Compositions and Averages of Two Resolvents: Relative Geometry of Fixed Points Sets and a Partial Answer to a Question by C. Byrne

Xianfu Wang* and Heinz H. Bauschke†

April 1, 2011 (originally submitted March 24, 2010)

Abstract

We show that the set of fixed points of the average of two resolvents can be found from the set of fixed points for compositions of two resolvents associated with scaled monotone operators. Recently, the proximal average has attracted considerable attention in convex analysis. Our results imply that the minimizers of proximal-average functions can be found from the set of fixed points for compositions of two proximal mappings associated with scaled convex functions. When both convex functions in the proximal average are indicator functions of convex sets, least squares solutions can be completely recovered from the limiting cycles given by compositions of two projection mappings. This provides a partial answer to a question posed by C. Byrne. A novelty of our approach is to use the notion of resolvent average and proximal average.

2010 Mathematics Subject Classification:
Primary 47H05, 47H09; Secondary 26B25, 47H10, 47J25, 90C25.

Keywords:
Attouch-Théra duality, convex function, Fenchel-Rockafellar duality, firmly nonexpansive mapping, fixed point, Hilbert space, monotone operator, Moreau envelope, projection, proximal average, proximal mapping, proximal point method, resolvent average, resolvent composition, strongly nonexpansive mapping, Yosida regularization.

*Mathematics, Irving K. Barber School, The University of British Columbia Okanagan, Kelowna, B.C. V1V 1V7, Canada. Email: shawn.wang@ubc.ca.

†Mathematics, Irving K. Barber School, The University of British Columbia Okanagan, Kelowna, B.C. V1V 1V7, Canada. Email: heinz.bauschke@ubc.ca.

1 Introduction

Throughout, $\mathcal{H}$ is a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\|\cdot\|$, and $\Gamma(\mathcal{H})$ is the set of proper lower semicontinuous convex functions on $\mathcal{H}$. Let $A : \mathcal{H} \rightrightarrows 2^{\mathcal{H}}$ be a set-valued operator with graph $\operatorname{gr} A := \big\{(x, u) \in \mathcal{H} \times \mathcal{H} \mid u \in Ax\big\}$. The set-valued inverse A^{-1} of A has graph $\big\{(u, x) \in \mathcal{H} \mid u \in Ax\big\}$, and the resolvent of A is $J_A := (A + \mathrm{Id})^{-1}$ where $\mathrm{Id} : \mathcal{H} \to \mathcal{H}$ denotes the identity mapping. The operator A is monotone if $\langle x - y, u - v\rangle \geq 0$ for all $(x, u), (y, v) \in \operatorname{gr} A$; A is maximal monotone if A is monotone and no proper enlargement of $\operatorname{gr} A$ is monotone.

Let A_1, A_2 be two maximal monotone operators, and $\lambda_1 + \lambda_2 = 1$ with $\lambda_i > 0$. The *resolvent average* of A_1, A_2 with weights λ_1, λ_2 is defined by

$$A := [\lambda_1 J_{A_1} + \lambda_2 J_{A_2}]^{-1} - \mathrm{Id},$$

and it owes its name to the identity

$$J_A = \lambda_1 J_{A_1} + \lambda_2 J_{A_2}.$$

This paper is concerned with the relationships among the fixed point sets of the resolvent average J_A*, the resolvent compositions* $J_{A_1/\lambda_2} J_{A_2/\lambda_1}$ *and* $J_{A_2/\lambda_1} J_{A_1/\lambda_2}$*. Although there appears to be no clear relationships between the fixed point sets of* $\operatorname{Fix}(\lambda_1 J_{A_1} + \lambda_2 J_{A_2})$*, and of* $\operatorname{Fix} J_{A_1} J_{A_2}$ *and* $\operatorname{Fix} J_{A_2} J_{A_1}$*, we will observe that* $\operatorname{Fix}(\lambda_1 J_{A_1} + \lambda_2 J_{A_2})$ *can be completely recovered from* $\operatorname{Fix}(J_{A_1/\lambda_2} J_{A_2/\lambda_1})$ *or* $\operatorname{Fix}(J_{A_2/\lambda_1} J_{A_1/\lambda_2})$*.*

Our investigation relies on the resolvent average and proximal average, [8, 9, 10, 12, 11]. Although compositions of resolvents (even more generally strongly nonexpansive mappings) have been studied [7, 18, 19, 16, 14, 5, 6, 21], the connections between the fixed point set of compositions and the fixed point set of the average of two resolvents appear to be new, *even when specialized to projection operators.*

The paper is organized as follows. Section 2 gathers some known facts used in later sections. In Section 3 we concentrate on resolvents. In order to find zeros of the resolvent average, we consider several inclusion problems. It turns out that their solution sets can be characterized in terms of fixed point sets associated with the resolvent average and with resolvent compositions. We provide homeomorphisms among these fixed point sets. In Section 4 we apply — and refine — the results of Section 3 to proximal mappings. The inclusion problems now translate into finding minimizers of proximal averages of convex functions. The Yosida regularization is the key tool for monotone operators, and its role is played by the Moreau envelope for convex functions. When specialized to projections, our results say that the least square solutions can be completed recovered from the solutions of alternating projections. This answers one of the question posed in [15, page 305] by Byrne for two sets, while the question for more than two sets is still open. In Section 5 we give three examples to illustrate our results. They illustrate that a recovery of $\operatorname{Fix}(\lambda_1 J_{A_1} + \lambda_2 J_{A_2})$ from $\operatorname{Fix}(J_{A_1} J_{A_2})$ and $\operatorname{Fix}(J_{A_2} J_{A_1})$ seems impossible.

Our notation is standard and follows, e.g., [26, 28, 29]. For a monotone operator $A : \mathcal{H} \rightrightarrows 2^{\mathcal{H}}$, the sets $\operatorname{dom} A := \big\{x \in \mathcal{H} \mid Ax \neq \varnothing\big\}, \operatorname{ran} A := \big\{u \in \mathcal{H} \mid (\exists x \in \mathcal{H})\ u \in Ax\big\}$ are the domain, range of

A respectively. It will be convenient to write $\widetilde{A} := (-\operatorname{Id}) \circ A^{-1} \circ (-\operatorname{Id})$. The Yosida approximation of A of index $\gamma \in (0, +\infty)$ is given by

$$(1) \qquad {}^{\gamma}A := (\operatorname{Id} - J_{\gamma A})/\gamma = (\gamma \operatorname{Id} + A^{-1})^{-1}.$$

For a mapping $T : D \to \mathcal{H}$, where $D \subseteq \mathcal{H}$, the fixed point set of T will be denoted by $\operatorname{Fix} T := \{x \in \mathcal{H} \mid Tx = x\}$. A mapping T between metric spaces X and Y is called a *homeomorphism* if T is a bijection (i.e., one-to-one and onto), T is continuous and its inverse T^{-1} is also continuous. For a sequence $(x_n)_{n\in\mathbb{N}}$ of $\mathcal{H}$, $x_n \rightharpoonup x \in \mathcal{H}$ means that $(x_n)_{n\in\mathbb{N}}$ converges weakly to x.

For a proper lower semicontinuous function $f \in \Gamma(\mathcal{H})$, the subdifferential operator $\partial f : \mathcal{H} \rightrightarrows \mathcal{H}$ of f which is given by $x \mapsto \partial f(x) := \{x^* \in \mathcal{H} \mid f(y) \geq f(x) + \langle x^*, y - x\rangle \ \forall y \in \mathcal{H}\}$ is maximal monotone. The resolvent of ∂f is called the *proximal mapping* of f, i.e., $\operatorname{Prox}_f := J_{\partial f}$. Note that Prox_f has a full domain. Also, f^* denotes the Fenchel conjugate of f, i.e., $(\forall x^* \in \mathcal{H})\ f^*(x^*) := \sup_x \left(\langle x^*, x\rangle - f(x)\right)$. The *Moreau envelope* of f with parameter γ is given by

$$e_\gamma f(x) := \inf_y \left(f(y) + \frac{1}{2\gamma}\|x - y\|^2 \right) \quad \text{for every } x \in \mathcal{H}.$$

The domain of f will be denoted by $\operatorname{dom} f$. For $f_1, f_2 \in \Gamma(\mathcal{H})$, $f_1 \oplus f_2$ means $(f_1 \oplus f_2)(x, y) := f_1(x) + f_2(y)$ for all $x, y \in \mathcal{H}$. We let $j(x) := \|x\|^2/2$ for every $x \in \mathcal{H}$ and we will use j and $\|\cdot\|^2/2$ interchangeably. For a subset $C \subseteq \mathcal{H}$, the indicator function is defined by $\iota_C(x) = 0$ if $x \in C$ and $+\infty$ otherwise. We use $d_C(x) := \inf\{\|x - y\| \mid y \in C\}$ for every $x \in \mathcal{H}$ for the distance function, $P_C := \operatorname{Prox}_{\iota_C}$ for the projection on set C, $N_C := \partial \iota_C$ for the normal cone operator, and $\operatorname{int} C$ for the interior of the set C.

2 Auxiliary results and facts

We gather some facts on strongly nonexpansive mapping, on the proximal point algorithm, and on fixed point sets of compositions of two resolvents.

Definition 2.1 *Let $T : D \to \mathcal{H}$, where $D \subseteq \mathcal{H}$. We say that*

(i) T *is* nonexpansive *if*

$$\|Tx - Ty\| \leq \|x - y\| \quad \forall x, y \in D;$$

(ii) T *is* strongly nonexpansive *if* T *is nonexpansive and* $(x_n - y_n) - (Tx_n - Ty_n) \to 0$ *whenever* $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}$ *are sequences in* D *such that* $(x_n - y_n)_{n\in\mathbb{N}}$ *is bounded and* $\|x_n - y_n\| - \|Tx_n - Ty_n\| \to 0$*;*

(iii) T *is* firmly nonexpansive *if*

$$\|Tx - Ty\|^2 \leq \langle Tx - Ty, x - y\rangle \quad \forall x, y \in D;$$

(iv) T *is* attracting *if* T *is nonexpansive and for every* $x \notin \operatorname{Fix} T, y \in \operatorname{Fix} T$ *one has*

$$\|Tx - Ty\| < \|x - y\|.$$

The following fact is well-known.

Fact 2.2 *Let* $B : \mathcal{H} \rightrightarrows 2^{\mathcal{H}}$ *be monotone operator and* $\gamma > 0$. *Then*

(i) $B^{-1}(0) = \operatorname{Fix}(J_{\gamma B}) = ({}^{\gamma}B)^{-1}(0)$.

(ii) J_B *is firmly nonexpansive.*

(iii) J_B *has a full domain if and only if* B *is maximal monotone.*

Proof. (i). This may be readily verified using definitions involved. (ii) and (iii): See [23] or [2, Fact 6.2, Corollary 6.3]. ■

Fact 2.3 (Bruck & Reich [14]) *Let* T, *and* $(T_i)_{1\le i\le m}$ *be operators from* $\mathcal{H}$ *to* $\mathcal{H}$. *Then the following properties hold:*

(i) *If* T *is firmly nonexpansive, then it is strongly nonexpansive.*

(ii) *If the operators* $(T_i)_{1\le i\le m}$ *are strongly nonexpansive, then the composition* $T_1 \cdots T_m$ *is also strongly nonexpansive.*

(iii) *If* T_1 *is strongly nonexpansive and* T_2 *is nonexpansive and* $0 < c < 1$, *then* $S = (1-c)T_1 + cT_2$ *is strongly nonexpansive.*

(iv) *Suppose that* T *is strongly nonexpansive and let* $x_0 \in \mathcal{H}$. *If* $\operatorname{Fix} T \neq \varnothing$, *then the sequence* $(T^n x_0)_n$ *converges weakly to some point in* $\operatorname{Fix} T$; *otherwise,* $\|T^n x_0\| \to \infty$.

Fact 2.2(ii) and Fact 2.3 immediately give the following result.

Corollary 2.4 *Let* $A_1, A_2 : \mathcal{H} \rightrightarrows \mathcal{H}$ *be maximal monotone operators. For* $x_0 \in \mathcal{H}$ *let* $(x_n)_{n\in\mathbb{N}}$ *be generated by*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = J_{A_1} J_{A_2} x_n;$$

For $y_0 \in \mathcal{H}$ *let* $(y_n)_{n\in\mathbb{N}}$ *be generated by*

$$(\forall n \in \mathbb{N}) \quad y_{n+1} = J_{A_2} J_{A_1} y_n.$$

If $\operatorname{Fix} J_{A_1} J_{A_2} \neq \varnothing$, *then* (x_n) *converges weakly to some point of* $\operatorname{Fix} J_{A_1} J_{A_2}$, *and* (y_n) *converges weakly to some point of* $\operatorname{Fix} J_{A_2} J_{A_1}$.

Fact 2.5 (Rockafellar [27]) *Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be maximal monotone. Assume that $A^{-1}(0) \neq \varnothing$. For any starting point x_0, the sequence (x_n) generated by the proximal point algorithm*

$$x_{n+1} = J_A(x_n) = (\mathrm{Id} + A)^{-1}(x_n)$$

converges weakly to a point in $A^{-1}(0)$ and $x_{n+1} - x_n \to 0$.

Let R denote the "transpose" mapping on $\mathcal{H} \times \mathcal{H}$, namely $R : \mathcal{H} \times \mathcal{H} \to \mathcal{H} \times \mathcal{H} : (x, y) \mapsto (y, x)$.

Fact 2.6 (See [7].) *Let $A, B : \mathcal{H} \rightrightarrows \mathcal{H}$ be maximal monotone operators and $\gamma \in (0, +\infty)$. Set*

$$\begin{aligned} S &:= (\mathrm{Id} - R + \gamma(A \times B))^{-1}(0, 0). \\ S^* &:= [(\mathrm{Id} - R)^{-1} + (A^{-1} \times B^{-1}) \circ (\mathrm{Id}/\gamma)]^{-1}(0, 0). \\ E &:= (A + {}^{\gamma}B)^{-1}(0), \quad F := (B + {}^{\gamma}A)^{-1}(0). \\ u^* &:= J_{(A^{-1} + \widetilde{B})/\gamma}(0) \text{ and } v^* := J_{(\widetilde{A} + B^{-1})/\gamma}(0). \end{aligned}$$

Then $S \neq \varnothing \Leftrightarrow S^* \neq \varnothing \Leftrightarrow E \neq \varnothing \Leftrightarrow F \neq \varnothing \Leftrightarrow u^*$ is well defined $\Leftrightarrow v^*$ is well defined, in which case the following hold.

(i) $E = \mathrm{Fix}\, J_{\gamma A} J_{\gamma B} = J_{\gamma A}(F)$ *and* $F = \mathrm{Fix}\, J_{\gamma B} J_{\gamma A} = J_{\gamma B}(E)$.

(ii) $S = \mathrm{Fix}\, J_{\gamma(A \times B)} R = (E \times F) \cap \mathrm{gr}\, J_{\gamma B}$.

(iii) $S^* = \{(\gamma u^*, \gamma v^*)\}$ *and* $u^* = -v^*$.

(iv) $S^* = (R - \mathrm{Id})(S)$.

(v) $J_{\gamma B}|_E : E \to F : x \mapsto x + \gamma u^*$ *is a bijection with inverse* $J_{\gamma A}|_F : F \to E : y \mapsto y + \gamma v^*$.

(vi) $E = A^{-1}(u^*) \cap ({}^{\gamma}B)^{-1}(v^*)$ *and* $F = ({}^{\gamma}A)^{-1}(u^*) \cap B^{-1}(v^*)$.

(vii) $S = (E \times F) \cap (R - \mathrm{Id})^{-1}(S^*)$.

Fact 2.7 (See, e.g., [5, Propositions 2.10, 2.12].) *Assume that T_1, T_2 are attracting and $\mathrm{Fix}\, T_1 \cap \mathrm{Fix}\, T_2 \neq \varnothing$. Let $\lambda_1 + \lambda_2 = 1$, with each $\lambda_i > 0$. Then*

$$\mathrm{Fix}(\lambda_1 T_1 + \lambda_2 T_2) = \mathrm{Fix}\, T_1 \cap \mathrm{Fix}\, T_2 = \mathrm{Fix}(T_1 \circ T_2) = \mathrm{Fix}(T_2 \circ T_1).$$

The class of attractive mappings properly contains the class of strongly nonexpansive mappings. See also [14, Lemma 2.1] for results related to Fact 2.7.

The following two facts relate the solutions of primal problems to the solutions of certain dual problems. For functions, a constraint qualification is needed; however, for monotone operators, the ensuing duality requires no constraint qualification.

Fact 2.8 (Fenchel-Rockafellar duality [28, 30]) *Assume that $f, g \in \Gamma(\mathcal{H})$ and $L : \mathcal{H} \to \mathcal{H}$ is a continuous linear operator. Suppose there exists $x_0 \in \operatorname{dom} f \cap L^{-1}(\operatorname{dom} g)$ such that g is continuous at Lx_0. Then*

$$\inf_{x\in\mathcal{H}} \big(f(x) + g(Lx)\big) = -\min_{y^*\in\mathcal{H}} \big(f^*(-L^*y^*) + g^*(y^*)\big).$$

Furthermore, $\bar{x}$ is a minimizer for $f + g \circ L$ if and only if there exists $\bar{y}^* \in \mathcal{H}$ such that

$$-L^*\bar{y}^* \in \partial f(\bar{x}), \quad \bar{y}^* \in \partial g(L\bar{x}).$$

Fact 2.9 (Attouch-Théra duality [1]) *Let $A, B : \mathcal{H} \rightrightarrows 2^{\mathcal{H}}$ be maximal monotone operators. Let S be the solution set of the primal problem*

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

Let S^* be the solution set of the dual problem

$$\text{find } x^* \in \mathcal{H} \text{ such that } 0 \in A^{-1}x^* + \widetilde{B}(x^*). \tag{2}$$

Then

(i) $S = \big\{x \in \mathcal{H} \mid (\exists\, x^* \in S^*)\ x^* \in Ax \text{ and } -x^* \in Bx\big\}$.

(ii) $S^* = \big\{x^* \in \mathcal{H} \mid (\exists\, x \in S)\ x \in A^{-1}x^* \text{ and } -x \in \widetilde{B}(x^*)\big\}$.

Moreover, let S_1^* be the solution to the dual problem given by

$$\text{find } y^* \in \mathcal{H} \text{ such that } 0 \in \widetilde{A}(y^*) + B^{-1}(y^*). \tag{3}$$

Then $S_1^* = -S^*$. Consequently, up to a change of sign in the dual variable, the Attouch-Théra duals (2) and (3) have the same solutions.

We are grateful to a referee for pointing out another reference relevant to Attouch-Théra duality, namely [22].

The last result recorded in this section concerns basic properties of the resolvent average.

Fact 2.10 (resolvent average) *Let $A_1, A_2 : \mathcal{H} \rightrightarrows 2^{\mathcal{H}}$ be maximal monotone operators, let $\lambda_1, \lambda_2, \gamma > 0$ with $\lambda_1 + \lambda_2 = 1$, and set*

$$A := \frac{(\lambda_1 J_{\gamma A_1} + \lambda_2 J_{\gamma A_2})^{-1} - \operatorname{Id}}{\gamma}.$$

Then

(i) $J_{\gamma A} = \lambda_1 J_{\gamma A_1} + \lambda_2 J_{\gamma A_2}$ *and* ${}^{\gamma}A = \lambda_1\, {}^{\gamma}A_1 + \lambda_2\, {}^{\gamma}A_2$.

(ii) *A is maximal monotone.*

Proof. (i) follows from the definitions involved. (ii): By Fact 2.2(iii) and maximal monotonicity of A_i, $J_{\gamma A_i}$ is firmly nonexpansive and has a full domain so that $J_{\gamma A}$ is firmly nonexpansive and has a full domain. Then by Fact 2.2(iii) again γA is maximal monotone, so is A. ∎

3 Fixed points of resolvent average and compositions

In this Section, we assume that $A_1, A_2 : \mathcal{H} \rightrightarrows 2^{\mathcal{H}}$ are maximal monotone operators, and that $\lambda_1 + \lambda_2 = 1$ where each $\lambda_i > 0$.

3.1 Inclusion problem formulations and their common Attouch-Théra dual

Consider the inclusion problems

$$(R_\gamma) \quad \text{find } z \text{ such that} \quad 0 \in \Big(\big(\lambda_1 J_{\gamma A_1} + \lambda_2 J_{\gamma A_2}\big)^{-1} - \operatorname{Id} \Big)(z); \tag{4}$$

$$(P_\gamma) \quad \text{find } z \text{ such that} \quad 0 = \big(\lambda_1 {}^{\gamma}A_1 + \lambda_2 {}^{\gamma}A_2\big)(z); \tag{5}$$

$$(P) \quad \text{find } (x,y) \text{ such that} \quad (0,0) \in \left(\frac{(\operatorname{Id} - R)}{\gamma} + \left(\frac{A_1}{\lambda_2} \times \frac{A_2}{\lambda_1} \right) \right)(x,y); \tag{6}$$

$$\text{find } x \text{ such that} \quad 0 \in \left(\frac{A_1}{\lambda_2} + {}^{\gamma}\!\left(\frac{A_2}{\lambda_1} \right) \right)(x); \tag{7}$$

$$\text{find } y \text{ such that} \quad 0 \in \left({}^{\gamma}\!\left(\frac{A_1}{\lambda_2} \right) + \frac{A_2}{\lambda_1} \right)(y). \tag{8}$$

Theorem 3.1 *Problems* (4)–(8) *either all have solutions or none admits a solution.*

Proof. (4)⇔(5): z solves (4) if and only if z solves $0 \in A(z)$ where

$$A = \frac{\big(\lambda_1 J_{\gamma A_1} + \lambda_2 J_{\gamma A_2}\big)^{-1} - \operatorname{Id}}{\gamma}.$$

It suffices to apply Fact 2.2(i) and Fact 2.10(i) to A.

(4)⇔(6): Note that z solves (4) if and only if $z = \lambda_1 J_{\gamma A_1}(z) + \lambda_2 J_{\gamma A_2}(z)$. Let $J_{\gamma A_1}(z) = x$, $J_{\gamma A_2}(z) = y$. We have z solves (4) if and only if

$$\begin{cases} z = \lambda_1 x + \lambda_2 y \\ x = J_{\gamma A_1}(z) \\ y = J_{\gamma A_2}(z). \end{cases} \tag{9}$$

We claim that (x,y) solves (6). Indeed, (9) gives $z \in \gamma A_1(x) + x$, $z \in \gamma A_2(y) + y$, i.e.,

$$0 \in \gamma A_1(x) + (x - z) \tag{10}$$

$$0 \in \gamma A_2(y) + (y - z). \tag{11}$$

As $z = \lambda_1 x + \lambda_2 y$, we have $x - z = \lambda_2(x - y)$, $y - z = \lambda_1(y - x)$. Therefore, (10) gives

$$0 \in \gamma A_1(x) + \lambda_2(x - y) \tag{12}$$

$$0 \in \gamma A_2(y) + \lambda_1(y - x), \tag{13}$$

equivalently,

$$0 \in \frac{\gamma A_1(x)}{\lambda_2} + (x - y) \tag{14}$$

$$0 \in \frac{\gamma A_2(y)}{\lambda_1} + (y - x). \tag{15}$$

In the product space setting,

$$(0,0) \in \left((\mathrm{Id} - R)(x,y) + \frac{\gamma A_1(x)}{\lambda_2} \times \frac{\gamma A_2(y)}{\lambda_1} \right). \tag{16}$$

Dividing both sides by γ gives

$$(0,0) \in \left(\frac{(\mathrm{Id} - R)}{\gamma}(x,y) + \frac{A_1(x)}{\lambda_2} \times \frac{A_2(y)}{\lambda_1} \right), \tag{17}$$

as required. Conversely, let (x,y) solves (6). Put $z = \lambda_1 x + \lambda_2 y$. Exactly reverse the arguments from (17) to (10) to get (9). Hence z solves (4).

(6)⇔(7): (x,y) solves (6) if and only if

$$0 \in \frac{\gamma A_1(x)}{\lambda_2} + (x - y) \tag{18}$$

$$0 \in \frac{\gamma A_2(y)}{\lambda_1} + (y - x). \tag{19}$$

From (19), $y = J_{\gamma A_2/\lambda_1}(x)$. Put this in (18) to get

$$0 \in \frac{\gamma A_1(x)}{\lambda_2} + x - J_{\gamma A_2/\lambda_1}(x). \tag{20}$$

Dividing both sides by γ gives

$$0 \in \frac{A_1(x)}{\lambda_2} + \frac{x - J_{\gamma A_2/\lambda_1}(x)}{\gamma} = \frac{A_1(x)}{\lambda_2} + {}^{\gamma}\!\left(\frac{A_2}{\lambda_1}\right)(x),$$

which says that x solves (7). Conversely, x solves (7) if and only if (20) holds. Put $y = J_{\gamma A_2/\lambda_1}(x)$. Then $x \in \gamma A_2(y)/\lambda_1 + y$, and (20) gives $0 \in \frac{\gamma A_1(x)}{\lambda_2} + x - y$. Hence (x,y) satisfies (18) and (19).

As in (6)⇔(7), similarly one can show (6)⇔(8). ■

We proceed to show that all of them share one common Attouch-Théra dual problem.

Theorem 3.2 *Up to a change of sign, the following inclusion problems have the same Attouch-Théra dual solution.*

(i) *(P_γ) find z such that* $0 \in \left(\frac{{}^{\gamma}A_1}{\lambda_2} + \frac{{}^{\gamma}A_2}{\lambda_1} \right)(z)$;

(ii) *(P)* *find* (x,y) *such that* $(0,0)\in\Big(\frac{(\operatorname{Id}-R)}{\gamma}+\big(\frac{A_1}{\lambda_2}\times\frac{A_2}{\lambda_1}\big)\Big)(x,y)$;

(iii) *find* x *such that* $0\in\Big(\frac{A_1}{\lambda_2}+{}^{\gamma}\big(\frac{A_2}{\lambda_1}\big)\Big)(x)$;

(iv) *find* y *such that* $0\in\Big({}^{\gamma}\big(\frac{A_1}{\lambda_2}\big)+\frac{A_2}{\lambda_1}\Big)(y)$.

Namely, up to a change of sign in the dual variable, their Attouch-Théra dual has the form

$$\text{(21)}\qquad \textit{find } z^* \textit{ such that}\quad 0\in\gamma z^*+(A_1/\lambda_2)^{-1}(z^*)+\widetilde{A_2/\lambda_1}(z^*).$$

Moreover, the set of solutions is either empty or a singleton.

Proof. By (1), we have

$$\widetilde{{}^{\gamma}A}=-({}^{\gamma}A)^{-1}(-\operatorname{Id})=\gamma\operatorname{Id}+\widetilde{A},$$
$$\widetilde{A/\lambda}=\widetilde{A}(\lambda\operatorname{Id}).$$

(i). The Attouch-Théra dual is:

$$0\in\Big[({}^{\gamma}A_1/\lambda_2)^{-1}+\widetilde{{}^{\gamma}A_2/\lambda_1}\Big](z^*).$$

We have

$$\begin{aligned}\big({}^{\gamma}A_1/\lambda_2\big)^{-1}+\widetilde{{}^{\gamma}A_2/\lambda_1}&=({}^{\gamma}A_1)^{-1}(\lambda_2\operatorname{Id})+\widetilde{{}^{\gamma}A_2}(\lambda_1\operatorname{Id})\\&=(\gamma\operatorname{Id}+A_1^{-1})(\lambda_2\operatorname{Id})+(\gamma\operatorname{Id}+\widetilde{A_2})(\lambda_1\operatorname{Id})\\&=\gamma\operatorname{Id}+A_1^{-1}(\lambda_2\operatorname{Id})+\widetilde{A_2}(\lambda_1\operatorname{Id})\\&=\gamma\operatorname{Id}+(A_1/\lambda_2)^{-1}+\widetilde{A_2/\lambda_1}.\end{aligned}$$

Hence the dual is

$$\text{(22)}\qquad 0\in[\gamma\operatorname{Id}+(A_1/\lambda_2)^{-1}+\widetilde{A_2/\lambda_1}](z^*).$$

(ii). The Attouch-Théra dual is

$$\text{(23)}\qquad (0,0)\in\left[\gamma(\operatorname{Id}-R)^{-1}+\left(\frac{A_1}{\lambda_2}\times\frac{A_2}{\lambda_1}\right)^{-1}\right](x^*,y^*).$$

Since $\operatorname{ran}(\operatorname{Id}-R)=\big\{(d,-d)\mid d\in\mathcal{H}\big\}$, we have $y^*=-x^*$. (23) reduces to find x^* such that

$$(0,0)\in\left[\gamma(\operatorname{Id}-R)^{-1}+\left(\frac{A_1}{\lambda_2}\times\frac{A_2}{\lambda_1}\right)^{-1}\right](x^*,-x^*).$$

Then x^* solves the dual if and only if there exists $y\in\mathcal{H}$ such that

$$(0,0)\in\gamma(y+x^*,y)+\left(\frac{A_1}{\lambda_2}\right)^{-1}(x^*)\times\left(\frac{A_2}{\lambda_1}\right)^{-1}(-x^*),$$

which transpires to

$$0 \in \gamma(y + x^*) + \left(\frac{A_1}{\lambda_2}\right)^{-1}(x^*), \quad 0 \in \gamma y + \left(\frac{A_2}{\lambda_1}\right)^{-1}(-x^*).$$

This is equivalent to find x^* such that

$$0 \in \gamma x^* + \left(\frac{A_1}{\lambda_2}\right)^{-1}(x^*) - \left(\frac{A_2}{\lambda_1}\right)^{-1}(-x^*) = \left[\gamma \operatorname{Id} + (A_1/\lambda_2)^{-1} + \widetilde{A_2/\lambda_1}\right](x^*).$$

(iii). The Attouch-Théra dual is

$$0 \in \left(\frac{A_1}{\lambda_2}\right)^{-1}(x^*) + \widetilde{{}^{\gamma}\left(\frac{A_2}{\lambda_1}\right)}(x^*).$$

The right-hand side becomes

$$\left(\frac{A_1}{\lambda_2}\right)^{-1} + \gamma \operatorname{Id} + \widetilde{\left(\frac{A_2}{\lambda_1}\right)} = \gamma \operatorname{Id} + (A_1/\lambda_2)^{-1} + \widetilde{A_2/\lambda_1}$$

Hence the dual is

$$0 \in \left[\gamma \operatorname{Id} + (A_1/\lambda_2)^{-1} + \widetilde{A_2/\lambda_1}\right](x^*).$$

(iv). The Attouch-Théra dual is

$$0 \in \widetilde{{}^{\gamma}(A_1/\lambda_2)}(y^*) + (A_2/\lambda_1)^{-1}(y^*).$$

We have

$$\widetilde{{}^{\gamma}(A_1/\lambda_2)} + (A_2/\lambda_1)^{-1} = \gamma \operatorname{Id} + \widetilde{A_1/\lambda_2} + (A_2/\lambda_1)^{-1}.$$

Then the dual becomes

$$0 \in \gamma y^* + \widetilde{A_1/\lambda_2}(y^*) + (A_2/\lambda_1)^{-1}(y^*),$$

that is,

$$0 \in \gamma y^* - (A_1/\lambda_2)^{-1}(-y^*) + (A_2/\lambda_1)^{-1}(y^*).$$

Multiplying both sides by -1, followed by making the substitution $z^* = -y^*$, we obtain

$$0 \in \gamma z^* + (A_1/\lambda_2)^{-1}(z^*) + \widetilde{A_2/\lambda_1}(z^*).$$

The proof is complete. ■

3.2 Characterization of solution sets

Problem (6) has its Attouch-Théra dual given by

$$\text{(24)} \qquad \text{(D)} \quad \text{find } (x^*, y^*) \text{ such that} \quad (0,0) \in \left(\left(\frac{\operatorname{Id} - R}{\gamma} \right)^{-1} + \left(\frac{A_1}{\lambda_2} \times \frac{A_2}{\lambda_1} \right)^{-1} \right)(x^*, y^*).$$

The following result gives a fixed point characterization to the solution sets of (4)–(8) when $\gamma = 1$.

Theorem 3.3 *The following assertions hold.*

(i) (**Fixed points of resolvent average**) *Let* $A = (\lambda_1 J_{A_1} + \lambda_2 J_{A_2})^{-1} - \operatorname{Id}$. *Then*

$$\operatorname{Fix} J_A = \big(\lambda_1 {}^1 A_1 + \lambda_2 {}^1 A_2\big)^{-1}(0) = \big\{ z \in \mathcal{H} \mid z = J_A(z) = \lambda_1 J_{A_1}(z) + \lambda_2 J_{A_2}(z) \big\}.$$

(ii) (**Fixed points of compositions**) *Set* $E := \Big(\frac{A_1}{\lambda_2} + {}^1\big(\frac{A_2}{\lambda_1}\big) \Big)^{-1}(0)$. *Then* $E = \operatorname{Fix} J_{A_1/\lambda_2} J_{A_2/\lambda_1} = J_{A_1/\lambda_2}(F)$.

(iii) (**Fixed points of compositions**) *Set* $F := \Big({}^1\big(\frac{A_1}{\lambda_2}\big) + \frac{A_2}{\lambda_1} \Big)^{-1}(0)$. *Then* $F = \operatorname{Fix} J_{A_2/\lambda_1} J_{A_1/\lambda_2} = J_{A_2/\lambda_1}(E)$.

(iv) (**Fixed points of alternating resolvents**) *Set* $S := \Big((\operatorname{Id} - R) + \big(\frac{A_1}{\lambda_2} \times \frac{A_2}{\lambda_1}\big) \Big)^{-1}(0,0)$. *Then*

$$S = \{(x,y) | x = J_{A_1/\lambda_2} y, y = J_{A_2/\lambda_1} x\} = \operatorname{Fix}\big(J_{A_1/\lambda_2 \times A_2/\lambda_1} \circ R\big) = (E \times F) \cap \operatorname{gr} J_{A_2/\lambda_1}.$$

(v) *Set* $S^* = \Big((\operatorname{Id} - R)^{-1} + \big(\frac{A_1}{\lambda_2} \times \frac{A_2}{\lambda_1}\big)^{-1} \Big)^{-1}(0,0)$. *Then* S^* *is at most a singleton with*

$$S^* = \big\{ (u^*, v^*) \mid u^* = J_{(A_1/\lambda_2)^{-1} + \widetilde{A_2/\lambda_1}}(0), v^* = J_{\widetilde{A_1/\lambda_2} + (A_2/\lambda_1)^{-1}}(0) \big\}.$$

Moreover, $u^* = -v^*$. *(Note that* S^* *may be empty, which is equivalent to the impossibility to compute the resolvents defining* u^* *and* v^*.*)*

(vi) $S^* = (R - \operatorname{Id})(S)$. *Consequently, for every* $(x,y) \in S$, $y - x = u^*$, *i.e., the gap vector is unique.*

(vii) $E = (A_1/\lambda_2)^{-1}(u^*) \cap \Big({}^1\big(A_2/\lambda_1\big)\Big)^{-1}(v^*)$ *and* $F = \Big({}^1\big(A_1/\lambda_2\big)\Big)^{-1}(u^*) \cap (A_2/\lambda_1)^{-1}(v^*)$.

(viii) $J_{A_2/\lambda_1}|_E : E \to F : x \mapsto x + u^*$ *is a bijection with inverse mapping* $J_{A_1/\lambda_2} : F \to E : y \mapsto y + v^*$.

(ix)

$$\text{(25)} \qquad S = (E \times F) \cap (R - \operatorname{Id})^{-1}(u^*, v^*)$$

$$\text{(26)} \qquad \quad = \left(\frac{A_1}{\lambda_2} \times \frac{A_2}{\lambda_1} \right)^{-1}(u^*, v^*) \cap (R - \operatorname{Id})^{-1}(u^*, v^*).$$

(x)

$$\operatorname{Fix} J_A = \left(\frac{{}^1A_1}{\lambda_2}\right)^{-1}(u^*) \bigcap \left(\frac{{}^1A_2}{\lambda_1}\right)^{-1}(v^*).$$

(xi) *The sets* $\operatorname{Fix}(J_A), E, F, S$ *are closed and convex.*

Proof. (i). $z \in \left(\lambda_1 {}^1A_1 + \lambda_2 {}^1A_2\right)^{-1}(0)$ if and only if

$$\begin{aligned} 0 &= \left(\lambda_1 {}^1A_1 + \lambda_2 {}^1A_2\right)(z) = \lambda_1(\operatorname{Id} - J_{A_1})(z) + \lambda_2(\operatorname{Id} - J_{A_2})(z) \\ &= z - (\lambda_1 J_{A_1} + \lambda_2 J_{A_2})(z) = z - J_A(z). \end{aligned}$$

(ii)–equation (25) of (ix) follow by applying Fact 2.6 with $A = A_1/\lambda_2$, $B = A_2/\lambda_1$ and $\gamma = 1$. To show (26), we assume that S and S^* are nonempty. Note that

$$S = \left\{(x, y) : \ (0, 0) \in \left((\operatorname{Id} - R) + \left(\frac{A_1}{\lambda_2} \times \frac{A_2}{\lambda_1}\right)\right)(x, y)\right\}$$

and that

$$S^* = \left((\operatorname{Id} - R)^{-1} + \left(\frac{A_1}{\lambda_2} \times \frac{A_2}{\lambda_1}\right)^{-1}\right)^{-1}(0, 0),$$

so we can use Fact 2.9(i) to get

$$S = \{(x, y) : \ (\exists (u^*, v^*) \in S^*) \ (u^*, v^*) \in \left(\frac{A_1}{\lambda_2} \times \frac{A_2}{\lambda_1}\right)(x, y), -(u^*, v^*) \in (\operatorname{Id} - R)(x, y)\}.$$

By (v), S^* is singleton so that $S^* = \{(u^*, v^*)\}$. Hence (26) holds.

(x). By (i), $z \in \operatorname{Fix} J_A \Leftrightarrow 0 \in \left(\frac{{}^1A_1}{\lambda_2} + \frac{{}^1A_2}{\lambda_1}\right)(z)$. The latter has its Attouch-Thera dual given by

$$0 \in \left(\frac{{}^1A_1}{\lambda_2}\right)^{-1}(x^*) + \widetilde{\left(\frac{{}^1A_2}{\lambda_2}\right)}(x^*),$$

equivalently by (22) (with $\gamma = 1$)

$$0 \in [\operatorname{Id} + (A_1/\lambda_2)^{-1} + \widetilde{A_2/\lambda_1}](z^*),$$

and it has a unique solution u^* by (v). Fact 2.9(i) gives that $z \in \operatorname{Fix} J_A$ if and only if

$$u^* \in \left(\frac{{}^1A_1}{\lambda_2}\right)(z), \quad -u^* = v^* \in \left(\frac{{}^1A_2}{\lambda_1}\right)(z), \text{ i.e.,}$$

$$z \in \left(\frac{{}^1A_1}{\lambda_2}\right)^{-1}(u^*) \bigcap \left(\frac{{}^1A_2}{\lambda_1}\right)^{-1}(v^*).$$

(xi). It is well-known that if $B : \mathcal{H} \rightrightarrows 2^{\mathcal{H}}$ is maximal monotone, then $B(x)$ is closed and convex for every $x \in \mathcal{H}$. Observe that $\lambda_1 {}^1A_1 + \lambda_2 {}^1A_2$, $\frac{A_1}{\lambda_2} + {}^1\left(\frac{A_2}{\lambda_1}\right)$, ${}^1\left(\frac{A_1}{\lambda_2}\right) + \frac{A_2}{\lambda_1}$, and $(\operatorname{Id} - R) + \left(\frac{A_1}{\lambda_2} \times \frac{A_2}{\lambda_1}\right)$ are maximal monotone operators by Rockafellar's sum theorem, see [29, pages 104–105] or [25]. Then $\left(\lambda_1 {}^1A_1 + \lambda_2 {}^1A_2\right)^{-1}$, $\left(\frac{A_1}{\lambda_2} + {}^1\left(\frac{A_2}{\lambda_1}\right)\right)^{-1}$, $\left({}^1\left(\frac{A_1}{\lambda_2}\right) + \frac{A_2}{\lambda_1}\right)^{-1}$, $\left((\operatorname{Id} - R) + \left(\frac{A_1}{\lambda_2} \times \frac{A_2}{\lambda_1}\right)\right)^{-1}$ are maximal monotone. Hence the result holds by the definitions of these sets given in (i)–(iv). ■

3.3 Relationship among solution sets

Recall that

$$J_A = \lambda_1 J_{A_1} + \lambda_2 J_{A_2} \quad \text{with } \lambda_1 + \lambda_2 = 1 \text{ and } \lambda_i > 0. \tag{27}$$

We now study the relationships among $\operatorname{Fix} J_A$, $E = \operatorname{Fix}(J_{A_1/\lambda_2} J_{A_2/\lambda_1})$, $F = \operatorname{Fix}(J_{A_2/\lambda_1} J_{A_1/\lambda_2})$, and

$$S = \big\{(x, y) \mid x = J_{A_1/\lambda_2} y, y = J_{A_2/\lambda_1} x\big\}.$$

See also [13] for some variants for more than 2 operators.

Lemma 3.4 (i) *If* $x = J_{A_1/\lambda_2} y, \quad y = J_{A_2/\lambda_1} x$, *then*

$$\lambda_1 x + \lambda_2 y \in \operatorname{Fix} J_A, \quad x \in \operatorname{Fix} J_{A_1/\lambda_2} J_{A_2/\lambda_1}, \quad y \in \operatorname{Fix} J_{A_2/\lambda_1} J_{A_1/\lambda_2}.$$

(ii) *If* $x = J_{A_1/\lambda_2} J_{A_2/\lambda_1} x$, *put* $y = J_{A_2/\lambda_1} x$, *then* $\lambda_1 x + \lambda_2 y \in \operatorname{Fix} J_A$.

(iii) *If* $y = J_{A_2/\lambda_1} J_{A_1/\lambda_2} y$, *put* $x = J_{A_1/\lambda_2} y$, *then* $\lambda_1 x + \lambda_2 y \in \operatorname{Fix} J_A$.

Proof. (i). We have

$$x \in J_{A_1/\lambda_2} y \quad \Leftrightarrow \quad y \in \frac{A_1}{\lambda_2} x + x,$$

$$y \in J_{A_2/\lambda_1} x \quad \Leftrightarrow \quad x \in \frac{A_2}{\lambda_1} y + y,$$

so that $-\lambda_2 x + \lambda_2 y \in A_1 x$ and $\lambda_1 x - \lambda_1 y \in A_2 y$. Then $\lambda_1 x + \lambda_2 y \in A_1 x + x$, $\lambda_1 x + \lambda_2 y \in A_2 y + y$, equivalently $x = J_{A_1}(\lambda_1 x + \lambda_2 y), y = J_{A_2}(\lambda_1 x + \lambda_2 y)$. This gives

$$\lambda_1 x + \lambda_2 y = \lambda_1 J_{A_1}(\lambda_1 x + \lambda_2 y) + \lambda_2 J_{A_2}(\lambda_1 x + \lambda_2 y) = [\lambda_1 J_{A_1} + \lambda_2 J_{A_2}](\lambda_1 x + \lambda_2 y).$$

Hence $\lambda_1 x + \lambda_2 y \in \operatorname{Fix} J_A$.

(ii) and (iii): In either (ii) or (iii), we have $x = J_{A_1/\lambda_2} y, \quad y = J_{A_2/\lambda_1} x$. Hence (i) applies. ■

Lemma 3.5 *If* $x \in \operatorname{Fix} J_A$, *then*

$$J_{A_1} x = J_{A_1/\lambda_2}(J_{A_2} x), \tag{28}$$

$$J_{A_2} x = J_{A_2/\lambda_1}(J_{A_1} x). \tag{29}$$

Consequently, $J_{A_1} x \in \operatorname{Fix} J_{A_1/\lambda_2} J_{A_2/\lambda_1}$ *and* $J_{A_2} x \in \operatorname{Fix} J_{A_2/\lambda_1} J_{A_1/\lambda_2}$.

Proof. Let us show (28). By assumption, $x = \lambda_1 J_{A_1} x + \lambda_2 J_{A_2} x$, we write

$$J_{A_2} x = \frac{x - \lambda_1 J_{A_1} x}{\lambda_2}. \tag{30}$$

We have

$$\begin{aligned}
J_{A_1}x = J_{A_1}x \Leftrightarrow\ & x \in (A_1 + \mathrm{Id})(J_{A_1}x) = A_1(J_{A_1}x) + J_{A_1}x \\
\Leftrightarrow\ & x \in A_1(J_{A_1}x) + \lambda_1 J_{A_1}x + \lambda_2 J_{A_1}x \quad (\text{since } \lambda_1 + \lambda_2 = 1) \\
\Leftrightarrow\ & x - \lambda_1 J_{A_1}x \in A_1(J_{A_1}x) + \lambda_2 J_{A_1}x = (A_1 + \lambda_2 \mathrm{Id})(J_{A_1}x) \\
\Leftrightarrow\ & \frac{x - \lambda_1 J_{A_1}x}{\lambda_2} \in \big(A_1/\lambda_2 + \mathrm{Id}\big)(J_{A_1}x) \\
\Leftrightarrow\ & J_{A_2}x \in \big(A_1/\lambda_2 + \mathrm{Id}\big)(J_{A_1}x) \quad \text{(by (30))} \\
\Leftrightarrow\ & J_{A_1}x = J_{A_1/\lambda_2}(J_{A_2}x).
\end{aligned}$$

The proof of (29) is similar. ■

Note that

$$S = \big\{(x, y) \mid x = J_{A_1/\lambda_2}y, y = J_{A_2/\lambda_1}x\big\}.$$

Theorem 3.6 *Define*

$$T : S \to \mathrm{Fix}\, J_A : (x, y) \mapsto \lambda_1 x + \lambda_2 y.$$

Then T is a homeomorphism, and the inverse of T is given by

$$T^{-1} : \mathrm{Fix}\, J_A \to S : z \mapsto (J_{A_1}z, J_{A_2}z).$$

Consequently, $\mathrm{Fix}\, J_A = L(S)$ *where* $L : \mathcal{H} \times \mathcal{H} \to \mathcal{H} : (x, y) \mapsto \lambda_1 x + \lambda_2 y.$

Proof. For every $(x, y) \in S$, by Lemma 3.4(i), $T(x, y) \in \mathrm{Fix}\, J_A$, so $T(S) \subseteq \mathrm{Fix}\, J_A$. For every $z \in \mathrm{Fix}\, J_A$, by Lemma 3.5, $(J_{A_1}z, J_{A_2}z) \in S$ and $z = \lambda_1 J_{A_1}(z) + \lambda_2 J_{A_2}(z) = T(J_{A_1}z, J_{A_2}z)$, thus $T(S) \supseteq \mathrm{Fix}\, J_A$. Hence $T(S) = \mathrm{Fix}\, J_A$, i.e., T is onto. To show that T is one-to-one, let $(x_i, y_i) \in S$ for $i \in \{1, 2\}$. If $T(x_1, y_1) = T(x_2, y_2)$, i.e., $\lambda_1 x_1 + \lambda_2 y_1 = \lambda_1 x_2 + \lambda_2 y_2$, then

$$\lambda_1(x_1 - y_1) + y_1 = \lambda_1(x_2 - y_2) + y_2.$$

By Theorem 3.3(v) and (vi), S^* is unique and $x_1 - y_1 = x_2 - y_2 = v^*$, thus $y_1 = y_2$ and $x_1 = x_2$.

Since for $z \in \mathrm{Fix}\, J_A$, $(J_{A_1}(z), J_{A_2}(z)) \in S$ and $z = T(J_{A_1}(z), J_{A_2}(z))$, T is one-to-one and onto, we obtain that $T^{-1}(z) = (J_{A_1}(z), J_{A_2}(z))$. In addition, both T, T^{-1} are continuous. Hence $T : S \to \mathrm{Fix}\, J_A$ is a homeomorphism. ■

Theorem 3.7 (i) *The mapping*

$$T_1 : E \to S : x \mapsto (x, J_{A_2/\lambda_1}x),$$

is a homeomorphism and its inverse is given by

$$T_1^{-1} : S \to E : (x, y) \mapsto x.$$

(ii) *The mapping*

$$T_2 : F \to S : y \mapsto (J_{A_1/\lambda_2} y, y),$$

is a homeomorphism and its inverse is given by

$$T_2^{-1} : S \to F : (x, y) \mapsto y.$$

Proof. We only prove (i), since (ii) can be proved similarly. To see (i), let $x \in E$. By the definition of E, $x = J_{A_1/\lambda_2} J_{A_2/\lambda_1} x$. Put $y = J_{A_2/\lambda_1} x$. We have

$$x = J_{A_1/\lambda_2} y, \quad y = J_{A_2/\lambda_1} x,$$

whence $T_1 x = (x, y) \in S$. Therefore, $T_1(E) \subseteq S$. Now for every $(x, y) \in S$, by the definition of S,

$$x = J_{A_1/\lambda_2} y, \quad y = J_{A_2/\lambda_1} x,$$

then $x \in E$ and $(x, y) = (x, J_{A_2/\lambda_1} x) = T_1 x$. Therefore, $S \subseteq T_1(E)$. Hence $T_1(E) = S$. Clearly, T_1 is one-to-one. Altogether, T_1 is one-to-one and onto. Since for every $(x, y) \in S$,

$$(x, y) = (x, J_{A_2/\lambda_1} x) = T_1 x,$$

we have $T_1^{-1}(x, y) = x$. ■

The next result provides a partial answer to a question raised by C. Byrne (see [15, page 305]). It provides the transformations to go back and forth between fixed point sets of compositions of resolvents and the fixed point set of the average.

Theorem 3.8 *Let u^* be given as in Theorem 3.3(v).*

(i) *The mapping*

$$H_1 : E \to \operatorname{Fix} J_A : x \mapsto \lambda_1 x + \lambda_2 J_{A_2/\lambda_1} x = x + \lambda_2 u^*,$$

is a homeomorphism. Moreover, $H_1^{-1} : \operatorname{Fix} J_A \to E$ is given by $H_1^{-1}(z) = J_{A_1}(z)$. Hence

$$\operatorname{Fix} J_A = E + \lambda_2 u^*. \tag{31}$$

(ii) *The mapping*

$$H_2 : F \to \operatorname{Fix} J_A : y \mapsto \lambda_1 J_{A_1/\lambda_2} y + \lambda_2 y = -\lambda_1 u^* + y,$$

is a homeomorphism. Moreover, $H_2^{-1} : \operatorname{Fix} J_A \to F$ is given by $H_2^{-1}(z) = J_{A_2}(z)$. Hence

$$\operatorname{Fix} J_A = F - \lambda_1 u^*. \tag{32}$$

Proof. Combine Theorem 3.6 and Theorem 3.7. Using the same notations as in Theorem 3.6 and Theorem 3.7, (i) follows from $H_1 = T \circ T_1$; (ii) follows from $H_2 = T \circ T_2$. Moreover,

$$(\forall\ x \in E)\ H_1(x) = \lambda_1 x + \lambda_2 J_{A_2/\lambda_1} x = x + \lambda_2 (J_{A_2/\lambda_1} x - x) = x + \lambda_2 u^*,$$

$$(\forall\ y \in F)\ H_2(y) = \lambda_1 J_{A_1/\lambda_2} y + \lambda_2 y = \lambda_1 (J_{A_1/\lambda_2} y - y) + y = \lambda_1(-u^*) + y,$$

by Theorem 3.3(vi). Hence (31) and (32) hold. ■

Corollary 3.9 *The following is true.*

(i) $E \neq \varnothing \Leftrightarrow F \neq \varnothing \Leftrightarrow S \neq \varnothing \Leftrightarrow S^* \neq \varnothing \Leftrightarrow \operatorname{Fix} J_A \neq \varnothing$.

(ii) E *is a singleton* $\Leftrightarrow$ F *is a singleton* $\Leftrightarrow$ S *is a singleton* $\Leftrightarrow$ $\operatorname{Fix} J_A$ *is a singleton.*

(iii) $\operatorname{Fix} J_A = \lambda_1 E + \lambda_2 F$.

Proof. (i) and (ii) follow from Theorems 3.6, 3.7 and Theorem 3.8. It remains to prove (iii). By Theorem 3.8,

$$\operatorname{Fix} J_A = E + \lambda_2 u^*, \quad \operatorname{Fix} J_A = F - \lambda_1 u^*.$$

As $\operatorname{Fix} J_A$ is convex by Theorem 3.3(xi), we obtain

$$\operatorname{Fix} J_A = \lambda_1 \operatorname{Fix} J_A + \lambda_2 \operatorname{Fix} J_A = \lambda_1(E + \lambda_2 u^*) + \lambda_2(F - \lambda_1 u^*) = \lambda_1 E + \lambda_2 F,$$

as claimed. ■

3.4 The case when $\operatorname{Fix} J_{A_1/\lambda_2} \cap \operatorname{Fix} J_{A_2/\lambda_1} \neq \varnothing$

Note that

$$\operatorname{Fix} J_{A_1} = \operatorname{Fix} J_{A_1/\lambda_2}, \quad \operatorname{Fix} J_{A_2} = \operatorname{Fix} J_{A_2/\lambda_1}. \tag{33}$$

Theorem 3.10 *Assume that* $\operatorname{Fix} J_{A_1/\lambda_2} \cap \operatorname{Fix} J_{A_2/\lambda_1} \neq \varnothing$. *Let* $\lambda_1 + \lambda_2 = 1$ *with* $\lambda_i > 0$. *Then*

$$\begin{aligned}\operatorname{Fix}(\lambda_1 J_{A_1} + \lambda_2 J_{A_2}) &= \operatorname{Fix} J_{A_1} \cap \operatorname{Fix} J_{A_2} = \operatorname{Fix} J_{A_1/\lambda_2} \cap \operatorname{Fix} J_{A_2/\lambda_1} \\ &= \operatorname{Fix}(J_{A_1/\lambda_2} \circ J_{A_2/\lambda_1}) = \operatorname{Fix}(J_{A_2/\lambda_1} \circ J_{A_1/\lambda_2}).\end{aligned}$$

Proof. In view of (33), $\operatorname{Fix} J_{A_1} \cap \operatorname{Fix} J_{A_2} \neq \varnothing$. Since every resolvent is attracting, it suffices to apply Fact 2.7. ■

4 Minimizers of the proximal average

We now specialize our results to $A_1 = \partial f_1$ and $A_2 = \partial f_2$ for two proper lower semicontinuous convex functions f_1, f_2. This allows us to understand the results of Section 3 from the variational analysis perspective. Let $f_1, f_2 \in \Gamma(\mathcal{H})$ and $\lambda_1 + \lambda_2 = 1$ with each $\lambda_i > 0$.

4.1 Minimization problem formulations and their common Fenchel-Rockafellar dual

We consider

$$\min_{(x,y)} g(x,y) := \left(\frac{f_1(x)}{\lambda_2} + \frac{f_2(y)}{\lambda_1} + \frac{\|x-y\|^2}{2}\right). \tag{34}$$

This turns out to be closely related to the proximal average of f_1, f_2, recently studied in [8, 9, 12, 11]. Recall

$$(\forall z \in X) \quad p_\gamma(\boldsymbol{f}, \boldsymbol{\lambda})(z) = \inf_{z=\lambda_1 x + \lambda_2 y} \left(\lambda_1 f_1(x) + \lambda_2 f_2(y) + \tfrac{\lambda_1\lambda_2}{2\gamma}\|x-y\|^2\right), \tag{35}$$

where $\boldsymbol{f} = (f_1, f_2), \boldsymbol{\lambda} = (\lambda_1, \lambda_2), \gamma > 0$. When $\gamma = 1$, we just write $p(\boldsymbol{f}, \boldsymbol{\lambda})$. Therefore, (34) has the same minimum value as the scaled proximal average

$$\min_z \frac{p(\boldsymbol{f}, \boldsymbol{\lambda})(z)}{\lambda_1\lambda_2}. \tag{36}$$

In terms of Moreau envelopes, (34) can be reformulated as

$$\min_x g_1(x) := f_1(x)/\lambda_2 + e_1(f_2/\lambda_1)(x) \tag{37}$$

and

$$\min_y g_2(y) := e_1(f_1/\lambda_2)(y) + f_2(y)/\lambda_1. \tag{38}$$

With regard to (36), we also consider

$$\min_z g_3(z) := \lambda_1 e_1 f_1(z) + \lambda_2 e_1 f_2(z) = \min_z \lambda_1\lambda_2\left(\frac{e_1 f_1(z)}{\lambda_2} + \frac{e_1 f_2(z)}{\lambda_1}\right). \tag{39}$$

The following facts about proximal average will be useful.

Fact 4.1 (i) (See [8, Theorem 4.10].) *For every $z \in \operatorname{dom} p_\gamma(\boldsymbol{f}, \boldsymbol{\lambda})$, there exist $x \in \operatorname{dom} f_1, y \in \operatorname{dom} f_2$ such that $z = \lambda_1 x + \lambda_2 y$ and*

$$p_\gamma(\boldsymbol{f}, \boldsymbol{\lambda})(z) = \lambda_1 f_1(x) + \lambda_2 f_2(y) + \frac{\lambda_1\lambda_2\|x-y\|^2}{2\gamma}.$$

(ii) (See [8, Theorem 6.2].) $e_\gamma p_\gamma(\boldsymbol{f}, \boldsymbol{\lambda}) = \lambda_1 e_\gamma f_1 + \lambda_2 e_\gamma f_2$.

(iii) (See [8, Theorem 6.7].) $\operatorname{Prox}_{p(\boldsymbol{f},\boldsymbol{\lambda})} = \lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2}$.

Fact 4.2 *Let $f \in \Gamma(\mathcal{H})$ and $\gamma \in (0, +\infty)$. Then*

(i) *$e_\gamma f$ is Fréchet differentiable on $\mathcal{H}$ and* $\nabla(e_\gamma f) = {}^{\gamma}(\partial f) = (\operatorname{Id} - \operatorname{Prox}_{\gamma f})/\gamma$.

(ii) $\inf(e_\gamma f) = \inf f$ *and* $\operatorname{argmin}(e_\gamma f) = \operatorname{argmin} f$.

Proof. See [19, Lemma 2.5] and [24]. See also [28, Example 1.46 and Theorem 2.26]. ■

Proposition 4.3 *For the minimization problems given by* (34)–(39)*, we have*

$$\min_{(x,y)} g(x,y) = \min_z \frac{p(\boldsymbol{f},\boldsymbol{\lambda})(z)}{\lambda_1\lambda_2} = \min_x g_1(x) = \min_y g_2(y) = \min_z \frac{g_3(z)}{\lambda_1\lambda_2}.$$

Proof. While the first three equalities are immediate, the fourth one follows from Fact 4.1(ii) and Fact 4.2(ii). ■

The following result is a convex-function refinement of Theorem 3.2. It says that convex optimization problems (34), (37), (38), (39) share one common Fenchel-Rockafellar dual problem.

Theorem 4.4 *Up to a change of sign, the following problems have the same Fenchel dual.*

(i)

$$\min_{(x,y)} \left(\frac{f_1(x)}{\lambda_2} + \frac{f_2(y)}{\lambda_1} + \frac{\|x-y\|^2}{2}\right).$$

(ii)

$$\min_x \; f_1(x)/\lambda_2 + e_1(f_2/\lambda_1)(x).$$

(iii)

$$\min_y \; e_1(f_1/\lambda_2)(y) + f_2(y)/\lambda_1.$$

(iv)

$$\min_z \; \frac{e_1 f_1(z)}{\lambda_2} + \frac{e_1 f_2(z)}{\lambda_1}.$$

Namely, up to a change of sign in the dual variable, their Fenchel-Rockafellar dual is given by

$$(40) \qquad (D) \quad \max_\phi \left[-\left(\frac{f_1}{\lambda_2}\right)^*(-\phi) - \left(\frac{f_2}{\lambda_1}\right)^*(\phi) - \frac{\|\phi\|^2}{2}\right] = -\min_\phi \left[\left(\frac{f_1}{\lambda_2}\right)^*(-\phi) + \left(\frac{f_2}{\lambda_1}\right)^*(\phi) + \frac{\|\phi\|^2}{2}\right].$$

Proof. (i). Using the Fenchel-Rockafellar Duality theorem (Fact 2.8) for $f_1/\lambda_2 \oplus f_2/\lambda_1$, $j \circ L$ with $j = \|\cdot\|^2/2$ and $L = (\mathrm{Id}, -\mathrm{Id}) : \mathcal{H}\times\mathcal{H} \to \mathcal{H}$, we obtain the dual problem (40).

(ii). The Fenchel dual is given by

$$\sup_\phi -(f_1/\lambda_2)^*(-\phi) - [e_1(f_2/\lambda_1)]^*(\phi).$$

As $[e_1(f_2/\lambda_1)]^* = (f_2/\lambda_1)^* + j$, its Fenchel dual becomes

$$\sup_\phi -(f_1/\lambda_2)^*(-\phi) - (f_2/\lambda_1)^*(\phi) - \|\phi\|^2/2.$$

(iii). The Fenchel dual is

$$\begin{aligned}
&\sup_\phi -[e_1(f_1/\lambda_2)]^*(-\phi) - (f_2/\lambda_1)^*(\phi) \\
&= \sup_\phi -(f_1/\lambda_2)^*(-\phi) - \|-\phi\|^2/2 - (f_2/\lambda_1)^*(\phi) \\
&= \sup_\phi -(f_1/\lambda_2)^*(-\phi) - (f_2/\lambda_1)^*(\phi) - \|\phi\|^2/2.
\end{aligned}$$

(iv). Its Fenchel dual is

$$\sup_\phi -\left(\frac{e_1 f_1}{\lambda_2}\right)^*(-\phi) - \left(\frac{e_1 f_2}{\lambda_1}\right)^*(\phi). \tag{41}$$

Now

$$\begin{aligned}
\left(\frac{e_1 f_1}{\lambda_2}\right)^* &= \frac{1}{\lambda_2}(e_1 f_1)^*(\lambda_2 \cdot) = \frac{1}{\lambda_2}(f_1^* + j)(\lambda_2 \cdot) \\
&= \frac{1}{\lambda_2} f_1^*(\lambda_2 \cdot) + \lambda_2 j = \left(\frac{f_1}{\lambda_2}\right)^* + \lambda_2 j.
\end{aligned}$$

Similarly,

$$\left(\frac{e_1 f_2}{\lambda_1}\right)^* = \left(\frac{f_2}{\lambda_1}\right)^* + \lambda_1 j.$$

Then (41) becomes

$$\begin{aligned}
&\sup_\phi -\left(\frac{f_1}{\lambda_2}\right)^*(-\phi) - \lambda_2 j(-\phi) - \left(\frac{f_2}{\lambda_1}\right)^*(\phi) - \lambda_1 j(\phi) \\
&= \sup_\phi -\left(\frac{f_1}{\lambda_2}\right)^*(-\phi) - \left(\frac{f_2}{\lambda_1}\right)^*(\phi) - \frac{\|\phi\|^2}{2}.
\end{aligned}$$

The proof is complete. ■

When the primal problem (34) has a finite infimum value, the primal optimal value and the dual optimal value are equal; moreover, the dual optimal value is attained.

While the solution set of the primal problem (34) (see Theorem 4.7(i)) may be empty, the solution set of the dual problem (40) is nonempty and a singleton as long as (34) has a finite infimum value. This feature of Fenchel-Rockafellar duality is in stark contrast to the Attouch-Théra duality of Section 3; see Corollary 3.9(i).

Theorem 4.5 *When the primal problem* (34) *has a finite infimum, the dual (D) has a unique solution*

$$\bar{\phi} = \mathrm{Prox}_{(f_1/\lambda_2)^* \circ (-\mathrm{Id}) + (f_2/\lambda_1)^*}(0). \tag{42}$$

If

(43) $$\operatorname{dom}(f_1/\lambda_2)^* \cap -\operatorname{int}\operatorname{dom}(f_2/\lambda_2)^* \neq \varnothing \quad or \quad \operatorname{int}\operatorname{dom}(f_1/\lambda_2)^* \cap -\operatorname{dom}(f_2/\lambda_2)^* \neq \varnothing,$$

then

(44) $$\bar{\phi} = J_{\widetilde{\partial f_1/\lambda_2}+(\partial f_2/\lambda_1)^{-1}}(0) = -J_{(\partial f_1/\lambda_2)^{-1}+\widetilde{\partial f_2/\lambda_1}}(0).$$

Proof. From (40), we have

$$0 \in \partial\big[(f_1/\lambda_2)^* \circ (-\operatorname{Id}) + (f_2/\lambda_1)^*\big](\bar{\phi}) + \bar{\phi},$$

so (42) holds.

Under the assumption (43), we can apply the chain rule so that

$$0 \in -(\partial f_1/\lambda_2)^{-1}(-\bar{\phi}) + (\partial f_2/\lambda_1)^{-1}(\bar{\phi}) + \bar{\phi} = \widetilde{\partial f_1/\lambda_2}(\bar{\phi}) + (\partial f_2/\lambda_1)^{-1}(\bar{\phi}) + \bar{\phi}.$$

Hence the first equality in (44) holds. Rewrite the dual problem (40) as

$$-\inf_{\psi}\left[\left(\frac{f_1}{\lambda_2}\right)^*(\psi) + \left(\frac{f_2}{\lambda_1}\right)^*(-\psi) + \frac{\|\psi\|^2}{2}\right]$$

and denote its optimal solution by $\bar{\psi}$. Then $-\bar{\psi} = \bar{\phi}$ and $\bar{\psi} = J_{(\partial f_1/\lambda_2)^{-1}+\widetilde{\partial f_2/\lambda_1}}(0)$. Therefore, the second equality in (44) holds also. ∎

Remark 4.6 Note that [7, Proposition 4.3] also implies (42) as well as

$$\bar{\phi} = -\operatorname{Prox}_{[(f_2/\lambda_1)\Box(f_1/\lambda_2\circ(-\operatorname{Id}))]^{**}}(0).$$

Observe that $\bar{\phi} = v^*$ as given in Theorem 3.3(v).

4.2 Characterization of minimizers

Set

$$\begin{aligned} S &:= \big\{(x,y) \mid x = \operatorname{Prox}_{f_1/\lambda_2} y,\ y = \operatorname{Prox}_{f_2/\lambda_1} x\big\},\\ E &:= \operatorname{Fix}(\operatorname{Prox}_{f_1/\lambda_2}\operatorname{Prox}_{f_2/\lambda_1}),\\ F &:= \operatorname{Fix}(\operatorname{Prox}_{f_2/\lambda_1}\operatorname{Prox}_{f_1/\lambda_2}). \end{aligned}$$

Theorem 4.7 *The following assertions hold.*

(i) (**Fixed points of alternating proximal mappings**) $S = \operatorname{argmin} g \subseteq E \times F$.

(ii) (**Fixed points of proximal mapping composition**) $E = \operatorname{argmin} g_1$.

(iii) (**Fixed points of proximal mapping composition**) $F = \operatorname{argmin} g_2$.

(iv) $\operatorname{argmin} p(\boldsymbol{f}, \boldsymbol{\lambda}) = \operatorname{argmin} g_3 = \left[\lambda_1\, {}^{1}(\partial f_1) + \lambda_2\, {}^{1}(\partial f_2)\right]^{-1}(0)$.

(v) (**Fixed points of the average of proximal mappings**)

$$\operatorname{argmin} p(\boldsymbol{f}, \boldsymbol{\lambda}) = \left\{\lambda_1 x + \lambda_2 y \mid (x, y) \in S\right\} \tag{45}$$

$$= \operatorname{Fix}(\operatorname{Prox}_{p(\boldsymbol{f},\boldsymbol{\lambda})}) = \operatorname{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2}). \tag{46}$$

(vi) *The sets* $S, E, F, \operatorname{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2})$ *are closed and convex.*

Proof. We have

$$\begin{aligned} \partial g(x, y) &= \left(\frac{\partial f_1(x)}{\lambda_2} + x - y, \frac{\partial f_2(y)}{\lambda_1} + y - x\right) \\ &= \left(\frac{\partial f_1}{\lambda_2} \times \frac{\partial f_2}{\lambda_1} + (\operatorname{Id} - R)\right)(x, y). \end{aligned}$$

Moreover, using $\nabla e_1 f_i = {}^{1}(\partial f_i)$,

$$\nabla e_1(f_1/\lambda_2) = {}^{1}[\partial f_1/\lambda_2] = \operatorname{Id} - \operatorname{Prox}_{f_1/\lambda_2}, \quad \nabla e_1(f_2/\lambda_1) = {}^{1}[\partial f_2/\lambda_1] = \operatorname{Id} - \operatorname{Prox}_{f_2/\lambda_1},$$

by Fact 4.2, we obtain

$$\partial g_1 = \partial f_1/\lambda_2 + {}^{1}[\partial f_2/\lambda_1] = \partial f_1/\lambda_2 + \operatorname{Id} - \operatorname{Prox}_{f_2/\lambda_1}, \tag{47}$$

$$\partial g_2 = \partial f_2/\lambda_1 + {}^{1}[\partial f_1/\lambda_2] = \partial f_2/\lambda_1 + \operatorname{Id} - \operatorname{Prox}_{f_1/\lambda_2}, \tag{48}$$

$$\nabla g_3 = \lambda_1\, {}^{1}(\partial f_1) + \lambda_2\, {}^{1}(\partial f_2) = \operatorname{Id} - (\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2}). \tag{49}$$

Then (i)–(iii) follows from Theorem 3.3 by using $A_1 = \partial f_1, A_2 = \partial f_2$, or [7, Proposition 4.1] by using $\gamma = 1$ and functions $f_1/\lambda_2, f_2/\lambda_1$. To show (iv), apply Fact 4.1(ii) to obtain $e_1 p(\boldsymbol{f}, \boldsymbol{\lambda}) = \lambda_1 e_1 f_1 + \lambda_2 e_1 f_2$. Since $\operatorname{argmin} e_1 p(\boldsymbol{f}, \boldsymbol{\lambda}) = \operatorname{argmin} p(\boldsymbol{f}, \boldsymbol{\lambda})$ by Fact 4.2, we have the first equality of (iv). Furthermore, (49) gives $\operatorname{argmin} g_3 = \nabla g_3^{-1}(0) = [\lambda_1\, {}^{1}(\partial f_1) + \lambda_2\, {}^{1}(\partial f_2)]^{-1}(0)$.

(v): We first show (45). Let $z \in \operatorname{argmin} p(\boldsymbol{f}, \boldsymbol{\lambda})$. By Fact 4.1(i), $z = \lambda_1 x + \lambda_2 y$ for some (x, y) with

$$\frac{p(\boldsymbol{f}, \boldsymbol{\lambda})(z)}{\lambda_1 \lambda_2} = \frac{f_1(x)}{\lambda_2} + \frac{f_2(y)}{\lambda_1} + \frac{\|x - y\|^2}{2}.$$

By Proposition 4.3, $\frac{p(\boldsymbol{f}, \boldsymbol{\lambda})(z)}{\lambda_1 \lambda_2} = \min g$, so $(x, y) \in \operatorname{argmin} g$. As $S = \operatorname{argmin} g$ by (i), we have $z \in \left\{\lambda_1 x + \lambda_2 y \mid (x, y) \in S\right\}$. Conversely, if $(x, y) \in S$, then by definition of $p(\boldsymbol{f}, \boldsymbol{\lambda})$ and Proposition 4.3,

$$\frac{p(\boldsymbol{f}, \boldsymbol{\lambda})(\lambda_1 x + \lambda_2 y)}{\lambda_1 \lambda_2} \leq \frac{f_1(x)}{\lambda_2} + \frac{f_2(y)}{\lambda_1} + \frac{\|x - y\|^2}{2} = \min g = \min \frac{p(\boldsymbol{f}, \boldsymbol{\lambda})}{\lambda_1 \lambda_2},$$

thus $\lambda_1 x + \lambda_2 y \in \operatorname{argmin} p(\boldsymbol{f}, \boldsymbol{\lambda})$. Therefore, (45) holds. Note that (46) follows from Fact 4.1(iii) and Fact 2.2(i) for $\gamma = 1$.

(vi): Indeed, these sets are argmin sets of lower semicontinuous convex functions $g, g_1, g_2, p(\boldsymbol{f}, \boldsymbol{\lambda})$ respectively. ■

Problem (39) is a least squares problem in terms of convex functions f_1, f_2. The next result is well known.

Corollary 4.8 (least square solution) *Let $f_1, f_2 \in \Gamma(\mathcal{H})$ and $\lambda_1 + \lambda_2 = 1$ with each $\lambda_i > 0$. Then*

$$\operatorname{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2}) = \operatorname{argmin}(\lambda_1 e_1 f_1 + \lambda_2 e_1 f_2). \tag{50}$$

When $f_i = \iota_{C_i}$ with $C_i \subseteq \mathcal{H}$ being nonempty closed convex, we have

$$\operatorname{Fix}(\lambda_1 P_{C_1} + \lambda_2 P_{C_2}) = \operatorname{argmin}\left(\lambda_1 \frac{d_{C_1}^2}{2} + \lambda_2 \frac{d_{C_2}^2}{2}\right). \tag{51}$$

Proof. Combining Theorem 4.7(iv) and (v) gives (50). Observe that $\operatorname{Prox}_{\iota_{C_i}} = P_{C_i}$ and $e_1 \iota_{C_i} = d_{C_i}^2/2$. Hence (51) follows from (50). ■

The following result says that when $S \neq \varnothing$, for every $(x, y) \in S$ the difference $x - y$, sometimes also called the *gap vector*, is the unique solution to the dual problem. Characterizations of S, E, F, and $\operatorname{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2})$ in terms of dual solution $\bar{\phi}$ come as follows.

Theorem 4.9 (i) *We have $(x, y) \in S$ and $\phi \in S^*$ if and only if*

$$(-\phi, \phi) \in \partial f_1(x)/\lambda_2 \times \partial f_2(y)/\lambda_1, \quad \phi = x - y.$$

(ii) *Let $\bar{\phi}$ be the unique solution to (D) and assume that $S \neq \varnothing$. Then for every $(x, y) \in S$, one has $x - y = \bar{\phi}$. Moreover,*

$$S = \big(\partial f_1/\lambda_2 \times \partial f_2/\lambda_1\big)^{-1}(-\bar{\phi}, \bar{\phi}) \cap (R - \operatorname{Id})^{-1}(-\bar{\phi}, \bar{\phi}).$$

Proof. (i). Use $L^* = (\operatorname{Id}, -\operatorname{Id}) : \mathcal{H} \to \mathcal{H} \times \mathcal{H}$, $f = f_1/\lambda_2 \oplus f_2/\lambda_1$ and $g = j$. By Fact 2.8 again, $(x, y) \in S$ and $\phi \in S^*$ if and only if

$$(-\phi, \phi) \in \partial f_1(x)/\lambda_2 \times \partial f_2(y)/\lambda_1, \quad \phi = x - y.$$

(ii). As the dual objective function is strictly concave, the dual solution is unique, say $\bar{\phi}$. It suffices to apply (i). ■

Theorem 4.10 (i) *$x \in E$ if and only if*

$$-\bar{\phi} \in \partial f_1(x)/\lambda_2, \quad \bar{\phi} = x - \operatorname{Prox}_{f_2/\lambda_1}(x).$$

(ii) *$y \in F$ if and only if*

$$-\bar{\phi} = y - \operatorname{Prox}_{f_1/\lambda_2}(y), \quad \bar{\phi} \in \partial f_2(y)/\lambda_1.$$

(iii) $z \in \operatorname{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2})$ *if and only if*

$$-\bar{\phi} = \lambda_2^{-1}(z - \operatorname{Prox}_{f_1}(z)), \quad \bar{\phi} = \lambda_1^{-1}(z - \operatorname{Prox}_{f_2}(z)).$$

Proof. This follows from Theorem 4.4 and Fact 2.8. ■

Remark 4.11 Theorems 4.9 and 4.10 are convex function analogues for Theorem 3.3(vi), (vii), (ix) and (x).

4.3 Relationship among minimizers

We now study the relationships among $\operatorname{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2})$, $E = \operatorname{Fix}(\operatorname{Prox}_{f_1/\lambda_2} \operatorname{Prox}_{f_2/\lambda_1})$, $F = \operatorname{Fix}(\operatorname{Prox}_{f_2/\lambda_1} \operatorname{Prox}_{f_1/\lambda_2})$, and

$$S = \big\{(x,y) \mid x = \operatorname{Prox}_{f_1/\lambda_2} y,\ y = \operatorname{Prox}_{f_2/\lambda_1} x\big\}.$$

Theorem 4.12 *Let $\bar{\phi}$ be the dual solution, i.e., the solution to* (40). *Define $T_1 : E \to \operatorname{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2})$ by*

$$T_1(x) = \lambda_1 x + \lambda_2 \operatorname{Prox}_{f_2/\lambda_1}(x) = x - \lambda_2 \bar{\phi}.$$

Then T_1 is a homeomorphism with $T_1^{-1}(z) = \operatorname{Prox}_{f_1}(z)$. Consequently,

$$\operatorname{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2}) = E - \lambda_2 \bar{\phi}. \tag{52}$$

Proof. Use Theorem 3.8(i) with $A_i = \partial f_i$ for $i = 1, 2$. ■

Theorem 4.13 *Let $\bar{\phi}$ be the dual solution. Define $T_2 : F \to \operatorname{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2})$ by*

$$T_2(y) = \lambda_1 \operatorname{Prox}_{f_1/\lambda_2}(y) + \lambda_2 y = y + \lambda_1 \bar{\phi}.$$

Then T_2 is a homeomorphism with $T_2^{-1}(z) = \operatorname{Prox}_{f_2}(z)$. Consequently,

$$\operatorname{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2}) = F + \lambda_1 \bar{\phi}. \tag{53}$$

Proof. Apply Theorem 3.8(ii) with $A_i = \partial f_i$ for $i = 1, 2$. ■

Theorem 4.14 *Define $T : S \to \operatorname{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2})$ by*

$$T(x,y) = \lambda_1 x + \lambda_2 y.$$

Then T is a homeomorphism. Moreover, for every $z \in \operatorname{Fix}(\operatorname{Prox}_{p(\boldsymbol{f},\boldsymbol{\lambda})})$ one has $T^{-1}(z) = (\operatorname{Prox}_{f_1}(z), \operatorname{Prox}_{f_2}(z))$. Consequently, $\operatorname{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2}) = L(S)$ where $L : \mathcal{H} \times \mathcal{H} \to \mathcal{H} : (x,y) \mapsto \lambda_1 x + \lambda_2 y$.

Proof. Use Theorem 3.6 with $A_i = \partial f_i$ for $i = 1, 2$. ■

Theorem 4.15 *The mapping* $\mathrm{Prox}_{f_2/\lambda_1}|_E : E \to F$ *is a homeomorphism with inverse* $\mathrm{Prox}_{f_1/\lambda_2}|_F$.

Proof. The results follow from Theorem 3.3(viii). ■

Corollary 4.16 (i) *$E \neq \varnothing$ if and only if $F \neq \varnothing$ if and only if $S \neq \varnothing$ if and only if* $\operatorname{argmin} p(\boldsymbol{f}, \boldsymbol{\lambda}) = \operatorname{Fix}(\lambda_1 \mathrm{Prox}_{f_1} + \lambda_2 \mathrm{Prox}_{f_2}) \neq \varnothing$.

(ii) *E is a singleton if and only if F is a singleton if and only if S is a singleton if and only if* $\operatorname{argmin} p(\boldsymbol{f}, \boldsymbol{\lambda}) = \operatorname{Fix}(\lambda_1 \mathrm{Prox}_{f_1} + \lambda_2 \mathrm{Prox}_{f_2})$ *is a singleton.*

(iii) $\operatorname{Fix}(\lambda_1 \mathrm{Prox}_{f_1} + \lambda_2 \mathrm{Prox}_{f_2}) = \lambda_1 E + \lambda_2 F$.

Proof. (i) and (ii): Combine Theorems 4.12, 4.13, 4.14. (iii): Apply Corollary 3.9(iii) with $A_i = \partial f_i$ for $i = 1, 2$. ■

Applying Theorem 4.14 to $\lambda_2 f_1, \lambda_1 f_2$ gives

Corollary 4.17

$$\operatorname{Fix}(\lambda_1 \mathrm{Prox}_{\lambda_2 f_1} + \lambda_2 \mathrm{Prox}_{\lambda_1 f_2}) = \big\{\lambda_1 x + \lambda_2 y \mid x = \mathrm{Prox}_{f_1} y, y = \mathrm{Prox}_{f_2}(x)\big\}.$$

4.4 The case when $\operatorname{argmin} f_1 \cap \operatorname{argmin} f_2 \neq \varnothing$

Note that

$$\operatorname{Fix}(\mathrm{Prox}_{f_1/\lambda_2}) = \operatorname{Fix}(\mathrm{Prox}_{f_1}) = \operatorname{argmin} f_1, \quad \operatorname{Fix}(\mathrm{Prox}_{f_2/\lambda_1}) = \operatorname{Fix}(\mathrm{Prox}_{f_2}) = \operatorname{argmin} f_2. \tag{54}$$

Theorem 4.18 *Assume that* $\operatorname{argmin} f_1 \cap \operatorname{argmin} f_2 \neq \varnothing$. *Then*

$$\operatorname{Fix}(\lambda_1 \mathrm{Prox}_{f_1} + \lambda_2 \mathrm{Prox}_{f_2}) = \operatorname{Fix}(\mathrm{Prox}_{f_1/\lambda_2}) \cap \operatorname{Fix}(\mathrm{Prox}_{f_2/\lambda_1}) = \operatorname{Fix}(\mathrm{Prox}_{f_1}) \cap \operatorname{Fix}(\mathrm{Prox}_{f_2}). \tag{55}$$

Moreover,

$$\operatorname{Fix}(\lambda_1 \mathrm{Prox}_{f_1} + \lambda_2 \mathrm{Prox}_{f_2}) = \operatorname{Fix}(\mathrm{Prox}_{f_1/\lambda_2} \mathrm{Prox}_{f_2/\lambda_1}) = \operatorname{Fix}(\mathrm{Prox}_{f_2/\lambda_1} \mathrm{Prox}_{f_1/\lambda_2}). \tag{56}$$

Proof. Apply Theorem 3.10. ■

4.5 Examples on projections

Projection algorithms, which are instances of the proximal point algorithm, are important in applications. Let $C_1, C_2 \subseteq \mathcal{H}$ be nonempty closed convex sets. With $f_i = \iota_{C_i}$, (34), (37), (38), (39) transpire to

$$\min_{(x,y)} g(x,y) = \left(\iota_{C_1}(x) + \iota_{C_2}(y) + \frac{\|x-y\|^2}{2}\right), \tag{57}$$

$$\min_x \; g_1(x) = \iota_{C_1}(x) + \tfrac{1}{2} d^2_{C_2}(x), \tag{58}$$

$$\min_y \; g_2(y) = \tfrac{1}{2} d^2_{C_1}(y) + \iota_{C_2}(y), \tag{59}$$

$$\min_z \; \frac{g_3(z)}{\lambda_1\lambda_2} = \lambda_2^{-1}\frac{d^2_{C_1}(z)}{2} + \lambda_1^{-1}\frac{d^2_{C_2}(z)}{2}. \tag{60}$$

The Fenchel-Rockafellar dual of (57) given by (40) transpires to

$$-\inf_\phi \left(\sigma_{C_2 - C_1}(\phi) + \frac{\|\phi\|^2}{2} \right),$$

with the unique solution $\bar\phi$, where $\sigma_{C_2-C_1}(\phi) = \sup\big\{\langle \phi, y - x\rangle \mid x \in C_1, y \in C_2\big\}$. In fact, convex calculus (see, e.g., [3, Theorem 2.1]) or Remark 4.6 yields

$$\bar\phi = -P_{\overline{C_2 - C_1}}(0). \tag{61}$$

The next two results are partially contained in [3], [4], and [17].

Theorem 4.19 *Let $C_1, C_2 \subseteq \mathcal{H}$ be nonempty closed convex sets. Then*

(i) (**Fixed points of alternating projections**) $S = \operatorname{argmin} g = \big\{(x, y) \mid x = P_{C_1}(y), y = P_{C_2}x\big\}$.

(ii) (**Fixed points of projection composition**) $E = \operatorname{argmin} g_1 = \big\{x \mid x = P_{C_1}P_{C_2}x\big\}$.

(iii) (**Fixed points of projection composition**) $F = \operatorname{argmin} g_2 = \big\{y \mid y = P_{C_2}P_{C_1}y\big\}$.

(iv) (**Fixed points of the average of projections**) $\operatorname{argmin} g_3 = \big\{z \mid z = \lambda_1 P_{C_1}(z) + \lambda_2 P_{C_2}(z)\big\}$.

Moreover,

(i) *The mapping $T : S \to \operatorname{Fix}(\lambda_1 P_{C_1} + \lambda_2 P_{C_2})$ given by*

$$T(x, y) = \lambda_1 x + \lambda_2 y,$$

is a homeomorphism with inverse $T^{-1}(z) = (P_{C_1}(z), P_{C_2}(z))$ for every $z \in \operatorname{Fix}(\lambda_1 P_{C_1} + \lambda_2 P_{C_2})$. Hence $\operatorname{Fix}(\lambda_1 P_{C_1} + \lambda_2 P_{C_2}) = L(S)$ where $L : \mathcal{H} \times \mathcal{H} \to \mathcal{H} : (x, y) \mapsto \lambda_1 x + \lambda_2 y$.

(ii) *The mapping $H_1 : E \to \operatorname{Fix}(\lambda_1 P_{C_1} + \lambda_2 P_{C_2})$ given by*

$$H_1(x) = \lambda_1 x + \lambda_2 P_{C_2} x = x - \lambda_2 \bar\phi,$$

is a homeomorphism with inverse $H_1^{-1}(z) = P_{C_1}(z)$ for every $z \in \operatorname{Fix}(\lambda_1 P_{C_1} + \lambda_2 P_{C_2})$. Hence $\operatorname{Fix}(\lambda_1 P_{C_1} + \lambda_2 P_{C_2}) = E - \lambda_2 \bar\phi$.

(iii) *The mapping* $H_2 : F \to \operatorname{Fix}(\lambda_1 P_{C_1} + \lambda_2 P_{C_2})$ *given by*

$$H_2(y) = \lambda_1 P_{C_1}(y) + \lambda_2 y = \lambda_1 \bar{\phi} + y,$$

is a homeomorphism with inverse $H_2^{-1}(z) = P_{C_2}(z)$ *for every* $z \in \operatorname{Fix}(\lambda_1 P_{C_1} + \lambda_2 P_{C_2})$*. Hence* $\operatorname{Fix}(\lambda_1 P_{C_1} + \lambda_2 P_{C_2}) = F + \lambda_1 \bar{\phi}$*.*

(iv) $\operatorname{Fix}(\lambda_1 P_{C_1} + \lambda_2 P_{C_2}) = \lambda_1 E + \lambda_2 F$*.*

Theorem 4.20 *Assume that* $C_1, C_2 \subseteq \mathcal{H}$ *are two closed convex sets such that* $C_1 \cap C_2 \neq \varnothing$*. Then*

$$\operatorname{Fix}(\lambda_1 P_{C_1} + \lambda_2 P_{C_2}) = \operatorname{Fix} P_{C_1} P_{C_2} = \operatorname{Fix} P_{C_2} P_{C_1} = C_1 \cap C_2.$$

Proof. As $\min g_3 = \min g_2 = \min g_1 = \min g = 0$ when $C_1 \cap C_2 \neq \varnothing$, we have

$$\operatorname{argmin} g_1 = \operatorname{argmin} g_2 = \operatorname{argmin} g_3 = C_1 \cap C_2,$$

and $\operatorname{argmin} g = \big\{(x, x) \mid x \in C_1 \cap C_2\big\}$. Alternatively, use Theorem 4.18 or Theorem 3.10. ■

As $\partial \iota_C = N_C$, $\operatorname{Prox}_{\iota_C} = P_C$, Theorems 4.9 and 4.10 give characterizations of $S, E, F, \operatorname{Fix}(\lambda_1 P_{C_1} + \lambda_2 P_{C_2})$ in terms of the dual solution $\bar{\phi}$:

Theorem 4.21 (i) $(x, y) \in S$ *if and only if*

$$-\bar{\phi} \in N_{C_1}(x), \quad \bar{\phi} \in N_{C_2}(y), \quad \bar{\phi} = x - y.$$

(ii) $x \in E$ *if and only if*

$$-\bar{\phi} \in N_{C_1}(x), \quad \bar{\phi} = x - P_{C_2}(x).$$

(iii) $y \in F$ *if and only if*

$$-\bar{\phi} = y - P_{C_1}(y), \quad \bar{\phi} \in N_{C_2}(y).$$

(iv) $z \in \operatorname{Fix}(\lambda_1 P_{C_1} + \lambda_2 P_{C_2})$ *if and only if*

$$-\bar{\phi} = \lambda_2^{-1}(z - P_{C_1}(z)), \quad \bar{\phi} = \lambda_1^{-1}(z - P_{C_2}(z)).$$

5 Algorithms and examples

In this section, notation is as in section 3.3 and we also assume that $\operatorname{Fix} J_A \neq \varnothing$. By Corollary 3.9, E, F, S all are nonempty. The following results give different algorithms to find a point in $\operatorname{Fix} J_A$.

Theorem 5.1 (Fixed point of resolvent average by alternating resolvent method) *Fix* $x_0 \in \mathcal{H}$ *and for every* $n \in \mathbb{N}$*, set*

$$y_n = J_{A_2/\lambda_1} x_n, \qquad x_{n+1} = J_{A_1/\lambda_2} y_n.$$

Then $\lambda_1 x_n + \lambda_2 y_n \rightharpoonup \lambda_1 \overline{x} + \lambda_2 \overline{y} \in \operatorname{Fix} J_A$*.*

Proof. By [7, Theorem 3.3], $(x_n, y_n) \rightharpoonup (\overline{x}, \overline{y}) \in S$. By Theorem 3.6, $\lambda_1\bar{x} + \lambda_2\bar{y} \in \operatorname{Fix} J_A$. Therefore $\lambda_1 x_n + \lambda_2 y_n \rightharpoonup \lambda_1\overline{x} + \lambda_2\overline{y} \in \operatorname{Fix} J_A$. ■

Theorem 5.2 (Fixed point of resolvent average by proximal point method) *Fix $x_0 \in \mathcal{H}$ and for every $n \in \mathbb{N}$, set*

$$x_{n+1} = (\lambda_1 J_{A_1} + \lambda_2 J_{A_2})(x_n). \tag{62}$$

Then $x_n \rightharpoonup \bar{x} \in \operatorname{Fix} J_A$.

Proof. As $\lambda_1 J_{A_1} + \lambda_2 J_{A_2} = J_A$, the iteration (62) is the proximal point algorithm. By Fact 2.5, $x_n \rightharpoonup \bar{x} \in \operatorname{Fix} J_A$. ■

Theorem 5.3 (Fixed point of resolvent average by resolvent compositions)

(i) *Fix $x_0 \in \mathcal{H}$ and for every $n \in \mathbb{N}$, set*

$$x_{n+1} = J_{A_1/\lambda_2} J_{A_2/\lambda_1} x_n.$$

Then $x_n \rightharpoonup x \in \operatorname{Fix} J_{A_1/\lambda_2} J_{A_2/\lambda_1}$ and $\lambda_1 x_n + \lambda_2 J_{A_2/\lambda_1} x_n \rightharpoonup \lambda_1 x + \lambda_2 J_{A_2/\lambda_1} x \in \operatorname{Fix} J_A$.

(ii) *Fix $y_0 \in \mathcal{H}$ and for every $n \in \mathbb{N}$, set*

$$y_{n+1} = J_{A_2/\lambda_1} J_{A_1/\lambda_2} y_n.$$

Then $y_n \rightharpoonup y \in \operatorname{Fix} J_{A_2/\lambda_1} J_{A_1/\lambda_2}$ and $\lambda_1 J_{A_1/\lambda_2} y_n + \lambda_2 y_n \rightharpoonup \lambda_1 J_{A_1/\lambda_2} y + \lambda_2 y \in \operatorname{Fix} J_A$.

Proof. (i). Since $J_{A_1/\lambda_2}, J_{A_2/\lambda_1}$ are firmly nonexpansive, Fact 2.3 shows that $J_{A_1/\lambda_2} J_{A_2/\lambda_1}$ is strongly nonexpansive and that $x_n \rightharpoonup x \in \operatorname{Fix} J_{A_1/\lambda_2} J_{A_2/\lambda_1}$. By [7, Theorem 3.3(iii)], $J_{A_2/\lambda_1} x_n - x_n \to u^*$, which implies that $J_{A_2/\lambda_1} x_n \rightharpoonup u^* + x$. Hence $\lambda_1 x_n + \lambda_2 J_{A_2/\lambda_1} x_n \rightharpoonup \lambda_1 x + \lambda_2(u^* + x) = x + \lambda_2 u^* = \lambda_1 x + \lambda_2 J_{A_2/\lambda_1} x \in \operatorname{Fix} J_A$ by Theorem 3.8(i). (ii). The proof is similar to the proof of (i). ■

Remark 5.4 (i). Note that when $\mathcal{H} = \mathbb{R}^N$, the weak convergence and norm convergence coincide. Hence in $\mathbb{R}^N$, the convergence in Theorems 5.1, 5.2, 5.3 is norm convergence.

(ii). As J_A (even a projection mapping) need not be weakly sequentially continuous, one cannot conclude directly that $\lambda_1 x_n + \lambda_2 J_{A_2/\lambda_1} x_n \rightharpoonup \lambda_1 x + \lambda_2 J_{A_2/\lambda_1} x$ in Theorem 5.3(i) or that $\lambda_1 J_{A_1/\lambda_2} y_n + \lambda_2 y_n \rightharpoonup \lambda_1 J_{A_1/\lambda_2} y + \lambda_2 y$ in Theorem 5.3(ii). Indeed, following Zarantonello [31, page 245], consider the real Hilbert sequence space ℓ^2. Let $\mathbb{B} := \{x \in \ell^2 \mid \|x\| \leq 1\}$ and $(e_n)_{n\in\mathbb{N}}$ be the basis vectors, i.e., $e_n = (\underbrace{0, \cdots, 0, 1}_{n \text{ terms}}, 0, \cdots)$. We have

$$e_1 + e_n \rightharpoonup e_1, \quad P_{\mathbb{B}}(e_1) = e_1,$$

$$(\forall n \geq 2)\ P_{\mathbb{B}}(e_1 + e_n) = \frac{e_1 + e_n}{\sqrt{2}} \rightharpoonup \frac{e_1}{\sqrt{2}} \neq P_{\mathbb{B}}(e_1).$$

Hence $P_{\mathbb{B}}$ is not weakly sequentially continuous. However, in the proof of Theorem 5.3, we invoked the analysis in [7, Section 3.2] which allowed us to obtain the weak convergence conclusion.

We end with three examples to illustrate our main results.

Example 5.5 Consider

$$C_1 = \{(x,y) \mid x^2 + (y-2)^2 \le 1\}, \quad C_2 = \{(x,0) \mid x \in \mathbb{R}\}.$$

Then $C_1 \cap C_2 = \varnothing$. *We claim that when $\lambda_1 + \lambda_2 = 1$ with $\lambda_i > 0$,*

$$\operatorname{Fix}(\lambda_1 P_{C_1} + \lambda_2 P_{C_2}) = \{(0, \lambda_1)\}.$$

In this example, $\operatorname{Fix} P_{C_1}P_{C_2}$ is easier to compute than $\operatorname{Fix}(\lambda_1 P_{C_1} + \lambda_2 P_{C_2})$. Indeed, we have

$$P_{C_1}(x,y) = \begin{cases} \left(\dfrac{x}{\sqrt{x^2+(y-2)^2}}, \dfrac{y-2}{\sqrt{x^2+(y-2)^2}} + 2\right), & \text{if } (x,y) \notin C_1 \\ (x,y), & \text{if } (x,y) \in C_1, \end{cases}$$

$$P_{C_2}(x,y) = (x,0).$$

Thus,

$$\begin{aligned} P_{C_1}P_{C_2}(x,y) &= P_{C_1}(x,0) \\ &= \left(\frac{x}{\sqrt{x^2+4}}, \frac{-2}{\sqrt{x^2+4}} + 2\right), \end{aligned}$$

since $(x,0) \notin C_1$. Start with (x_0,y_0). Consider the composition algorithm $(x_{n+1}, y_{n+1}) = P_{C_1}P_{C_2}(x_n,y_n)$. We have

$$(x_{n+1}, y_{n+1}) = \left(\frac{x_n}{\sqrt{x_n^2+4}}, \frac{-2}{\sqrt{x_n^2+4}} + 2\right).$$

It follows that

$$|x_{n+1}| = \frac{|x_n|}{\sqrt{x_n^2+4}} \le \frac{|x_n|}{2} \le \cdots \le \frac{|x_0|}{2^{n+1}},$$

and this gives $x_{n+1} \to 0$, consequently $y_{n+1} \to 1$. Therefore,

$$(0,1) \in \operatorname{Fix} P_{C_1}P_{C_2}.$$

In fact, by using

$$(x,y) = \left(\frac{x}{\sqrt{x^2+4}}, \frac{-2}{\sqrt{x^2+4}} + 2\right),$$

we see that $(x,y) = (0,1)$. Hence $\operatorname{Fix} P_{C_1}P_{C_2} = \{(0,1)\}$. Therefore, by Theorem 4.19(ii)

$$\operatorname{Fix}(\lambda_1 P_{C_1} + \lambda_2 P_{C_2}) = \lambda_1(0,1) + \lambda_2 P_{C_2}(0,1) = (0,\lambda_1), \tag{63}$$

since $P_{C_2}(0,1) = (0,0)$.

For $P_{C_2}P_{C_1}$, since $P_{C_2}P_{C_1}(x_1,y_1)\in C_2$ (not in C_1) we have

$$(x_{n+1},y_{n+1}) = P_{C_2}P_{C_1}(x_n,y_n) = \left(\frac{x_n}{\sqrt{x_n^2+4}},0\right) \quad \forall\, n\geq 2,$$

and $\mathrm{Fix}(P_{C_2}P_{C_1}) = \{(0,0)\}$. Then for $(0,0)\in \mathrm{Fix}(P_{C_2}P_{C_1})$,

$$\lambda_1 P_{C_1}(0,0) + \lambda_2(0,0) = \lambda_1(0,1) = (0,\lambda_1), \tag{64}$$

which shows also that $\mathrm{Fix}(\lambda_1 P_{C_1}+\lambda_2 P_{C_2}) = \{(0,\lambda_1)\}$.

On the other hand, the averaged projection method proceeds as follows.

$$(\lambda_1 P_{C_1}+\lambda_2 P_{C_2})(x,y) = \begin{cases} \lambda_1\left(\frac{x}{\sqrt{x^2+(y-2)^2}}, \frac{y-2}{\sqrt{x^2+(y-2)^2}}+2\right)+\lambda_2(x,0), & \text{if } (x,y)\notin C_1;\\ \lambda_1(x,y)+\lambda_2(x,0), & \text{if } (x,y)\in C_1,\end{cases}$$

$$(x_{n+1},y_{n+1}) = (\lambda_1 P_{C_1}+\lambda_2 P_{C_2})(x_n,y_n). \tag{65}$$

Start with any $(x_0,y_0)\in\mathbb{R}^2$.

Claim: If n is sufficiently large, then $y_n<2$ and $(x_n,y_n)\notin C_1$.

To see that, for $y_n\geq 2$ consider two cases: if $(x_n,y_n)\in C_1$, then $y_n\geq 1$, and

$$(x_{n+1},y_{n+1}) = \lambda_1(x_n,y_n)+\lambda_2(x_n,0) = (x_n,\lambda_1 y_n),$$

which gives $y_{n+1}=\lambda_1 y_n$; if $(x_n,y_n)\notin C_1$, then $\sqrt{x_n^2+(y_n-2)^2}\geq 1$ and

$$(x_{n+1},y_{n+1}) = \lambda_1\left(\frac{x_n}{\sqrt{x_n^2+(y_n-2)^2}}, \frac{y_n-2}{\sqrt{x_n^2+(y_n-2)^2}}+2\right)+\lambda_2(x_n,0),$$

so that

$$y_{n+1} = \lambda_1\left(\frac{y_n-2}{\sqrt{x_n^2+(y_n-2)^2}}+2\right) \leq \lambda_1(y_n-2+2) = \lambda_1 y_n.$$

Furthermore, whenever $(x_n,y_n)\in C_1$, we have $y_{n+1}=\lambda_1 y_n$ and $y_n\geq 1$. Altogether, $y_{n+1}\leq\lambda_1 y_n$. This implies that the averaged projection iterations can only stay in C_1 for only a finite number of times and that for n sufficiently large $y_n<2$. Hence for all n sufficiently large, the average projection algorithm (65) gives $y_n<2$ and

$$(x_{n+1},y_{n+1}) = \left(\lambda_1\frac{x_n}{\sqrt{x_n^2+(y_n-2)^2}}+\lambda_2 x_n, \lambda_1\left(\frac{y_n-2}{\sqrt{x_n^2+(y_n-2)^2}}+2\right)\right).$$

Moreover, as for n sufficiently large

$$1\leq \frac{y_n-2}{\sqrt{x_n^2+(y_n-2)^2}}+2\leq 2,$$

we have $\lambda_1 \le y_{n+1} \le \lambda_1 2 < 2$. Then $(x,y) \in \mathrm{Fix}(\lambda_1 P_{C_1} + \lambda_2 P_{C_2})$ means

$$(x,y) = \left(\lambda_1 \frac{x}{\sqrt{x^2+(y-2)^2}} + \lambda_2 x, \lambda_1\left(\frac{y-2}{\sqrt{x^2+(y-2)^2}} + 2\right)\right),$$

which gives only one solution $(x,y) = (0,\lambda_1)$ in view of $\lambda_1 > 0, \lambda_1 \le y < 2$. Again this shows that

$$\mathrm{Fix}(\lambda_1 P_{C_1} + \lambda_2 P_{C_2}) = \{(0,\lambda_1)\},$$

which is consistent with the results given by (63) and (64). (See also [4, Example 5.3] for more on the rate of convergence of alternating projections.)

Now what can one say about the relationships among $\mathrm{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2})$, $\mathrm{Fix}(\operatorname{Prox}_{f_1} \operatorname{Prox}_{f_2})$, $\mathrm{Fix}(\operatorname{Prox}_{f_2} \operatorname{Prox}_{f_1})$? It is tempting to conjecture that for fixed points of alternating iterations:

$$x = \operatorname{Prox}_{f_1} y, \quad y = \operatorname{Prox}_{f_2} x,$$

one has $\lambda_1 x + \lambda_2 y \in \mathrm{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2})$ — and this is true for projections — but this is not right in general, as the following examples show.

Example 5.6 Consider $f_1(x) = x^2, f_2(x) = (x-1)^2$ for $x \in \mathbb{R}$. Let $\lambda_1 + \lambda_2 = 1$ with $\lambda_i > 0$. Then $\operatorname{argmin} f_1 = \{0\}, \operatorname{argmin} f_2 = \{1\}$, so $\operatorname{argmin} f_1 \cap \operatorname{argmin} f_2 = \varnothing$. As $\nabla f_1(x) = 2x, \nabla f_1(x)/\lambda_2 = 2x/\lambda_2, \nabla f_2(x) = 2(x-1), \nabla f_2(x)/\lambda_1 = 2(x-1)/\lambda_1$, for every $z \in \mathbb{R}$,

$$\operatorname{Prox}_{f_1}(z) = \frac{z}{3}, \quad \operatorname{Prox}_{f_2}(z) = \frac{z+2}{3},$$

$$\operatorname{Prox}_{f_1/\lambda_2}(z) = \frac{\lambda_2 z}{2+\lambda_2}, \quad \operatorname{Prox}_{f_2/\lambda_1}(z) = \frac{\lambda_1 z + 2}{2+\lambda_1}.$$

Moreover,

$$x \mapsto \operatorname{Prox}_{f_1/\lambda_2} \operatorname{Prox}_{f_2/\lambda_1}(x) = \frac{\lambda_2}{2+\lambda_2}\frac{\lambda_1 x + 2}{2+\lambda_1},$$

$$y \mapsto \operatorname{Prox}_{f_2/\lambda_1} \operatorname{Prox}_{f_1/\lambda_2}(y) = \frac{1}{2+\lambda_1}\left(\frac{\lambda_1\lambda_2 y}{2+\lambda_2} + 2\right),$$

$$z \mapsto (\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2})(z) = \frac{z}{3} + \frac{2\lambda_2}{3}.$$

We have

$$\mathrm{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2}) = \{\lambda_2\},$$
$$\mathrm{Fix}(\operatorname{Prox}_{f_1/\lambda_2} \operatorname{Prox}_{f_2/\lambda_1}) = \{\lambda_2/3\},$$
$$\mathrm{Fix}(\operatorname{Prox}_{f_2/\lambda_1} \operatorname{Prox}_{f_1/\lambda_2}) = \{(2+\lambda_2)/3\},$$
$$S = \{(x,y) |\ x = \operatorname{Prox}_{f_1/\lambda_2}(y), y = \operatorname{Prox}_{f_2/\lambda_1}(x)\} = \{(\lambda_2/3, (\lambda_2+2)/3)\}.$$

As in Theorem 4.12, for $x \in E = \{\lambda_2/3\}$,

$$\lambda_1 x + \lambda_2 \operatorname{Prox}_{f_2/\lambda_1}(x) = \lambda_1\lambda_2/3 + \lambda_2 \frac{\lambda_1\lambda_2/3 + 2}{2+\lambda_1} = \lambda_2 \in \mathrm{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2}).$$

As in Theorem 4.13, for $y \in F = \{(\lambda_2 + 2)/3\}$,

$$\lambda_1 \operatorname{Prox}_{f_1/\lambda_2}(y) + \lambda_2 y = \lambda_1 \frac{\lambda_2}{2+\lambda_2} \frac{2+\lambda_2}{3} + \lambda_2 \frac{2+\lambda_2}{3} = \lambda_2 \in \operatorname{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2}).$$

As in Theorem 4.14, for $z \in \operatorname{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2}) = \{\lambda_2\}$,

$$T^{-1}(\lambda_2) = (\operatorname{Prox}_{f_1}(\lambda_2), \operatorname{Prox}_{f_2}(\lambda_2)) = (\lambda_2/3, (\lambda_2 + 2)/3) \in S.$$

As in Theorem 4.9, the dual solution satisfies

$$-\bar{\phi} = y - x = 2/3 = \operatorname{Prox}_{f_2/\lambda_1}(x) - x = y - \operatorname{Prox}_{f_1/\lambda_2}(y),$$

for $(x, y) \in S$.

We now show that for fixed points of alternating iterations

$$x = \operatorname{Prox}_{f_1} y, \quad y = \operatorname{Prox}_{f_2} x,$$

one has $\lambda_1 x + \lambda_2 y \notin \operatorname{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2})$. Indeed, as

$$x \mapsto \operatorname{Prox}_{f_1} \operatorname{Prox}_{f_2}(x) = \frac{x+2}{9},$$

$\operatorname{Fix}(\operatorname{Prox}_{f_1} \operatorname{Prox}_{f_2}) = \{1/4\}$. With $x = 1/4$, we have

$$\lambda_1 x + \lambda_2 \operatorname{Prox}_{f_2}(x) = \lambda_1 1/4 + \lambda_2 \frac{1/4 + 2}{3} = 1/4 + \lambda_2/2 \neq \lambda_2 \in \operatorname{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2}),$$

unless $\lambda_2 = 1/2$. Similarly, one can also show that

$$y \mapsto \operatorname{Prox}_{f_2} \operatorname{Prox}_{f_1}(y) = \frac{y+6}{9},$$

has $\operatorname{Fix}(\operatorname{Prox}_{f_2} \operatorname{Prox}_{f_1}) = \{3/4\}$. With $y = 3/4$,

$$\lambda_1 \operatorname{Prox}_{f_1}(y) + \lambda_2 y = 1/4 + \lambda_2/2 \neq \lambda_2 \in \operatorname{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2}),$$

unless $\lambda_2 = 1/2$.

However, for $\lambda_2 f_1(x) = \lambda_2 x^2$, $\lambda_1 f_2(x) = \lambda_1 (x-1)^2$ with

$$\operatorname{Prox}_{\lambda_2 f_1}(x) = \frac{x}{2\lambda_2 + 1}, \quad \operatorname{Prox}_{\lambda_1 f_2}(x) = \frac{x + 2\lambda_1}{2\lambda_1 + 1},$$

for every $x \in \mathbb{R}$, we have

$$(\lambda_1 \operatorname{Prox}_{\lambda_2 f_1} + \lambda_2 \operatorname{Prox}_{\lambda_1 f_2})(x) = \lambda_1 \frac{x}{2\lambda_2 + 1} + \lambda_2 \frac{x + 2\lambda_1}{2\lambda_1 + 1}.$$

By solving

$$x = \lambda_1 \frac{x}{2\lambda_2 + 1} + \lambda_2 \frac{x + 2\lambda_1}{2\lambda_1 + 1},$$

one indeed has

$$\operatorname{Fix}(\lambda_1 \operatorname{Prox}_{\lambda_2 f_1} + \lambda_2 \operatorname{Prox}_{\lambda_1 f_2}) = \{1/4 + \lambda_2/2\} = \big\{\lambda_1 x + \lambda_2 y \mid x = \operatorname{Prox}_{f_1} y, y = \operatorname{Prox}_{f_2}(x)\big\},$$

as Theorem 4.14 or Corollary 4.17 shows.

Example 5.7 Consider $f_1(x) = |x|, f_2(x) = (x-1)^2$ for every $x \in \mathbb{R}$ and $\lambda_1 + \lambda_2 = 1$ with $\lambda_i > 0$. Then $\operatorname{argmin} f_1 = \{0\}, \operatorname{argmin} f_2 = \{1\}$, so $\operatorname{argmin} f_1 \cap \operatorname{argmin} f_2 = \varnothing$. As

$$\partial f_1(x) = \begin{cases} 1 & \text{if } x > 0, \\ [-1, 1] & \text{if } x = 0, \\ -1 & \text{if } x < 0, \end{cases}$$

we have

$$\operatorname{Prox}_{f_1}(x) = \begin{cases} x - 1 & \text{if } x > 1, \\ 0 & \text{if } -1 \le x \le 1, \\ x + 1 & \text{if } x < -1, \end{cases}$$

$$\operatorname{Prox}_{f_2}(x) = \frac{x+2}{3}.$$

Then

$$(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2})(x) = \begin{cases} \lambda_1(x-1) + \frac{\lambda_2(x+2)}{3} & \text{if } x > 1, \\ \frac{\lambda_2(x+2)}{3} & \text{if } -1 \le x \le 1, \\ \lambda_1(x+1) + \frac{\lambda_2(x+2)}{3} & \text{if } x < -1, \end{cases}$$

and

$$\operatorname{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2}) = \left\{ \frac{2\lambda_2}{2 + \lambda_1} \right\}.$$

Moreover,

$$\begin{aligned} \operatorname{Prox}_{f_1} \operatorname{Prox}_{f_2}(x) &= \operatorname{Prox}_{f_1}((x+2)/3) \\ &= \begin{cases} \frac{x-1}{3} & \text{if } x > 1, \\ 0 & \text{if } -5 \le x \le 1, \\ \frac{x+5}{3} & \text{if } x < -5, \end{cases} \end{aligned}$$

and

$$\operatorname{Fix}(\operatorname{Prox}_{f_1} \operatorname{Prox}_{f_2}) = \{0\}.$$

For $x \in E = \operatorname{Fix}(\operatorname{Prox}_{f_1} \operatorname{Prox}_{f_2})$, i.e., $x = 0$,

$$\begin{aligned} \lambda_1 x + \lambda_2 \operatorname{Prox}_{f_2}(x) &= \lambda_1 0 + \lambda_2 \operatorname{Prox}_{f_2}(0) \\ &= \frac{2\lambda_2}{3} \notin \operatorname{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2}). \end{aligned}$$

Therefore, one cannot use the fixed point set of

$$x = \operatorname{Prox}_{f_1} y, \quad y = \operatorname{Prox}_{f_2} x,$$

to recover $\operatorname{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2})$.

Now let us consider $\mathrm{Prox}_{f_1/\lambda_2}, \mathrm{Prox}_{f_2/\lambda_1}$. As

$$\partial f_1(x)/\lambda_2 = \begin{cases} 1/\lambda_2 & \text{if } x > 0, \\ [-1/\lambda_2, 1/\lambda_2] & \text{if } x = 0, \\ -1/\lambda_2 & \text{if } x < 0, \end{cases}$$

$$\partial f_2(x)/\lambda_1 = \frac{2(x-1)}{\lambda_1},$$

we have

$$\mathrm{Prox}_{f_1/\lambda_2}(x) = \begin{cases} x - 1/\lambda_2 & \text{if } x > 1/\lambda_2, \\ 0 & \text{if } -1/\lambda_2 \leq x \leq 1/\lambda_2, \\ x + 1/\lambda_2 & \text{if } x < -1/\lambda_2, \end{cases}$$

$$\mathrm{Prox}_{f_2/\lambda_1}(x) = \frac{\lambda_1 x + 2}{2 + \lambda_1}.$$

It follows that

$$\begin{aligned} \mathrm{Prox}_{f_1/\lambda_2} \mathrm{Prox}_{f_2/\lambda_1}(x) &= \mathrm{Prox}_{f_1/\lambda_2}((\lambda_1 x + 2)/(2 + \lambda_1)) \\ &= \begin{cases} \frac{\lambda_1 x + 2}{2+\lambda_1} - \frac{1}{\lambda_2} & \text{if } x > \frac{3}{\lambda_2}, \\ 0 & \text{if } \frac{-(3+\lambda_2)}{\lambda_1 \lambda_2} \leq x \leq \frac{3}{\lambda_2}, \\ \frac{\lambda_1 x + 2}{2+\lambda_1} + \frac{1}{\lambda_2} & \text{if } x < \frac{-(3+\lambda_2)}{\lambda_1 \lambda_2}, \end{cases} \end{aligned}$$

and that

$$\mathrm{Fix}(\mathrm{Prox}_{f_1/\lambda_2} \mathrm{Prox}_{f_2/\lambda_1}) = \{0\}.$$

For $x \in E = \mathrm{Fix}(\mathrm{Prox}_{f_1/\lambda_2} \mathrm{Prox}_{f_2/\lambda_1})$, i.e., $x = 0$, we have

$$\lambda_1 0 + \lambda_2 \mathrm{Prox}_{f_2/\lambda_1}(0) = \lambda_2 \frac{\lambda_1 0 + 2}{2 + \lambda_1} = \frac{2\lambda_2}{2 + \lambda_1} \in \mathrm{Fix}(\lambda_1 \mathrm{Prox}_{f_1} + \lambda_2 \mathrm{Prox}_{f_2}).$$

Again, these testify that

$$\mathrm{Fix}(\lambda_1 \mathrm{Prox}_{f_1} + \lambda_2 \mathrm{Prox}_{f_2}) = \{\lambda_1 x + \lambda_2 \mathrm{Prox}_{f_2/\lambda_1}(x) |\ x \in E\}.$$

Similarly, one can verify that

$$\mathrm{Prox}_{f_2/\lambda_1} \mathrm{Prox}_{f_1/\lambda_2}(x) = \begin{cases} \frac{\lambda_1(x - 1/\lambda_2) + 2}{2+\lambda_1} & \text{if } x > 1/\lambda_2, \\ \frac{2}{2+\lambda_1} & \text{if } -1/\lambda_2 \leq x \leq 1/\lambda_2, \\ \frac{\lambda_1(x + 1/\lambda_2) + 2}{2+\lambda_1} & \text{if } x < -1/\lambda_2, \end{cases}$$

and

$$F = \mathrm{Fix}(\mathrm{Prox}_{f_2/\lambda_1} \mathrm{Prox}_{f_1/\lambda_2}) = \left\{ \frac{2}{2 + \lambda_1} \right\}.$$

For $y \in F$, i.e., $0 < y = 2/(2 + \lambda_1) < 1$,

$$\lambda_1 \mathrm{Prox}_{f_1/\lambda_2}(y) + \lambda_2 y = \lambda_1 0 + \lambda_2 \frac{2}{(2 + \lambda_1)} = \frac{2\lambda_2}{2 + \lambda_1} \in \mathrm{Fix}(\lambda_1 \mathrm{Prox}_{f_1} + \lambda_2 \mathrm{Prox}_{f_2}).$$

Again,

$$\mathrm{Fix}(\lambda_1 \operatorname{Prox}_{f_1} + \lambda_2 \operatorname{Prox}_{f_2}) = \{\lambda_1 \operatorname{Prox}_{f_1/\lambda_2}(y) + \lambda_2 y | \mid y \in F\}.$$

Finally, since

$$S = \{(x, \operatorname{Prox}_{f_2/\lambda_1}(x)) | \; x \in E\} = \left\{ \left(0, \frac{2}{2+\lambda_1}\right) \right\},$$

we have $\bar{\phi} = -2/(2+\lambda_1)$.

More examples can be constructed by using the proximal mapping calculus developed by Combettes and Wajs [19].

Remark 5.8 We conclude by pointing out that the situation for three or more functions (or sets if we work with indicator functions) is not clear at the moment. For instance, as pointed out by De Pierro and attributed to Iusem [20], one may have 3 sets such that least squares solutions exist but the existence of fixed points of compositions depends on the *order* of the projections. To fully understand these situations is an interesting topic for further research.

Acknowledgments

The authors thank a referee for insightful comments and additional bibliographical pointers. Xianfu Wang was partially supported by the Natural Sciences and Engineering Research Council of Canada. Heinz Bauschke was partially supported by the Natural Sciences and Engineering Research Council of Canada and by the Canada Research Chair Program.

References

[1] H. Attouch and M. Théra, A general duality principle for the sum of two operators, *J. Convex Anal.* 3 (1996), 1–24.

[2] S. Bartz, H. H. Bauschke, J. M. Borwein, S. Reich and X. Wang, Fitzpatrick functions, cyclic monotonicity and Rockafellar's antiderivative, *Nonlinear Anal.* 66 (2007), 1198–1223.

[3] H. H. Bauschke and J. M. Borwein, On the convergence of von Neumann's alternating projection algorithm for two sets, *Set-Valued Anal.* 1 (1993), 185–212.

[4] H. H. Bauschke and J. M. Borwein, Dykstra's alternating projection algorithm for two sets, *J. Approx. Theory* 79 (1994), 418–443.

[5] H. H. Bauschke and J. M. Borwein, On projection algorithms for solving convex feasibility problems, *SIAM Rev.* 38 (1996), 367–426.

[6] H. H. Bauschke, P. L. Combettes and D. R. Luke, Finding best approximation pairs relative to two closed convex sets in Hilbert spaces, *J. Approx. Theory* 127 (2004), 178–192.

[7] H. H. Bauschke, P. L. Combettes and S. Reich, The asymptotic behavior of the composition of two resolvents, *Nonlinear Anal.* 60 (2005), 283–301.

[8] H. H. Bauschke, R. Goebel, Y. Lucet, and X. Wang, The proximal average: basic theory, *SIAM J. Optim.* 19 (2008), 766–785.

[9] H. H. Bauschke, Y. Lucet and M. Trienis, How to transform one convex function continuously into another, *SIAM Rev.* 50 (2008), 115–132.

[10] H. H. Bauschke, E. Matoušková and S. Reich, Projection and proximal point methods: convergence results and counterexamples, *Nonlinear Anal.* 56 (2004), 715–738.

[11] H. H. Bauschke, S. M. Moffat, and X. Wang, The resolvent average for positive semidefinite matrices, *Linear Algebra Appl.* 432 (2010), 1757–1771.

[12] H. H. Bauschke and X. Wang, The kernel average for two convex functions and its application to the extension and representation of monotone operators, *Trans. Amer. Math. Soc.* 361 (2009), 5947–5965.

[13] H. H. Bauschke, X. Wang, and C. J. S. Wylie, Fixed points of averages of resolvents: geometry and algorithms, preprint, `http://arxiv.org/abs/1102.1478`, February 2011.

[14] R. E. Bruck and S. Reich, Nonexpansive projections and resolvents of accretive operators in Banach spaces, *Houston J. Math.* 3 (1977), 459–470.

[15] C. L. Byrne, *Signal Processing: A Mathematical Approach*, A.K. Peters, 2005.

[16] W. Cheney and A. A. Goldstein, Proximity maps for convex sets, *Proc. Amer. Math. Soc.* 10 (1959), 448–450.

[17] P. L. Combettes, Inconsistent signal feasibility problems: least-squares solutions in a product space, *IEEE Trans. Signal Processing* 42 (1994), 2955–2966.

[18] P. L. Combettes, Solving monotone inclusions via compositions of nonexpansive averaged operators, *Optimization* 53 (2004), 475–504.

[19] P. L. Combettes and V. R. Wajs, Signal recovery by proximal forward-backward splitting, *Multiscale Model. Simul.* 4 (2005), 1168–1200.

[20] A. R. De Pierro, From parallel to sequential projection methods and vice versa in convex feasibility: results and conjectures, in *Inherently Parallel Algorithms in Feasibility and Optimization and their Applications* (D. Butnariu, Y. Censor and S. Reich, Eds.) pp. 187–201. Elsevier, Amsterdam, 2001.

[21] D. R. Luke, Finding best approximation pairs relative to a convex and prox-regular set in a Hilbert space, *SIAM J. Optim.* 19 (2008), 714–739.

[22] B. Mercier, *Inéquations variationnelles de la mécanique*, Publications Mathématiques d'Orsay 80, vol. 1, Université de Paris-Sud Département de Mathématique, Orsay, 1980.

[23] G. J. Minty, Monotone (nonlinear) operators in Hilbert space, *Duke Math. J.* 29 (1962), 341–346.

[24] J. J. Moreau, Proximitet dualit dans un espace hilbertien, *Bull. Soc. Math. France* 93 (1965), 273–299.

[25] R. T. Rockafellar, On the maximality of sums of nonlinear monotone operators, *Trans. Amer. Math. Soc.* 149 (1970), 75–88.

[26] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

[27] R. T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM J. Control Optim.* 14 (1976), 877–898.

[28] R. T. Rockafellar and R. J-B Wets, *Variational Analysis*, Springer-Verlag, New York, 1998.

[29] S. Simons, *From Hahn-Banach to Monotonicity*, Lecture Notes in Mathematics, Vol. 1693, Springer-Verlag, 2008.

[30] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific Publishing, 2002.

[31] E. H. Zarantonello, Projections on convex sets in Hilbert space and spectral theory, in: *Contributions to Nonlinear Functional Analysis,* (E. H. Zarantonello, Ed.) pp. 237–424, Academic Press, New York, 1971.