Optimization Letters (2020) 14:1193–1205
https://doi.org/10.1007/s11590-019-01432-x

ORIGINAL PAPER

Computing the resolvent of the sum of operators with application to best approximation problems

Minh N. Dao[1] · **Hung M. Phan**[2]

Received: 11 September 2018 / Accepted: 3 May 2019 / Published online: 17 May 2019

Abstract
We propose a flexible approach for computing the resolvent of the sum of weakly monotone operators in real Hilbert spaces. This relies on splitting methods where strong convergence is guaranteed. We also prove linear convergence under Lipschitz continuity assumption. The approach is then applied to computing the proximity operator of the sum of weakly convex functions, and particularly to finding the best approximation to the intersection of convex sets.

Keywords Best approximation · Douglas–Rachford algorithm · Linear convergence · Operator splitting · Peaceman–Rachford algorithm · Projector · Proximity operator · Resolvent · Strong convergence

1 Introduction

In this paper, we explore a straightforward path to the problem of *computing the resolvent of the sum of two (not necessarily monotone) operators* using resolvents of individual operators. When applied to normal cones of convex sets, this computation solves the *best approximation problem* of finding the projection onto the intersection of these sets.

In general, computations involving simultaneously two or more operators are usually difficult. One popular approach is to treat each operator individually, then use these calculations to construct the desired answer. Prominent examples of such splitting strategy include the *Douglas–Rachford algorithm* [9,11] and the *Peaceman–Rachford*

✉ Hung M. Phan
hung_phan@uml.edu

Minh N. Dao
daonminh@gmail.com

1 CARMA, University of Newcastle, Callaghan, NSW 2308, Australia

2 Department of Mathematical Sciences, Kennedy College of Sciences, University of Massachusetts Lowell, Lowell, MA 01854, USA

algorithm [12] that apply to the problem of finding a zero of the sum of maximally monotone operators. In [3], the authors proposed an extension of *Dykstra's algorithm* [10] for constructing the resolvent of the sum of two maximally monotone operators. By product space reformulation, this problem was then handled in [5] for finitely many operators. Recently, the so-called *averaged alternating modified reflections algorithm* was used in [2] to study this problem, and was soon after re-derived in [1] from the view point of the proximal and resolvent average. Because computing the resolvent of a finite sum of operators can be transformed into that of the sum of two operators by a standard product space setting, as done in [2,5], we will focus on the case of two operators for simplicity.

The goal of this paper is to provide a flexible approach for computing the resolvent of the sum of two weakly monotone operators from individual resolvents. Our work extends and complements recent results in this direction. We also present applications to computing the proximity operator of the sum of two weakly convex functions and to finding the best approximation to the intersection of two convex sets.

The paper is organized as follows. In Sect. 2, we provide necessary materials. Sect. 3 contains our main results. Finally, applications are presented in Sect. 4.

2 Preparation

We assume throughout that X is a real Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and induced norm $\|\cdot\|$. The set of nonnegative integers is denoted by $\mathbb{N}$, the set of real numbers by $\mathbb{R}$, the set of nonnegative real numbers by $\mathbb{R}_+ := \{x \in \mathbb{R} \mid x \geq 0\}$, and the set of the positive real numbers by $\mathbb{R}_{++} := \{x \in \mathbb{R} \mid x > 0\}$. The notation $A\colon X \rightrightarrows X$ indicates that A is a set-valued operator on X.

Given an operator A on X, its *domain* is denoted by $\operatorname{dom} A := \{x \in X \mid Ax \neq \varnothing\}$, its *range* by $\operatorname{ran} A := A(X)$, its *graph* by $\operatorname{gra} A := \{(x,u) \in X \times X \mid u \in Ax\}$, its set of *zeros* by $\operatorname{zer} A := \{x \in X \mid 0 \in Ax\}$, and its *fixed point* set by $\operatorname{Fix} A := \{x \in X \mid x \in Ax\}$. The *inverse* of A, denoted by A^{-1}, is the operator with graph $\operatorname{gra} A^{-1} := \{(u,x) \in X \times X \mid u \in Ax\}$. Recall from [8, Definition 3.1] that an operator $A\colon X \rightrightarrows X$ is said to be *α-monotone* if $\alpha \in \mathbb{R}$ and

$$\forall (x,u),(y,v) \in \operatorname{gra} A, \quad \langle x-y, u-v\rangle \geq \alpha\|x-y\|^2. \tag{1}$$

In this case, we say that A is *monotone* if $\alpha = 0$, *strongly monotone* if $\alpha > 0$, and *weakly monotone* if $\alpha < 0$. The operator A is said to be *maximally α-monotone* if it is α-monotone and there is no α-monotone operator $B\colon X \rightrightarrows X$ such that $\operatorname{gra} B$ properly contains $\operatorname{gra} A$. It is worth mentioning that if A is maximally α-monotone with $\alpha \in \mathbb{R}_+$, then it is maximally monotone (see [8, Section 3]. Furthermore, it is also clear that A is (resp. maximally) α-monotone if and only if $A - \alpha \operatorname{Id}$ is (resp. maximally) monotone, where Id is the identity operator.

The *resolvent* of $A\colon X \rightrightarrows X$ is defined by $J_A := (\operatorname{Id} + A)^{-1}$. We conclude this section by an elementary formula for computing the resolvent of special composition via resolvents of its components.

Proposition 2.1 (Resolvent of composition) *Let* $A\colon X \rightrightarrows X$, $q, r \in X$, $\theta \in \mathbb{R}_{++}$, *and* $\sigma \in \mathbb{R}$. *Define* $\bar{A} := A \circ (\theta \operatorname{Id} - q) + \sigma \operatorname{Id} - r$ *and let* $\gamma \in \mathbb{R}_{++}$. *Then the following hold:*

(i) *A is (resp. maximally) α-monotone if and only if $\bar{A}$ is (resp. maximally) $(\theta\alpha+\sigma)$-monotone.*
(ii) *If* $1 + \gamma\sigma \neq 0$, *then*

$$J_{\gamma\bar{A}} = \frac{1}{\theta}\left(J_{\frac{\gamma\theta}{1+\gamma\sigma}A} \circ \left(\frac{\theta}{1+\gamma\sigma}\operatorname{Id} + \frac{\gamma\theta}{1+\gamma\sigma}r - q\right) + q\right); \tag{2}$$

and if, in addition, A is maximally α-monotone and $1 + \gamma(\theta\alpha + \sigma) > 0$, then $J_{\gamma\bar{A}}$ and $J_{\frac{\gamma\theta}{1+\gamma\sigma}A}$ are single-valued and have full domain.

Proof (i): This is straightforward from the definition.

(ii): We note that $(\theta \operatorname{Id} - q)^{-1} = \frac{1}{\theta}(\operatorname{Id} + q)$, that $(T - z)^{-1} = T^{-1} \circ (\operatorname{Id} + z)$, and that $(\alpha T)^{-1} = T^{-1} \circ (\frac{1}{\alpha}\operatorname{Id})$ for any operator T, any $z \in X$, and any $\alpha \in \mathbb{R} \smallsetminus \{0\}$. Using these facts yields

$$\begin{aligned}
J_{\gamma\bar{A}} &= \Big((1+\gamma\sigma)\operatorname{Id} + \gamma A \circ (\theta \operatorname{Id} - q) - \gamma r\Big)^{-1} && \text{(3a)}\\
&= \left(\left(\frac{1+\gamma\sigma}{\theta}(\operatorname{Id} + q) + \gamma A\right) \circ (\theta \operatorname{Id} - q)\right)^{-1} \circ (\operatorname{Id} + \gamma r) && \text{(3b)}\\
&= (\theta \operatorname{Id} - q)^{-1} \circ \left(\frac{1+\gamma\sigma}{\theta}\operatorname{Id} + \gamma A + \frac{1+\gamma\sigma}{\theta}q\right)^{-1} \circ (\operatorname{Id} + \gamma r) && \text{(3c)}\\
&= (\theta \operatorname{Id} - q)^{-1} \circ \left(\frac{1+\gamma\sigma}{\theta}\left(\operatorname{Id} + \frac{\gamma\theta}{1+\gamma\sigma}A\right)\right)^{-1}\\
&\quad \circ \left(\operatorname{Id} - \frac{1+\gamma\sigma}{\theta}q\right) \circ (\operatorname{Id} + \gamma r) && \text{(3d)}\\
&= \frac{1}{\theta}(\operatorname{Id} + q) \circ \left(\operatorname{Id} + \frac{\gamma\theta}{1+\gamma\sigma}A\right)^{-1} \circ \left(\frac{\theta}{1+\gamma\sigma}\operatorname{Id}\right)\\
&\quad \circ \left(\operatorname{Id} + \gamma r - \frac{1+\gamma\sigma}{\theta}q\right) && \text{(3e)}\\
&= \frac{1}{\theta}(\operatorname{Id} + q) \circ J_{\frac{\gamma\theta}{1+\gamma\sigma}A} \circ \left(\frac{\theta}{1+\gamma\sigma}\operatorname{Id} + \frac{\gamma\theta}{1+\gamma\sigma}r - q\right) && \text{(3f)}\\
&= \frac{1}{\theta}\left(J_{\frac{\gamma\theta}{1+\gamma\sigma}A} \circ \left(\frac{\theta}{1+\gamma\sigma}\operatorname{Id} + \frac{\gamma\theta}{1+\gamma\sigma}r - q\right) + q\right). && \text{(3g)}
\end{aligned}$$

Since A is maximally α-monotone, $\bar{A}$ is maximally $(\theta\alpha + \sigma)$-monotone. Now, since $1 + \gamma(\theta\alpha + \sigma) > 0$, [8, Proposition 3.4] implies the conclusion. □

3 Main results

In this section, let $A, B\colon X \rightrightarrows X$, $\omega \in \mathbb{R}_{++}$, and $r \in X$. We present a flexible approach to the computation of the resolvent at r of the scaled sum $\omega(A+B)$, that is to

$$\text{compute } J_{\omega(A+B)}(r). \tag{4}$$

Our analysis relies on the observation that this problem can be reformulated into the problem of finding a zero of the sum of two suitable operators. Indeed, when $r \in \operatorname{dom} J_{\omega(A+B)} = \operatorname{ran}(\operatorname{Id} + \omega(A+B))$, we have by definition that

$$x \in J_{\omega(A+B)}(r) \iff r \in x+\omega(A+B)x \iff 0 \in (A+B)x+\frac{1}{\omega}x-\frac{1}{\omega}r. \tag{5}$$

By writing $\frac{1}{\omega} = \sigma + \tau$ and $\frac{1}{\omega}r = r_A + r_B$, the last inclusion is equivalent to

$$0 \in (A + \sigma \operatorname{Id} - r_A)x + (B + \tau \operatorname{Id} - r_B)x, \tag{6}$$

which leads to finding a zero of the sum of two new operators $A + \sigma \operatorname{Id} - r_A$ and $B + \tau \operatorname{Id} - r_B$.

Based on the above observation, we proceed with a more general formulation. Assume throughout that

$$\theta \in \mathbb{R}_{++} \quad \text{and} \quad q \in X, \tag{7}$$

that $(\sigma, \tau) \in \mathbb{R}^2$ and $(r_A, r_B) \in X^2$ satisfy

$$\sigma + \tau = \frac{\theta}{\omega} \quad \text{and} \quad r_A + r_B = \frac{1}{\omega}(q + r), \tag{8}$$

and that

$$A_\sigma := A \circ (\theta \operatorname{Id} - q) + \sigma \operatorname{Id} - r_A \quad \text{and} \quad B_\tau := B \circ (\theta \operatorname{Id} - q) + \tau \operatorname{Id} - r_B. \tag{9}$$

Now, we will derive the formula for the resolvent of the scaled sum via zeros of the sum of these newly defined operators.

Proposition 3.1 (Resolvent via zeros of sum of operators) *Suppose that* $r \in \operatorname{ran}(\operatorname{Id} + \omega(A+B))$. *Then*

$$J_{\omega(A+B)}(r) = \theta \operatorname{zer}(A_\sigma + B_\tau) - q \neq \varnothing. \tag{10}$$

Consequently, if $A_\sigma + B_\tau$ *is strongly monotone, then* $J_{\omega(A+B)}(r)$ *and* $\operatorname{zer}(A_\sigma + B_\tau)$ *are singletons.*

Proof By assumption, $J_{\omega(A+B)}(r) \neq \varnothing$. For every $z \in X$, we derive from (8) and (9) that

$$\begin{aligned}
\theta z - q \in J_{\omega(A+B)}(r) &\iff r \in (\theta z - q) + \omega(A+B)(\theta z - q) && \text{(11a)}\\
&\iff 0 \in (A+B)(\theta z - q) + \frac{\theta}{\omega} z - \frac{1}{\omega}(q + r) && \text{(11b)}\\
&\iff 0 \in (A+B)(\theta z - q) + (\sigma + \tau) z - (r_A + r_B) && \text{(11c)}\\
&\iff 0 \in \big(A(\theta z - q) + \sigma z - r_A\big)\\
&\qquad + \big(B(\theta z - q) + \tau z - r_B\big) && \text{(11d)}\\
&\iff z \in \operatorname{zer}(A_\sigma + B_\tau). && \text{(11e)}
\end{aligned}$$

The remaining conclusion follows from [4, Proposition 23.35]. □

The new operators A_σ and B_τ along with Proposition 3.1 allow for the flexibility in chosing (σ, τ) and (r_A, r_B) as one can decide the values of these parameters as long as (8) is satisfied. We are now ready for our main result.

Theorem 3.2 (Resolvent of sum of α- and β-monotone operators) *Suppose that A and B are respectively maximally α- and β-monotone with $\alpha + \beta > -1/\omega$, that $r \in \operatorname{ran}(\operatorname{Id} + \omega(A+B))$, and that (σ, τ) satisfies*

$$\theta\alpha + \sigma > 0 \quad \textit{and} \quad \theta\beta + \tau \geq 0. \tag{12}$$

Let $\gamma \in \mathbb{R}_{++}$ be such that $1 + \gamma\sigma \neq 0$ and $1 + \gamma\tau \neq 0$. Given any $\kappa \in \,]0, 1]$ and $x_0 \in X$, define the sequence $(x_n)_{n\in\mathbb{N}}$ by

$$\forall n \in \mathbb{N}, \quad x_{n+1} := (1-\kappa)x_n + \kappa(2J_{\gamma B_\tau} - \operatorname{Id}) \circ (2J_{\gamma A_\sigma} - \operatorname{Id})x_n, \tag{13}$$

with explicit formulas

$$\begin{aligned}
J_{\gamma A_\sigma} &= \frac{1}{\theta}\left(J_{\frac{\gamma\theta}{1+\gamma\sigma}A} \circ \left(\frac{\theta}{1+\gamma\sigma}\operatorname{Id} + \frac{\gamma\theta}{1+\gamma\sigma} r_A - q\right) + q\right) && \text{(14a)}\\
\textit{and } J_{\gamma B_\tau} &= \frac{1}{\theta}\left(J_{\frac{\gamma\theta}{1+\gamma\tau}B} \circ \left(\frac{\theta}{1+\gamma\tau}\operatorname{Id} + \frac{\gamma\theta}{1+\gamma\tau} r_B - q\right) + q\right). && \text{(14b)}
\end{aligned}$$

Then $J_{\omega(A+B)}(r) = J_{\frac{\gamma\theta}{1+\gamma\sigma}A}\left(\frac{\theta}{1+\gamma\sigma}\overline{x} + \frac{\gamma\theta}{1+\gamma\sigma} r_A - q\right)$ with $\overline{x} \in \operatorname{Fix}(2J_{\gamma B_\tau} - \operatorname{Id}) \circ (2J_{\gamma A_\sigma} - \operatorname{Id})$ and the following hold:

(i) $\left(J_{\frac{\gamma\theta}{1+\gamma\sigma}A}\left(\frac{\theta}{1+\gamma\sigma}x_n + \frac{\gamma\theta}{1+\gamma\sigma} r_A - q\right)\right)_{n\in\mathbb{N}}$ *converges strongly to $J_{\omega(A+B)}(r)$.*

(ii) *If $\kappa < 1$, then $(x_n)_{n\in\mathbb{N}}$ converges weakly to $\overline{x}$.*

(iii) *If A is Lipschitz continuous, then the convergences in* (i) *and* (ii) *are linear.*

Proof We first note that the existence of $(\sigma, \tau) \in \mathbb{R}^2$ satisfying (8) and (12) is ensured since $\alpha + \beta > -1/\omega$. By Proposition 2.1(i) and (12), A_σ and B_τ are respectively maximally $(\theta\alpha + \sigma)$- and $(\theta\beta + \tau)$-monotone with $\theta\alpha + \sigma > 0$ and $\theta\beta + \tau \geq 0$, hence, by Proposition 2.1(ii), $J_{\gamma A_\sigma}$ and $J_{\gamma B_\tau}$ are single-valued and have full domain. We also see that A_σ and B_τ are maximally monotone and that A_σ and $A_\sigma + B_\tau$

are strongly monotone. Using Proposition 3.1 and [4, Proposition 26.1(iii)(b)], we have $\operatorname{zer}(A_\sigma + B_\tau) = \{J_{\gamma A_\sigma}(\overline{x})\}$ with $\overline{x} \in \operatorname{Fix}(2J_{\gamma B_\tau} - \operatorname{Id}) \circ (2J_{\gamma A_\sigma} - \operatorname{Id})$ and $J_{\omega(A+B)}(r) = \theta J_{\gamma A_\sigma}(\overline{x}) - q$.

Now, Proposition 2.1(ii) implies (14), which yields

$$\theta J_{\gamma A_\sigma} - q = J_{\frac{\gamma\theta}{1+\gamma\sigma}A} \circ \left(\frac{\theta}{1+\gamma\sigma} \operatorname{Id} + \frac{\gamma\theta}{1+\gamma\sigma} r_A - q \right). \tag{15}$$

Therefore,

$$J_{\omega(A+B)}(r) = \theta J_{\gamma A_\sigma}(\overline{x}) - q = J_{\frac{\gamma\theta}{1+\gamma\sigma}A} \left(\frac{\theta}{1+\gamma\sigma} \overline{x} + \frac{\gamma\theta}{1+\gamma\sigma} r_A - q \right). \tag{16}$$

(i): By applying [4, Theorem 26.11(vi)(b)] with all $\lambda_n = \kappa$ if $\kappa < 1$ and applying [4, Proposition 26.13] if $\kappa = 1$, we obtain that $J_{\gamma A_\sigma}(x_n) \to J_{\gamma A_\sigma}(\overline{x})$. Now combine with (15) and (16).

(ii): Again apply [4, Theorem 26.11] with all $\lambda_n = \kappa$.

(iii): Assume that A is Lipschitz continuous with constant ℓ. It is straightforward to see that A_σ is Lipschitz continuous with constant $(\theta\ell + |\sigma|)$. The conclusion follows from [8, Theorem 4.8] with $\lambda = \mu = 2$ and $\delta = \gamma$. □

Remark 3.3 Some remarks regarding Theorem 3.2 are in order.

(i) Under the assumptions made, $A + B$ is $(\alpha + \beta)$-monotone but not necessarily maximal. If, in addition, $A + B$ is indeed maximally $(\alpha + \beta)$-monotone, then $J_{\omega(A+B)}$ has full domain by [8, Proposition 3.4(ii)]; thus, the condition $r \in \operatorname{ran}(\operatorname{Id} + \omega(A + B))$ can be removed.

(ii) The iterative scheme (13) is the Douglas–Rachford algorithm if $\kappa = 1/2$ and the Peaceman–Rachford algorithm if $\kappa = 1$. For a more general version of (13), we refer the readers to [8]; see also [6,7].

(iii) If the condition (12) is replaced by

$$\theta\alpha + \sigma \geq 0 \quad \text{and} \quad \theta\beta + \tau > 0, \tag{17}$$

then the conclusions in Theorem 3.2(ii)–(iii) still hold, while Theorem 3.2(i) only holds for $\kappa < 1$; see also [5, Theorem 2.1(ii) and Remark 2.2(iv)].

(iv) One can simply choose $\theta = 1$ and $q = 0$, in which case, (14) reduces to

$$J_{\gamma A_\sigma} = J_{\frac{\gamma}{1+\gamma\sigma}A} \circ \frac{1}{1+\gamma\sigma}(\operatorname{Id} + \gamma r_A) \quad \text{and} \quad J_{\gamma B_\tau} = J_{\frac{\gamma}{1+\gamma\tau}B} \circ \frac{1}{1+\gamma\tau}(\operatorname{Id} + \gamma r_B). \tag{18}$$

(v) When A and B are maximally monotone, i.e., $\alpha = \beta = 0$, (12) is satisfied whenever $\sigma > 0$ and $\tau \geq 0$. One thus can choose for instance $\sigma = \tau = \frac{\theta}{2\omega}$.

(vi) It is always possible to find $\gamma \in \mathbb{R}_{++}$ satisfying even $1+\gamma\sigma > 0$ and $1+\gamma\tau > 0$. In fact, these inequalities are automatic regardless of $\gamma \in \mathbb{R}_{++}$ as long as σ and τ are both nonnegative.

When A and B are maximally monotone, the following result gives an iterative method for computing the resolvent of $A + B$ where each iteration relies only on the computations of J_A and J_B.

Theorem 3.4 (Resolvent of sum of two maximally monotone operators) *Suppose that A and B are maximally monotone, that $\omega > 1/2$, and that $r \in \operatorname{ran}(\mathrm{Id} + \omega(A + B))$. Define*

$$\bar{A} := \frac{2\omega}{\theta(2\omega - 1)} A \circ (\theta \, \mathrm{Id} - q) + \frac{1}{\theta(2\omega - 1)}(\theta \, \mathrm{Id} - q - r) \tag{19a}$$

$$\text{and } \bar{B} := \frac{2\omega}{\theta(2\omega - 1)} B \circ (\theta \, \mathrm{Id} - q) + \frac{1}{\theta(2\omega - 1)}(\theta \, \mathrm{Id} - q - r). \tag{19b}$$

Let $\kappa \in \,]0, 1]$, let $x_0 \in X$, and define the sequence $(x_n)_{n\in\mathbb{N}}$ by

$$\forall n \in \mathbb{N}, \quad x_{n+1} := (1 - \kappa)x_n + \kappa(2J_{\bar{B}} - \mathrm{Id}) \circ (2J_{\bar{A}} - \mathrm{Id})x_n, \tag{20}$$

with explicit formulas

$$J_{\bar{A}} = \frac{1}{\theta}\left(J_A \circ \left(\left(1 - \frac{1}{2\omega}\right)(\theta \, \mathrm{Id} - q) + \frac{1}{2\omega} r\right) + q\right) \tag{21a}$$

$$\text{and } J_{\bar{B}} = \frac{1}{\theta}\left(J_B \circ \left(\left(1 - \frac{1}{2\omega}\right)(\theta \, \mathrm{Id} - q) + \frac{1}{2\omega} r\right) + q\right). \tag{21b}$$

Then $J_{\omega(A+B)}(r) = J_A\left((1 - \frac{1}{2\omega})(\theta\overline{x} - q) + \frac{1}{2\omega}r\right)$ with $\overline{x} \in \operatorname{Fix}(2J_{\bar{B}} - \mathrm{Id}) \circ (2J_{\bar{A}} - \mathrm{Id})$ and the following hold:

(i) $\left(J_A\left((1 - \frac{1}{2\omega})(\theta x_n - q) + \frac{1}{2\omega}r\right)\right)_{n\in\mathbb{N}}$ *converges strongly to* $J_{\omega(A+B)}(r)$.
(ii) *If $\kappa < 1$, then $(x_n)_{n\in\mathbb{N}}$ converges weakly to $\overline{x}$.*
(iii) *If A is Lipschitz continuous, then the convergences in* (i) *and* (ii) *are linear.*

Proof Choosing

$$\sigma = \tau = \frac{\theta}{2\omega} > 0, \quad r_A = r_B = \frac{1}{2\omega}(q + r), \quad \text{and} \quad \gamma = \frac{2\omega}{\theta(2\omega - 1)} > 0, \tag{22}$$

we have that (8) is satisfied and that

$$A_\sigma = A \circ (\theta \, \mathrm{Id} - q) + \frac{1}{2\omega}(\theta \, \mathrm{Id} - q - r) \tag{23a}$$

$$\text{and } B_\tau = B \circ (\theta \, \mathrm{Id} - q) + \frac{1}{2\omega}(\theta \, \mathrm{Id} - q - r), \tag{23b}$$

which yields $\gamma A_\sigma = \bar{A}$ and $\gamma B_\tau = \bar{B}$. Since $1 + \gamma\sigma = 1 + \gamma\theta/(2\omega) = 2\omega/(2\omega - 1) = \gamma\theta$, we get (21) from (14). Now apply Theorem 3.2 with $\alpha = \beta = 0$. ◻

Having the freedom of choice, one can certainly use appropriate parameters to obtain simpler formulations. The following corollary illustrates one of such instances.

Corollary 3.5 *Suppose that A is maximally α-monotone with $\alpha > -1/(2\omega)$, that B is maximally β-monotone with $\beta \geq -1/(2\omega)$, and that $r \in \operatorname{ran}(\operatorname{Id}+\omega(A+B))$. Let $\eta \in \mathbb{R}_{++}$, $\kappa \in \,]0,1]$, $x_0 \in X$, and define the sequence $(x_n)_{n\in\mathbb{N}}$ by*

$$\forall n \in \mathbb{N},\ x_{n+1} := (1-\kappa)x_n + \kappa\left(2\eta J_{\omega B} \circ \frac{1}{2\eta}\operatorname{Id} + 2\eta r - \operatorname{Id}\right) \circ \left(2\eta J_{\omega A} \circ \frac{1}{2\eta}\operatorname{Id} + 2\eta r - \operatorname{Id}\right) x_n. \tag{24}$$

Then $J_{\omega(A+B)}(r) = J_{\omega A}(\frac{1}{2\eta}\overline{x})$ with $\overline{x} \in \operatorname{Fix}\left(2\eta J_{\omega B} \circ \frac{1}{2\eta}\operatorname{Id} + 2\eta r - \operatorname{Id}\right) \circ \left(2\eta J_{\omega A} \circ \frac{1}{2\eta}\operatorname{Id} + 2\eta r - \operatorname{Id}\right)$ and the following hold:

(i) $\left(J_{\omega A}(\frac{1}{2\eta}x_n)\right)_{n\in\mathbb{N}}$ *converges strongly to* $J_{\omega(A+B)}(r)$.
(ii) *If $\kappa < 1$, then $(x_n)_{n\in\mathbb{N}}$ converges weakly to $\overline{x}$.*
(iii) *If A is Lipschitz continuous, then the above convergences are linear.*

Proof We first see that $\alpha + \beta > -1/\omega$. Now choose

$$\theta = \frac{1}{\eta} > 0, \quad q = r, \quad \sigma = \tau = \frac{\theta}{2\omega}, \quad r_A = r_B = \frac{1}{2\omega}(q+r) = \frac{1}{\omega}r, \quad \text{and} \quad \gamma = \frac{2\omega}{\theta} > 0. \tag{25}$$

Then (8) and (12) are satisfied, while $\gamma\theta = 2\omega$ and $1 + \gamma\sigma = 1 + \gamma\tau = 2$. We have that

$$A_\sigma = A \circ \left(\frac{1}{\eta}\operatorname{Id} - r\right) + \frac{1}{2\eta\omega}\operatorname{Id} - \frac{1}{\omega}r \tag{26a}$$

$$\text{and} \quad B_\tau = B \circ \left(\frac{1}{\eta}\operatorname{Id} - r\right) + \frac{1}{2\eta\omega}\operatorname{Id} - \frac{1}{\omega}r. \tag{26b}$$

Noting from (14) that

$$J_{\gamma A_\sigma} = \eta\left(J_{\omega A} \circ \frac{1}{2\eta}\operatorname{Id} + r\right) \quad \text{and} \quad J_{\gamma B_\tau} = \eta\left(J_{\omega B} \circ \frac{1}{2\eta}\operatorname{Id} + r\right), \tag{27}$$

we get the conclusion due to Theorem 3.2. ◻

Again thanks to the flexibility of the parameters, our results recapture the formulation and convergence analysis of recent methods. In particular, Corollaries 3.6 and 3.7 are in the spirit of [2, Theorem 3.1] and [1, Theorem 3.2], respectively.

Corollary 3.6 *Let $\eta \in \,]0,1[$ and $\gamma \in \mathbb{R}_{++}$. Suppose that A and B are maximally monotone and that $r \in \operatorname{ran}\left(\operatorname{Id} + \frac{\gamma}{2(1-\eta)}(A+B)\right)$. Let $\kappa \in \,]0,1]$, let $x_0 \in X$, and define the sequence $(x_n)_{n\in\mathbb{N}}$ by*

$$\forall n \in \mathbb{N},\ x_{n+1} := (1-\kappa)x_n + \kappa\Big(2\eta J_{\gamma B}\circ(\operatorname{Id}+r) - 2\eta r - \operatorname{Id}\Big) \circ \Big(2\eta J_{\gamma A}\circ(\operatorname{Id}+r) - 2\eta r - \operatorname{Id}\Big)x_n. \tag{28}$$

Then $J_{\frac{\gamma}{2(1-\eta)}(A+B)}(r) = J_{\gamma A}(\overline{x}+r)$ *with* $\overline{x} \in \operatorname{Fix}(2\eta J_{\gamma B}(\operatorname{Id}+r) - 2\eta r - \operatorname{Id}) \circ (2\eta J_{\gamma A}(\operatorname{Id}+r) - 2\eta r - \operatorname{Id})$ *and the following hold:*

(i) $\big(J_{\gamma A}(x_n + r)\big)_{n\in\mathbb{N}}$ *converges strongly to* $J_{\frac{\gamma}{2(1-\eta)}(A+B)}(r)$.
(ii) *If* $\kappa < 1$, *then* $(x_n)_{n\in\mathbb{N}}$ *converges weakly to* $\overline{x}$.
(iii) *If A is Lipschitz continuous, then the above convergences are linear.*

Proof Let $\omega = \frac{\gamma}{2(1-\eta)}$, $\theta = \frac{1}{\eta}$, $q = -r$, $\sigma = \tau = \frac{\theta}{2\omega} = \frac{1-\eta}{\gamma\eta}$, and $r_A = r_B = 0$. Then (8) is satisfied,

$$A_\sigma = A \circ \left(\frac{1}{\eta}\operatorname{Id} + r\right) + \frac{1-\eta}{\gamma\eta}\operatorname{Id} \quad \text{and} \quad B_\tau = B \circ \left(\frac{1}{\eta}\operatorname{Id} + r\right) + \frac{1-\eta}{\gamma\eta}\operatorname{Id}. \tag{29}$$

Noting that $1 + \gamma\sigma = 1 + \gamma\tau = 1/\eta = \theta$, we have from (14) that

$$J_{\gamma A_\sigma} = \eta\left(J_{\gamma A}\circ(\operatorname{Id}+r) - r\right) \quad \text{and} \quad J_{\gamma B_\tau} = \eta\left(J_{\gamma B}\circ(\operatorname{Id}+r) - r\right). \tag{30}$$

Applying Theorem 3.2 with $\alpha = \beta = 0$ completes the proof. □

Corollary 3.7 *Suppose that A and B are maximally monotone and that* $A + B$ *is also maximally monotone. Let* $\eta \in\ ]0,1[$, $\kappa \in\ ]0,1]$, $x_0 \in X$, *and define the sequence* $(x_n)_{n\in\mathbb{N}}$ *by*

$$\forall n \in \mathbb{N},\ x_{n+1} := (1-\kappa)x_n + \kappa\Big(2\eta J_B + 2(1-\eta)r - \operatorname{Id}\Big) \circ \Big(2\eta J_A + 2(1-\eta)r - \operatorname{Id}\Big)x_n. \tag{31}$$

Then $J_{\frac{1}{2(1-\eta)}(A+B)}(r) = J_A(\overline{x})$ *with* $\overline{x} \in \operatorname{Fix}(2\eta J_B + 2(1-\eta)r - \operatorname{Id}) \circ (2\eta J_A + 2(1-\eta)r - \operatorname{Id})$ *and the following hold:*

(i) $(J_A(x_n))_{n\in\mathbb{N}}$ *converges strongly to* $J_{\frac{1}{2(1-\eta)}(A+B)}(r)$.
(ii) *If* $\kappa < 1$, *then* $(x_n)_{n\in\mathbb{N}}$ *converges weakly to* $\overline{x}$.
(iii) *If A is Lipschitz continuous, then the above convergences are linear.*

Proof Apply Theorem 3.4 with $\omega = \frac{1}{2(1-\eta)}$, $\theta = \frac{1}{\eta}$, and $q = \frac{1-\eta}{\eta}r = \frac{1}{2\omega-1}r$ and note that $J_{\frac{1}{2(1-\eta)}(A+B)}$ has full domain due to Remark 3.3(i). □

4 Applications

In this section, we provide transparent applications of our result to computing the proximity operator of the sum of two weakly convex functions and to finding the closest point in the intersection of closed convex sets.

We recall that a function $f\colon X \to]-\infty,+\infty]$ is *proper* if $\operatorname{dom} f := \{x \in X \mid f(x) < +\infty\} \neq \varnothing$ and *lower semicontinuous* if $\forall x \in \operatorname{dom} f$, $f(x) \leq \liminf_{z\to x} f(z)$. The function f is said to be *α-convex* for some $\alpha \in \mathbb{R}$ if $\forall x, y \in \operatorname{dom} f$, $\forall \kappa \in\,]0,1[$,

$$f((1-\kappa)x + \kappa y) + \frac{\alpha}{2}\kappa(1-\kappa)\|x-y\|^2 \leq (1-\kappa)f(x) + \kappa f(y). \tag{32}$$

We say that f is *convex* if $\alpha = 0$, *strongly convex* if $\alpha > 0$, and *weakly convex* if $\alpha < 0$.

Let $f\colon X \to]-\infty,+\infty]$ be proper. The *Fréchet subdifferential* of f at x is given by

$$\widehat{\partial} f(x) := \left\{ u \in X \;\middle|\; \liminf_{z\to x} \frac{f(z) - f(x) - \langle u, z-x\rangle}{\|z-x\|} \geq 0 \right\}. \tag{33}$$

It is known that if f is differentiable at x, then $\widehat{\partial} f(x) = \{\nabla f(x)\}$, and that if f is convex, then the Fréchet subdifferential coincides with the *convex subdifferential*, i.e.,

$$\widehat{\partial} f(x) = \partial f(x) := \{u \in X \mid \forall z \in X,\ f(z) - f(x) \geq \langle u, z-x\rangle\}. \tag{34}$$

The *proximity operator* of f with parameter $\gamma \in \mathbb{R}_{++}$ is the mapping $\operatorname{Prox}_{\gamma f}\colon X \rightrightarrows X$ defined by

$$\forall x \in X, \quad \operatorname{Prox}_{\gamma f}(x) := \operatorname*{argmin}_{z\in X} \left(f(z) + \frac{1}{2\gamma}\|z-x\|^2 \right). \tag{35}$$

Given a nonempty closed subset C of X, the *indicator function* ι_C of C is defined by $\iota_C(x) = 0$ if $x \in C$ and $\iota_C(x) = +\infty$ if $x \notin C$. It is clear that $\operatorname{Prox}_{\gamma\iota_C} = P_C$, where $P_C\colon X \rightrightarrows C$ is the *projector* onto C given by

$$\forall x \in X, \quad P_C x := \operatorname*{argmin}_{c\in C} \|x-c\|. \tag{36}$$

If C is convex, then the *normal cone* to C is the operator $N_C\colon X \rightrightarrows X$ defined by

$$\forall x \in X, \quad N_C(x) := \begin{cases} \{u \in X \mid \forall c \in C,\ \langle u, c-x\rangle \leq 0\} & \text{if } x \in C, \\ \varnothing & \text{otherwise.} \end{cases} \tag{37}$$

Lemma 4.1 *Let $f\colon X \to]-\infty,+\infty]$ and $g\colon X \to]-\infty,+\infty]$ be proper, lower semicontinuous, and respectively α- and β-convex, let $\omega \in \mathbb{R}_{++}$, and let* $r \in \operatorname{ran}(\operatorname{Id} + \omega(\widehat{\partial} f + \widehat{\partial} g))$. *Suppose that $\alpha + \beta > -1/\omega$. Then*

$$J_{\omega(\widehat{\partial} f + \widehat{\partial} g)}(r) = J_{\omega\widehat{\partial}(f+g)}(r) = \operatorname{Prox}_{\omega(f+g)}(r). \tag{38}$$

Consequently, if C and D are closed convex subsets of X with $C \cap D \neq \varnothing$ and $r \in \operatorname{ran}(\operatorname{Id} + N_C + N_D)$*, then*

$$J_{N_C+N_D}(r) = P_{C\cap D}(r). \tag{39}$$

Proof On the one hand, noting that $\operatorname{ran}(\mathrm{Id} + \omega(\widehat{\partial} f + \widehat{\partial} g)) = \operatorname{dom} J_{\omega(\widehat{\partial} f + \widehat{\partial} g)}$ and that $\widehat{\partial} f + \widehat{\partial} g \subseteq \widehat{\partial}(f + g)$, we have for any $p \in X$ that

$$\begin{aligned} p \in J_{\omega(\widehat{\partial} f + \widehat{\partial} g)}(r) &\iff r \in p + \omega(\widehat{\partial} f + \widehat{\partial} g)(p) && (40a)\\ &\implies r \in p + \omega\widehat{\partial}(f + g)(p) && (40b)\\ &\iff p \in J_{\omega\widehat{\partial}(f+g)}(r). && (40c)\end{aligned}$$

On the other hand, it is straightforward from definition that $f + g$ is $(\alpha + \beta)$-convex. Since $1 + \omega(\alpha + \beta) > 0$, we learn from [8, Lemma 5.2] that $\operatorname{Prox}_{\omega(f+g)} = J_{\omega\widehat{\partial}(f+g)}$ is single-valued and has full domain. Combining with (40) implies (38).

Now, since C and D are closed convex sets, ι_C and ι_D are convex functions, and therefore, $\widehat{\partial}\iota_C = \partial\iota_C = N_C$ and $\widehat{\partial}\iota_D = \partial\iota_D = N_D$ (see, e.g., [4, Example 16.13]). Applying (38) to $(f, g) = (\iota_C, \iota_D)$ and $\omega = 1$ yields

$$J_{N_C+N_D}(r) = \operatorname{Prox}_{\iota_C+\iota_D}(r) = \operatorname{Prox}_{\iota_{C\cap D}}(r) = P_{C\cap D}(r), \tag{41}$$

which completes the proof. □

We now derive some applications of Theorem 3.2. In what follows, $\theta \in \mathbb{R}_{++}$ and $q \in X$.

Corollary 4.2 (Proximity operator of sum of α- and β-convex functions) *Let $f : X \to]-\infty, +\infty]$ and $g : X \to]-\infty, +\infty]$ be proper, lower semicontinuous, and respectively α- and β-convex, let $\omega \in \mathbb{R}_{++}$, and let $r \in \operatorname{ran}(\mathrm{Id} + \omega(\widehat{\partial} f + \widehat{\partial} g))$. Suppose that $\alpha + \beta > -1/\omega$ and let $(\sigma, \tau) \in \mathbb{R}^2$ and $(r_f, r_g) \in X^2$ be such that $\sigma + \tau = \theta/\omega$, $r_f + r_g = (q + r)/\omega$,*

$$\theta\alpha + \sigma > 0 \quad \text{and} \quad \theta\beta + \tau \geq 0. \tag{42}$$

Let $\gamma \in \mathbb{R}_{++}$ be such that $1 + \gamma\sigma > 0$ and $1 + \gamma\tau > 0$. Given any $\kappa \in]0, 1]$ and $x_0 \in X$, define the sequence $(x_n)_{n\in\mathbb{N}}$ by

$$\forall n \in \mathbb{N}, \quad x_{n+1} := (1 - \kappa)x_n + \kappa R_g R_f x_n, \tag{43}$$

where

$$R_f := \frac{2}{\theta}\left(\operatorname{Prox}_{\frac{\gamma\theta}{1+\gamma\sigma} f} \circ \left(\frac{\theta}{1 + \gamma\sigma}\,\mathrm{Id} + \frac{\gamma\theta}{1 + \gamma\sigma} r_f - q\right) + q\right) - \mathrm{Id} \tag{44a}$$

$$\text{and } R_g := \frac{2}{\theta}\left(\operatorname{Prox}_{\frac{\gamma\theta}{1+\gamma\tau} g} \circ \left(\frac{\theta}{1 + \gamma\tau}\,\mathrm{Id} + \frac{\gamma\theta}{1 + \gamma\tau} r_g - q\right) + q\right) - \mathrm{Id}\,. \tag{44b}$$

Then $\operatorname{Prox}_{\omega(f+g)}(r) = \operatorname{Prox}_{\frac{\gamma\theta}{1+\gamma\sigma} f}\left(\frac{\theta}{1+\gamma\sigma}\overline{x} + \frac{\gamma\theta}{1+\gamma\sigma} r_f - q\right)$ with $\overline{x} \in \operatorname{Fix} R_g R_f$ and the following hold:

(i) $\left(\operatorname{Prox}_{\frac{\gamma\theta}{1+\gamma\sigma} f}\left(\frac{\theta}{1+\gamma\sigma} x_n + \frac{\gamma\theta}{1+\gamma\sigma} r_f - q\right)\right)_{n\in\mathbb{N}}$ *converges strongly to* $\operatorname{Prox}_{\omega(f+g)}(r)$.

(ii) *If* $\kappa < 1$, *then* $(x_n)_{n\in\mathbb{N}}$ *converges weakly to* $\overline{x}$.
(iii) *If* f *is differentiable with Lipschitz continuous gradient, then the above convergences are linear.*

Proof By assumption, [8, Lemma 5.2] implies that $\widehat{\partial} f$ and $\widehat{\partial} g$ are respectively maximally α- and β-monotone, and that

$$J_{\frac{\gamma}{1+\gamma\sigma}\widehat{\partial} f} = \mathrm{Prox}_{\frac{\gamma}{1+\gamma\sigma}f} \quad \text{and} \quad J_{\frac{\gamma}{1+\gamma\tau}\widehat{\partial} g} = \mathrm{Prox}_{\frac{\gamma}{1+\gamma\tau}g}\,. \tag{45}$$

Next, Lemma 4.1 implies that $J_{\omega(\widehat{\partial} f+\widehat{\partial} g)}(r) = \mathrm{Prox}_{\omega(f+g)}(r)$. The conclusion then follows from Theorem 3.2 applied to $(A, B) = (\widehat{\partial} f, \widehat{\partial} g)$. □

Corollary 4.3 (Projection onto intersection of two closed convex sets) *Let* C *and* D *be closed convex subsets of* X *with* $C \cap D \neq \varnothing$ *and let* $r \in \mathrm{ran}(\mathrm{Id} + N_C + N_D)$. *Let* $\sigma \in \mathbb{R}_{++}$, $\tau \in \mathbb{R}_{+}$, *and* $(r_C, r_D) \in X^2$ *with* $r_C + r_D = (\sigma + \tau)(q + r)/\theta$. *Let also* $\gamma \in \mathbb{R}_{++}$, $\kappa \in\ ]0, 1]$, $x_0 \in X$, *and define the sequence* $(x_n)_{n\in\mathbb{N}}$ *by*

$$\forall n \in \mathbb{N}, \quad x_{n+1} := (1-\kappa)x_n + \kappa \bar{R}_D \bar{R}_C x_n, \tag{46}$$

where

$$\bar{R}_C := \frac{2}{\theta}\left(P_C \circ \left(\frac{\theta}{1+\gamma\sigma}\,\mathrm{Id} + \frac{\gamma\theta}{1+\gamma\sigma} r_C - q\right) + q\right) - \mathrm{Id} \tag{47a}$$

$$\text{and } \bar{R}_D := \frac{2}{\theta}\left(P_D \circ \left(\frac{\theta}{1+\gamma\tau}\,\mathrm{Id} + \frac{\gamma\theta}{1+\gamma\tau} r_D - q\right) + q\right) - \mathrm{Id}\,. \tag{47b}$$

Then $\left(P_C\left(\frac{\theta}{1+\gamma\sigma}x_n + \frac{\gamma\theta}{1+\gamma\sigma}r_C - q\right)\right)_{n\in\mathbb{N}}$ *converges strongly to* $P_{C\cap D}(r) = P_C \left(\frac{\theta}{1+\gamma\sigma}\overline{x} + \frac{\gamma\theta}{1+\gamma\sigma}r_C - q\right)$ *with* $\overline{x} \in \mathrm{Fix}\, \bar{R}_D \bar{R}_C$. *Furthermore, if* $\kappa < 1$, *then* $(x_n)_{n\in\mathbb{N}}$ *converges weakly to* $\overline{x}$.

Proof We first derive from [4, Example 20.26] that N_C and N_D are maximally monotone and from [4, Example 23.4] that

$$J_{\frac{\gamma}{1+\gamma\sigma}N_C} = J_{N_C} = P_C \quad \text{and} \quad J_{\frac{\gamma}{1+\gamma\tau}N_D} = J_{N_D} = P_D. \tag{48}$$

Setting $\omega := \theta/(\sigma+\tau) > 0$, we note that $r \in \mathrm{ran}(\mathrm{Id} + N_C + N_D) = \mathrm{ran}\left(\mathrm{Id} + \omega(N_C + N_D)\right)$ and from Lemma 4.1 that $J_{\omega(N_C+N_D)}(r) = J_{N_C+N_D}(r) = P_{C\cap D}(r)$. Now apply Theorem 3.2 to $(A, B) = (N_C, N_D)$. □

As in the proof of Corollary 3.6, if we choose $\theta = \frac{1}{\eta}$ (with $\eta \in\ ]0, 1[$), $q = -r$, $\sigma = \tau = \frac{1-\eta}{\gamma\eta}$, and $r_C = r_D = 0$, then Corollary 4.3 reduces to [2, Corollary 3.1] where (46) reads as

$$\forall n \in \mathbb{N},\ x_{n+1} := (1-\kappa)x_n + \kappa\Big(2\eta P_D \circ (\mathrm{Id} + r) - 2\eta r - \mathrm{Id}\Big) \circ \Big(2\eta P_C \circ (\mathrm{Id} + r) - 2\eta r - \mathrm{Id}\Big)x_n. \tag{49}$$

Similarly, if $\theta = \frac{1}{\eta}$ (with $\eta \in \mathbb{R}_{++}$), $q = r$, $\sigma = \tau = \frac{1}{\gamma}$, and $r_C = r_D = \frac{2\eta}{\gamma}r$, then (46) is simplified to

$$\forall n \in \mathbb{N},\ x_{n+1} := (1-\kappa)x_n + \kappa\left(2\eta P_D \circ \frac{1}{2\eta}\,\mathrm{Id} + 2\eta r - \mathrm{Id}\right) \circ \left(2\eta P_C \circ \frac{1}{2\eta}\,\mathrm{Id} + 2\eta r - \mathrm{Id}\right)x_n. \tag{50}$$

Acknowledgements The authors thank two anonymous referees for their careful reading and helpful comments. MND was partially supported by Discovery Project 160101537 from the Australian Research Council. HMP was partially supported by Autodesk, Inc. via a gift made to the Department of Mathematical Sciences, University of Massachusetts Lowell.

References

1. Alwadani, S., Bauschke, H.H., Moursi, W.M., Wang, X.: On the asymptotic behaviour of the Aragón Artacho–Campoy algorithm. Oper. Res. Lett. **46**(6), 585–587 (2018)
2. Aragón Artacho, F.J., Campoy, R.: Computing the resolvent of the sum of maximally monotone operators with the averaged alternating modified reflections algorithm. J. Optim. Theory Appl. (2019)
3. Bauschke, H.H., Combettes, P.L.: A Dykstra-like algorithm for two monotone operators. Pac. J. Optim. **4**(3), 383–391 (2008)
4. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces, 2nd edn. Springer, New York (2017)
5. Combettes, P.L.: Iterative construction of the resolvent of a sum of maximal monotone operators. J. Convex Anal. **16**(3), 727–748 (2009)
6. Dao, M.N., Phan, H.M.: Linear convergence of projection algorithms. Math. Oper. Res. (2018). https://doi.org/10.1287/moor.2018.0942
7. Dao, M.N., Phan, H.M.: Linear convergence of the generalized Douglas–Rachford algorithm for feasibility problems. J. Glob. Optim. **72**(3), 443–474 (2018)
8. Dao, M.N., Phan, H.M.: Adaptive Douglas–Rachford splitting algorithm for the sum of two operators (2018). arXiv:1809.00761
9. Douglas, J., Rachford, H.H.: On the numerical solution of heat conduction problems in two and three space variables. Trans. Am. Math. Soc. **82**, 421–439 (1956)
10. Dykstra, R.L.: An algorithm for restricted least squares regression. J. Am. Stat. Assoc. **78**(384), 837–842 (1983)
11. Lions, P.-L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. **16**(6), 964–979 (1979)
12. Peaceman, D.W., Rachford, H.H.: The numerical solution of parabolic and elliptic differential equations. J. Soc. Ind. Appl. Math. **3**, 28–41 (1955)