and H. M. Phan, *Conical averagedness and convergence analysis of fixed point algorithms*. J. Global Optim. **82**(2022), 351–373.
[4] S. Bartz, H. H. Bauschke, S. M. Moffat, and X. Wang, *The resolvent average of monotone operators: dominant and recessive properties*. SIAM J. Optim. **26**(2016), 602–634.
[5] H. H. Bauschke, T. Bendit, and W. M. Moursi, *How averaged is the composition of two linear projections?* Numer. Funct. Anal. Optim. **44**(2023), 1652–1668
[6] H. H. Bauschke and P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*. 2nd ed., Springer, Cham, 2017.
[7] H. H. Bauschke, M. N. Dao, D. Noll, and H. M. Phan, *On Slater's condition and finite convergence of the Douglas–Rachford algorithm for solving convex feasibility problems in Euclidean spaces*. J. Global Optim. **65**(2016), 329–349.
[8] H. H. Bauschke, S. M. Moffat, and X. Wang, *Firmly nonexpansive mappings and maximally monotone operators: Correspondence and duality*. Set-Valued Var. Anal. **20**(2012), 131–153.
[9] H. H. Bauschke and W. M. Moursi, *An introduction to convexity, optimization, and algorithms*. SIAM, Philadelphia, PA, 2023.
[10] H. H. Bauschke, W. M. Moursi, and X. Wang, *Generalized monotone operators and their averaged resolvents*. Math. Program. **189**(2021), 55–74.
[11] A. Beck, *First-order methods in optimization*. SIAM, Philadelphia, PA, 2017.
[12] J. M. Borwein and Q. J. Zhu, *Techniques of variational analysis*. Springer-Verlag, New York, NY, 2005.
[13] A. Cegielski, *Iterative methods for fixed point problems in Hilbert spaces*. Springer, Heidelberg, 2012.
[14] P. L. Combettes, *Solving monotone inclusions via compositions of nonexpansive averaged operators*. Optimization **53**(2004), 475–504.
[15] P. L. Combettes and I. Yamada, *Compositions and convex combinations of averaged nonexpansive operators*. J. Math. Anal. Appl. **425**(2015), 55–70.
[16] F. Deutsch, *The angle between subspaces of a Hilbert space*. In: S. P. Singh (ed.), *Approximation theory, wavelets and applications*, NATO ASI Series. Series C, Mathematical and Physical Sciences, 454, Kluwer Academic Publishers Group, Dordrecht, 1995, pp. 107–130.
[17] K. Falconer, *Fractal geometry: Mathematical foundations and applications*. John Wiley & Sons, Chichester, 2014.
[18] H. Luo, X. Wang, and X. Yang, *Various notions of nonexpansiveness coincide for proximal mappings of functions*. SIAM J. Optim. **34**(2024), 642–653.

[19] B. S. Mordukhovich, *Variational analysis and generalized differentiation: I. Basic theory*. Springer-Verlag, Berlin, 2006.
[20] N. Ogura and I. Yamada, *Non-strictly convex minimization over the fixed point set of an asymptotically shrinking nonexpansive mapping*. Numer. Funct. Anal. Optim. **23**(2002), 113–137.
[21] C. Planiden and X. Wang, *Strongly convex functions, Moreau envelopes, and the generic nature of convex functions with strong minimizers*. SIAM J. Optim. **26**(2016), 1341–1364.
[22] R. T. Rockafellar and R. J.-B. Wets, *Variational analysis*. Springer, Berlin, 2004.
[23] W. Rudin, *Functional analysis*. 2nd ed., McGraw-Hill, New York, 1991.
[24] H. K. Xu, *Averaged mappings and the gradient-projection algorithm*. J. Optim. Theory Appl. **150**(2011), 360–378.

Department of Mathematics, I.K. Barber Faculty of Science, The University of British Columbia, Kelowna, BC, Canada
e-mail: cat688@student.ubc.ca shawn.wang@ubc.ca