



HAL
open science

Convex Minimization With Nonlinear Compositions

Luis M Briceño-Arias, Patrick L Combettes

► **To cite this version:**

Luis M Briceño-Arias, Patrick L Combettes. Convex Minimization With Nonlinear Compositions. 2021. hal-03441875

HAL Id: hal-03441875

<https://cnrs.hal.science/hal-03441875v1>

Preprint submitted on 22 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Convex Minimization With Nonlinear Compositions*

Luis M. Briceño-Arias¹ and Patrick L. Combettes²

¹Universidad Técnica Federico Santa María, Departamento de Matemática, Santiago, Chile

luis.briceno@usm.cl

²North Carolina State University, Department of Mathematics, Raleigh, NC 27695-8205, USA

plc@math.ncsu.edu

Abstract. We investigate the duality properties of a minimization problem involving the sum of a nonlinearly composed convex function and a linearly composed convex function. A Kuhn–Tucker operator is constructed for this problem as an extension of the operator found in classical Fenchel–Rockafellar duality theory. Monotone splitting algorithms are applied to this Kuhn–Tucker operator to solve the composite problem.

*Contact author: P. L. Combettes. Email: plc@math.ncsu.edu. Phone: +1 919 515 2671. The work of L. M. Briceño-Arias was supported by Centro de Modelamiento Matemático (CMM), ACE210010 and FB210005, BASAL funds for centers of excellence and FONDECYT 1190871 from ANID-Chile, and the work of P. L. Combettes was supported by the National Science Foundation under grant DMS-1818946.

1 Introduction

The goal of this paper is to investigate the following minimization problem, and to propose splitting algorithms to solve it. Given a real Banach space \mathcal{X} , $\Gamma_0(\mathcal{X})$ designates the class of proper lower semicontinuous convex functions from \mathcal{X} to $] -\infty, +\infty]$.

Problem 1.1 Let \mathcal{X} and \mathcal{Y} be reflexive real Banach spaces, let $f \in \Gamma_0(\mathcal{X})$, let $g \in \Gamma_0(\mathcal{Y})$, let $L: \mathcal{X} \rightarrow \mathcal{Y}$ be a bounded linear operator, and let $\phi \in \Gamma_0(\mathbb{R})$ be increasing. Set

$$\phi \circ f: \mathcal{X} \rightarrow] -\infty, +\infty] : x \mapsto \begin{cases} \phi(f(x)), & \text{if } f(x) \in \text{dom } \phi; \\ +\infty, & \text{if } f(x) \notin \text{dom } \phi. \end{cases} \quad (1.1)$$

The goal is to

$$\underset{x \in \mathcal{X}}{\text{minimize}} \quad \phi(f(x)) + g(Lx), \quad (1.2)$$

and \mathcal{P} denotes its set of solutions.

To motivate this formulation, let us consider a few special cases of interest (see Section 2 for notation).

Example 1.2 In Problem 1.1 suppose that $\phi = \iota_{]-\infty, 0]}$. Then (1.1) reduces to the constrained minimization problem

$$\underset{\substack{x \in \mathcal{X} \\ f(x) \leq 0}}{\text{minimize}} \quad g(Lx), \quad (1.3)$$

which is pervasive in nonlinear programming.

Example 1.3 Let $\theta \in \Gamma_0(\mathbb{R})$ be an increasing function such that $\text{dom } \theta =] -\infty, \eta[$ for some $\eta \in \mathbb{R}$, $\lim_{\xi \uparrow \eta} \theta(\xi) = +\infty$, and $(\text{rec } \theta)(1) > 0$. Let $\alpha:]0, +\infty[\rightarrow]0, +\infty[$ be such that $\lim_{\rho \downarrow 0} \alpha(\rho) = 0$ and $\underline{\lim}_{\rho \downarrow 0} \alpha(\rho)/\rho > 0$. Set $\mathcal{X} = \mathbb{R}^N$ and $L = \text{Id}$. Given $\rho \in]0, +\infty[$, set $\phi: \xi \mapsto \alpha(\rho)\theta(\xi/\rho)$. Then Problem 1.1 becomes

$$\underset{x \in \mathbb{R}^N}{\text{minimize}} \quad \alpha(\rho)\theta(f(x)/\rho) + g(x). \quad (1.4)$$

The asymptotic behavior of this family of penalty-barrier minimization problems as $\rho \downarrow 0$ is investigated in [1].

Example 1.4 In Problem 1.1 suppose that $\phi = \theta \circ \max\{0, \cdot - \rho\}$, where $\theta \in \Gamma_0(\mathbb{R})$ is increasing and $\rho \in \mathbb{R}$. Then (1.1) reduces to

$$\underset{x \in \mathcal{X}}{\text{minimize}} \quad \theta(\max\{0, f(x) - \rho\}) + g(Lx). \quad (1.5)$$

For instance, if C is a nonempty closed convex subset of \mathcal{X} and $f = d_C$, we recover the setting of [14, Example 2.4].

Example 1.5 Let $p \in [1, +\infty[$ and let $(\mathcal{X}_i)_{i \in I}$ and $(\mathcal{Y}_k)_{k \in K}$ be finite families of reflexive real Banach spaces. For every $i \in I$, let C_i be a nonempty closed convex subset of \mathcal{X}_i , and, for every $k \in K$, let $L_{ik} : \mathcal{X}_i \rightarrow \mathcal{Y}_k$ be a bounded linear operator. Set $\mathcal{X} = \bigoplus_{i \in I} \mathcal{X}_i$, $\mathcal{Y} = \bigoplus_{k \in K} \mathcal{Y}_k$, $C = \times_{i \in I} C_i$, $f : (x_i)_{i \in I} \mapsto (\sum_{i=1}^p d_{C_i}^p(x_i))^{1/p}$, $\phi = (\max\{0, \cdot\})^p$, $g : (y_k)_{k \in K} \mapsto \sum_{k \in K} g_k(y_k)$, and $L : (x_i)_{i \in I} \mapsto (\sum_{i \in I} L_{ik} x_i)_{k \in K}$. Then (1.1) reduces to

$$\underset{(x_i)_{i \in I} \in \times_{i \in I} \mathcal{X}_i}{\text{minimize}} \sum_{i \in I} d_{C_i}^p(x_i) + \sum_{k \in K} g_k \left(\sum_{i \in I} L_{ik} x_i \right). \quad (1.6)$$

This formulation covers problems in signal processing and location problems [4, 14, 21].

There is a vast literature on Problem 1.1 in the case of linear compositions, that is, when $\phi : t \mapsto t$. In this context, the duality theory goes back to [23], and various solution methods are available, e.g., [5, 9, 15, 30]. In the nonlinear setting, in terms of convex analysis, the conjugate and the subdifferential of $\phi \circ f$ in (1.1) have been derived in [11, 12], building up on the work of [18, 19, 17, 27]. However, duality theory for the minimization problem (1.1) does not seem to have been studied. On the numerical side, in the finite dimensional setting, with f smooth and $g = 0$, Problem 1.1 has been studied in [3] in the case when f is vector-valued by linearizing the objective function (see also [8, 20, 22, 31] and the references therein for the case when f is not convex in this scenario). However, in the general setting of Problem 1.1, solution methods are not available. The goal of the present paper is to address these gaps by first extending the classical Fenchel–Rockafellar theory to Problem 1.1, and then exploiting it to design proximal splitting algorithms to solve it.

We introduce our notation in Section 2. In Section 3, we develop a duality theory for Problem 1.1. In particular, we establish a connection between the conjugate of composite functions and perspective functions. We then derive a dual problem and show that primal-dual solutions can be obtained as the zeros of a nonlinear monotone Kuhn–Tucker operator. This property is used in Section 4 to derive splitting algorithms that use ϕ , f , g , and L separately to solve Problem 1.1 via the computation of resolvents of suitable monotone operators.

2 Notation and background

Let \mathcal{X} and \mathcal{Y} be reflexive real Banach spaces, and let \mathcal{X}^* and \mathcal{Y}^* denote their respective topological duals. The symbol $\mathcal{X} \oplus \mathcal{Y}$ designates the standard vector space $\mathcal{X} \times \mathcal{Y}$ equipped with the pairing

$$(\forall (x, y) \in \mathcal{X} \times \mathcal{Y})(\forall (x^*, y^*) \in \mathcal{X}^* \times \mathcal{Y}^*) \quad \langle (x, y), (x^*, y^*) \rangle = \langle x, x^* \rangle + \langle y, y^* \rangle \quad (2.1)$$

and the norm

$$(\forall (x, y) \in \mathcal{X} \times \mathcal{Y}) \quad \|(x, y)\| = \sqrt{\|x\|^2 + \|y\|^2}. \quad (2.2)$$

The power set of \mathcal{X}^* is denoted by $2^{\mathcal{X}^*}$. Let $A : \mathcal{X} \rightarrow 2^{\mathcal{X}^*}$ be a set-valued operator. We denote by $\text{ran } A = \{x^* \in \mathcal{X}^* \mid (\exists x \in \mathcal{X}) x^* \in Ax\}$ the range of A , by $\text{dom } A = \{x \in \mathcal{X} \mid Ax \neq \emptyset\}$ the domain of A , by $\text{zer } A = \{x \in \mathcal{X} \mid 0 \in Ax\}$ the set of zeros of A , by $\text{gra } A = \{(x, x^*) \in \mathcal{X} \times \mathcal{X}^* \mid x^* \in Ax\}$ graph of A , and by A^{-1} the inverse of A , which has graph $\{(x^*, x) \in \mathcal{X}^* \times \mathcal{X} \mid x^* \in Ax\}$. Moreover, A is monotone if

$$(\forall (x, y) \in \mathcal{X} \times \mathcal{X})(\forall (x^*, y^*) \in Ax \times Ay) \quad \langle x - y, x^* - y^* \rangle \geq 0, \quad (2.3)$$

and maximally so if there exists no monotone operator $B: \mathcal{X} \rightarrow 2^{\mathcal{X}^*}$ such that $\text{gra } A \subset \text{gra } B \neq \text{gra } A$. If \mathcal{X} is Hilbertian, the resolvent of A is $J_A = (\text{Id} + A)^{-1}$.

A function $f: \mathcal{X} \rightarrow [-\infty, +\infty]$ is proper if $-\infty \notin f(\mathcal{X}) \neq \{+\infty\}$. The domain of $f: \mathcal{X} \rightarrow [-\infty, +\infty]$ is $\text{dom } f = \{x \in \mathcal{X} \mid f(x) < +\infty\}$ and its epigraph is $\text{epi } f = \{(x, \xi) \in \mathcal{X} \times \mathbb{R} \mid f(x) \leq \xi\}$. Let $f \in \Gamma_0(\mathcal{X})$. The recession function of f is $\text{rec } f$, the conjugate of f is the function $f^* \in \Gamma_0(\mathcal{X}^*)$ defined by $f^*: x^* \mapsto \sup_{x \in \mathcal{X}} (\langle x, x^* \rangle - f(x))$, and the perspective of f is the function $\tilde{f} \in \Gamma_0(\mathcal{X} \oplus \mathbb{R})$ defined by

$$\tilde{f}: \mathcal{X} \times \mathbb{R} \rightarrow]-\infty, +\infty] : (x, \xi) \mapsto \begin{cases} \xi f(x/\xi), & \text{if } \xi > 0; \\ (\text{rec } f)(x), & \text{if } \xi = 0; \\ +\infty, & \text{otherwise.} \end{cases} \quad (2.4)$$

If f is proper, its subdifferential is the maximally monotone operator

$$\partial f: \mathcal{X} \rightarrow 2^{\mathcal{X}^*} : x \mapsto \{x^* \in \mathcal{X}^* \mid (\forall y \in \text{dom } f) \langle y - x, x^* \rangle + f(x) \leq f(y)\}. \quad (2.5)$$

Let C be a convex subset of \mathcal{X} . The indicator function of C is denoted by ι_C , the strong relative interior of C , i.e., the set of points $x \in C$ such that the cone generated by $-x + C$ is a closed vector subspace of \mathcal{X} , by $\text{sri } C$, and the distance to C is the function $d_C: x \mapsto \inf_{y \in C} \|x - y\|$. If \mathcal{X} is Hilbertian, for every $x \in \mathcal{X}$, $\text{prox}_f x = J_{\partial f} x$ denotes the unique minimizer of $f + \|\cdot - x\|^2/2$.

For background on convex analysis and monotone operators, see [2, 32].

3 Duality theory for nonlinear composite minimization

We start with a technical fact.

Lemma 3.1 *Let $\phi: \mathbb{R} \rightarrow]-\infty, +\infty]$ be an increasing proper convex function. Then the following hold:*

- (i) $\text{dom } \phi$ is an interval and $\inf \text{dom } \phi = -\infty$.
- (ii) $\text{dom } \phi^* \subset [0, +\infty[$.

Proof. (i): Since ϕ is convex, $\text{dom } \phi$ is convex, hence an interval. Now take $\xi \in \text{dom } \phi$ and $\eta < \xi$. Then $\phi(\eta) \leq \phi(\xi) < +\infty$ and hence $\eta \in \text{dom } \phi$. Consequently, $\inf \text{dom } \phi = -\infty$.

(ii): Since ϕ is increasing, it follows from (i) that

$$(\forall \xi^* \in]-\infty, 0[) \quad \phi^*(\xi^*) = \sup_{\xi \in \text{dom } \phi} (\xi \xi^* - \phi(\xi)) = +\infty, \quad (3.1)$$

which yields $\text{dom } \phi^* \subset [0, +\infty[$. \square

The next result concerns convex analytical properties of $\phi \circ f$ and establishes in particular a connection between the conjugate of a composite function and function and the marginal of a function that involves the perspective (see (2.4)) of its conjugate.

Proposition 3.2 *Let \mathcal{X} be a real Banach space, let $f \in \Gamma_0(\mathcal{X})$, and let $\phi \in \Gamma_0(\mathbb{R})$ be an increasing function such that $\text{dom } \phi \cap f(\text{dom } f) \neq \emptyset$. Then*

(i) $\text{dom}(\phi \circ f) = f^{-1}(\text{dom} \phi)$.

(ii) $\phi \circ f \in \Gamma_0(\mathcal{X})$.

(iii) Suppose that there exists $x \in \mathcal{X}$ such that $f(x) \in \text{int dom } \phi$. Then

$$(\forall x^* \in \mathcal{X}^*) \quad (\phi \circ f)^*(x^*) = \min_{\xi^* \in \mathbb{R}} (\phi^*(\xi^*) + \widetilde{f}^*(x^*, \xi^*)). \quad (3.2)$$

(iv) Suppose that there exists $z \in \text{dom } f$ such that $f(z) \in \text{int dom } \phi$. Then

$$(\forall x \in \mathcal{X}) \quad \partial(\phi \circ f)(x) = \bigcup_{\xi^* \in \partial\phi(f(x))} \xi^* \partial f(x). \quad (3.3)$$

(v) Suppose that there exists $z \in \text{dom } f$ such that $f(z) \in \text{int dom } \phi$, and let $x \in \mathcal{X}$ and $\xi^* \in \mathbb{R}$. Then

$$\begin{cases} \xi^* \in \partial\phi(f(x)) \\ x^* \in \xi^* \partial f(x) \end{cases} \Leftrightarrow \begin{cases} x^* \in \partial(\phi \circ f)(x) \\ \xi^* \in \text{Argmin}(\phi^* + \widetilde{f}^*(x^*, \cdot)). \end{cases} \quad (3.4)$$

Proof. (i): See (1.1).

(ii): In view of (1.1), convexity is established as in [2, Proposition 8.21], lower semicontinuity follows from [10, Proposition II.8.4], and properness follows from (i).

(iii): It follows from [11, Proposition 4.11ii)] and Lemma 3.1(ii) that

$$\begin{aligned} (\forall x^* \in \mathcal{X}^*) \quad (\phi \circ f)^*(x^*) &= \min_{\xi^* \in [0, +\infty[} (\phi^*(\xi^*) + (\xi^* f)^*(x^*)) \\ &= \min_{\xi^* \in \mathbb{R}} (\phi^*(\xi^*) + \widetilde{f}^*(x^*, \xi^*)), \end{aligned} \quad (3.5)$$

which completes the proof.

(iv): See [11, Proposition 4.11i)].

(v): It follows from (iv) and the Fenchel–Young identity that

$$\begin{aligned} \begin{cases} \xi^* \in \partial\phi(f(x)) \\ x^* \in \xi^* \partial f(x) \end{cases} &\Leftrightarrow (\exists z^* \in \mathcal{X}^*) \begin{cases} (\xi^*, z^*) \in \partial\phi(f(x)) \times \partial f(x) \\ x^* = \xi^* z^* \in \partial(\phi \circ f)(x) \end{cases} \\ &\Leftrightarrow (\exists z^* \in \mathcal{X}^*) \begin{cases} \phi(f(x)) + \phi^*(\xi^*) = \xi^* f(x) \\ f(x) + f^*(z^*) = \langle x, z^* \rangle \\ x^* = \xi^* z^* \in \partial(\phi \circ f)(x) \end{cases} \\ &\Leftrightarrow (\exists z^* \in \mathcal{X}^*) \begin{cases} (\phi \circ f)(x) + \phi^*(\xi^*) = \xi^* (\langle x, z^* \rangle - f^*(z^*)) \\ f(x) + f^*(z^*) = \langle x, z^* \rangle \\ x^* = \xi^* z^* \in \partial(\phi \circ f)(x) \end{cases} \\ &\Leftrightarrow (\exists z^* \in \mathcal{X}^*) \begin{cases} (\phi \circ f)(x) + \phi^*(\xi^*) + \xi^* f^*(z^*) = \langle x, x^* \rangle \\ f(x) + f^*(z^*) = \langle x, z^* \rangle \\ x^* = \xi^* z^* \in \partial(\phi \circ f)(x). \end{cases} \end{aligned} \quad (3.6)$$

Let $\mu = (\phi \circ f)^*(x^*) = \min(\phi^* + \widetilde{f}^*(x^*, \cdot))$ and note that, by Fenchel–Young inequality [32, Theorem 2.3.1(ii)], we have

$$\mu \geq \langle x^*, x \rangle - (\phi \circ f)(x). \quad (3.7)$$

We consider two cases.

- $\xi^* = 0$: In this case $x^* = 0$, $\mu \leq \phi^*(0) + \widetilde{f}^*(0, 0) = \phi^*(0)$, and (3.7) yields $-(\phi \circ f)(x) \leq \mu$. Altogether,

$$(3.6) \Leftrightarrow (\exists z^* \in \mathcal{X}^*) \begin{cases} \mu \leq \phi^*(0) \\ (\phi \circ f)(x) + \phi^*(0) = 0 \\ f(x) + f^*(z^*) = \langle x, z^* \rangle \end{cases} \quad (3.8)$$

$$\Leftrightarrow (\exists z^* \in \mathcal{X}^*) \begin{cases} \mu \leq \phi^*(0) = -(\phi \circ f)(x) \\ f(x) + f^*(z^*) = \langle x, z^* \rangle \\ (\phi \circ f)(x) + \mu \geq 0 \end{cases}$$

$$\Leftrightarrow (\exists z^* \in \mathcal{X}^*) \begin{cases} (\phi \circ f)(x) = -\phi^*(0) \\ f(x) + f^*(z^*) = \langle x, z^* \rangle \\ \mu = \phi^*(0), \end{cases} \quad (3.9)$$

which is equivalent to $0 \in \text{Argmin}(\phi^* + \widetilde{f}^*(0, \cdot))$.

- $\xi^* \neq 0$: In this case, it follows from Lemma 3.1(ii) that $\xi^* > 0$. Therefore, (3.7) yields

$$(3.6) \Leftrightarrow (\exists z^* \in \mathcal{X}^*) \begin{cases} \phi^*(\xi^*) + \widetilde{f}^*(x^*, \xi^*) = \langle x, x^* \rangle - (\phi \circ f)(x) \leq \mu \\ f(x) + f^*(z^*) = \langle x, z^* \rangle \\ x^* = \xi^* z^*, \end{cases} \quad (3.10)$$

and the result follows.

□

Next, we provide a characterization of the solutions to Problem 1.1 in terms of a monotone inclusion problem in $\mathcal{X} \times \mathbb{R}$. It will necessitate the following qualification conditions.

Assumption 3.3 *In Problem 1.1,*

$$0 \in \text{sri}(L(f^{-1}(\text{dom } \phi)) - \text{dom } g) \quad \text{and} \quad (\exists z \in \text{dom } f) \quad f(z) \in \text{int dom } \phi. \quad (3.11)$$

Proposition 3.4 *Consider the setting of Problem 1.1 under Assumption 3.3, and set*

$$\begin{cases} \mathbf{A}: \mathcal{X} \oplus \mathbb{R} \rightarrow 2^{\mathcal{X}^* \oplus \mathbb{R}}: (x, \xi^*) \mapsto \begin{cases} \xi^* \partial f(x) \times (\partial \phi^*(\xi^*) - f(x)), & \text{if } (x, \xi^*) \in \text{dom } \partial f \times \text{dom } \partial \phi^*; \\ \emptyset, & \text{otherwise.} \end{cases} \\ \mathbf{B}: \mathcal{X} \oplus \mathbb{R} \rightarrow 2^{\mathcal{X}^* \oplus \mathbb{R}}: (x, \xi^*) \mapsto (L^*(\partial g(Lx))) \times \{0\}. \end{cases} \quad (3.12)$$

Then the following hold:

- (i) $\text{dom } \mathbf{A} = \text{dom } \partial f \times \text{dom } \partial \phi^* \subset \text{dom } f \times [0, +\infty[.$
- (ii) \mathbf{A} and \mathbf{B} are maximally monotone.
- (iii) $\mathcal{P} = \bigcup_{\xi^* \in \mathbb{R}} \{x \in \mathcal{X} \mid (x, \xi^*) \in \text{zer}(\mathbf{A} + \mathbf{B})\}.$

Proof. (i): This follows from (3.12) and Lemma 3.1(ii).

(ii): Set

$$F: \mathcal{X} \times \mathbb{R} \rightarrow [-\infty, +\infty]: (x, \xi^*) \mapsto \begin{cases} +\infty, & \text{if } x \notin \text{dom } f; \\ \xi^* f(x) - \phi^*(\xi^*), & \text{if } x \in \text{dom } f \text{ and } \xi^* \in \text{dom } \phi^*; \\ -\infty, & \text{if } x \in \text{dom } f \text{ and } \xi^* \notin \text{dom } \phi^*. \end{cases} \quad (3.13)$$

Note that, for every $(x, \xi^*) \in \mathcal{X} \times \mathbb{R}$, $-F(x, \cdot) \in \Gamma_0(\mathbb{R})$ and, in view of Lemma 3.1(ii), $F(\cdot, \xi^*) \in \Gamma_0(\mathcal{X})$. As a result, the associated saddle operator \mathbf{A} is maximally monotone [24, 25]. On the other hand, it follows from Assumption 3.3 and [32, Theorem 2.8.3(vii)] that $L^* \circ \partial g \circ L = \partial(g \circ L)$ is maximally monotone, which implies that \mathbf{B} is maximally monotone.

(iii): Since $f^{-1}(\text{int dom } \phi) \neq \emptyset$, it follows from [32, Theorems 2.8.3(vii)], Proposition 3.2(iv), and [32, Theorem 2.4.2(iii)] that

$$\begin{aligned} x \in \mathcal{P} &\Leftrightarrow 0 \in \partial(\phi \circ f + g \circ L)(x) \\ &\Leftrightarrow (\exists \xi^* \in \partial \phi(f(x))) \quad 0 \in \xi^* \partial f(x) + L^*(\partial g(Lx)) \end{aligned} \quad (3.14)$$

$$\Leftrightarrow (\exists \xi^* \in \mathbb{R}) \quad \begin{cases} 0 \in \partial \phi^*(\xi^*) - f(x) \\ 0 \in \xi^* \partial f(x) + L^*(\partial g(Lx)), \end{cases} \quad (3.15)$$

and the result follows. \square

To advance further our the investigation of Problem 1.1, we introduce an auxiliary problem.

Problem 3.5 Let \mathcal{X} and \mathcal{Y} be reflexive real Banach spaces, let $f \in \Gamma_0(\mathcal{X})$, let $g \in \Gamma_0(\mathcal{Y})$, let $L: \mathcal{X} \rightarrow \mathcal{Y}$ be a bounded linear operator, and let $\phi \in \Gamma_0(\mathbb{R})$ be increasing. The goal is to

$$\underset{\substack{\xi^* \in \mathbb{R} \\ y^* \in \mathcal{Y}^*}}{\text{minimize}} \quad \phi^*(\xi^*) + \widetilde{f^*}(-L^* y^*, \xi^*) + g^*(y^*), \quad (3.16)$$

and \mathcal{D} denotes its set of solutions.

Proposition 3.6 Consider the setting of Problem 3.5. Then the following hold:

- (i) The Fenchel–Rockafellar of Problem 3.5 in $\mathcal{X} \oplus \mathbb{R}$ is

$$\underset{(x, \xi) \in \text{epi } f}{\text{minimize}} \quad \phi(\xi) + g(Lx). \quad (3.17)$$

- (ii) Let $\bar{x} \in f^{-1}(\text{dom } \phi)$. Then \bar{x} solves Problem 1.1 if and only if $(\bar{x}, f(\bar{x}))$ solves (3.17).

Proof. (i): Set

$$\begin{cases} \Phi: \mathcal{Y}^* \times \mathbb{R} \rightarrow]-\infty, +\infty]: (y^*, \xi^*) \mapsto g^*(y^*) + \phi^*(\xi^*) \\ \Psi: \mathcal{X}^* \times \mathbb{R} \rightarrow]-\infty, +\infty]: (x^*, \xi^*) \mapsto \widetilde{f^*}(x^*, -\xi^*) \\ \Lambda: \mathcal{Y}^* \times \mathbb{R} \rightarrow \mathcal{X}^* \times \mathbb{R}: (y^*, \xi^*) \mapsto (-L^* y^*, -\xi^*) \\ C = \{(x, \xi) \in \mathcal{X} \times \mathbb{R} \mid f(x) + \xi \leq 0\}. \end{cases} \quad (3.18)$$

We have $\Phi^*: \mathcal{Y} \times \mathbb{R} \rightarrow]-\infty, +\infty]: (y, \xi) \mapsto g(y) + \phi(\xi)$ and $\Lambda^*: \mathcal{X} \times \mathbb{R} \rightarrow \mathcal{Y} \times \mathbb{R}: (x, \xi) \mapsto (-Lx, -\xi)$. Moreover, $\Psi^*: (x, \xi) \mapsto \iota_C(x, -\xi) = \iota_{\text{epi } f}(x, \xi)$. Hence, we rewrite (3.16) as

$$\underset{(y^*, \xi^*) \in \mathcal{Y}^* \times \mathbb{R}}{\text{minimize}} \quad \Phi(y^*, \xi^*) + \Psi(\Lambda(y^*, \xi^*)), \quad (3.19)$$

and its Fenchel–Rockafellar dual is

$$\underset{(x, \xi) \in \mathcal{X} \times \mathbb{R}}{\text{minimize}} \quad \Phi^*(-\Lambda^*(x, \xi)) + \Psi^*(x, \xi), \quad (3.20)$$

which is precisely (3.17).

(ii): Since ϕ is increasing, it follows from (1.1) that

$$(\forall x \in f^{-1}(\text{dom } \phi)) \quad \phi(f(x)) = \min_{\xi \in [f(x), +\infty[} \phi(\xi). \quad (3.21)$$

Hence, since $(\bar{x}, f(\bar{x})) \in \text{epi } f$, \bar{x} solves Problem 1.1 if and only if

$$\min_{x \in \mathcal{X}} \phi(f(x)) + g(Lx) = \phi(f(\bar{x})) + g(L\bar{x}) = \min_{(x, \xi) \in \text{epi } f} \phi(\xi) + g(Lx), \quad (3.22)$$

which is equivalent to saying that $(\bar{x}, f(\bar{x}))$ solves (3.17). \square

We are now ready to present the main result of this section, which connects Problems 1.1 and 3.5 with a monotone inclusion in $\mathcal{X} \oplus \mathbb{R} \oplus \mathcal{Y}^*$.

Theorem 3.7 Consider the framework of Problem 1.1 under Assumption 3.3, set $\mathcal{X} = \mathcal{X} \oplus \mathbb{R} \oplus \mathcal{Y}^*$, let A be as in (3.12), and set

$$\begin{cases} M: \mathcal{X} \rightarrow 2^{\mathcal{X}^*}: (x, \xi^*, y^*) \mapsto A(x, \xi^*) \times \partial g^*(y^*) \\ S: \mathcal{X} \rightarrow \mathcal{X}^*: (x, \xi^*, y^*) \mapsto (L^*y^*, 0, -Lx). \end{cases} \quad (3.23)$$

Then the following hold:

- (i) The Fenchel–Rockafellar dual of Problem 1.1 is represented by Problem 3.5 and $\mathcal{D} \neq \emptyset$.
- (ii) M is maximally monotone.
- (iii) S is skew and $\|L\|$ -Lipschitzian.
- (iv) $\text{zer}(M + S)$ is a closed convex subset of $\mathcal{P} \times \mathcal{D}$.
- (v) $\mathcal{P} = \bigcup_{(\xi^*, y^*) \in \mathbb{R} \times \mathcal{Y}^*} \{x \in \mathcal{X} \mid (x, \xi^*, y^*) \in \text{zer}(M + S)\}$.

Proof. (i): The Fenchel–Rockafellar dual of (1.2) is

$$\underset{y^* \in \mathcal{Y}^*}{\text{minimize}} \quad (\phi \circ f)^*(-L^*y^*) + g^*(y^*). \quad (3.24)$$

In view of Proposition 3.2(iii), this gives (3.16). In addition, (3.11) and Proposition 3.2(i) entail that

$$0 \in \text{sri}(L(\text{dom}(\phi \circ f)) - \text{dom } g), \quad (3.25)$$

and it therefore follows from [32, Corollary 2.8.5] that (3.24) admits a solution.

(ii): As seen in Proposition 3.4(ii), \mathbf{A} is maximally monotone. Therefore, it follows from [32, Theorems 2.3.3 and 3.2.8] that \mathbf{M} is maximally monotone. On the other hand, \mathbf{S} is maximally monotone by [28, Chapter 8].

(iii): It is clear that $\mathbf{S}^* = -\mathbf{S}$. Now let $\mathbf{x} = (x, \xi^*, y^*) \in \mathcal{X}$. Then (2.2) yields

$$\|\mathbf{S}\mathbf{x}\|^2 = \|L^*y^*\|^2 + \|Lx\|^2 \leq \|L\|^2\|\mathbf{x}\|^2, \quad (3.26)$$

which shows that $\|\mathbf{S}\| \leq \|L\|$. Conversely, suppose that $\|\mathbf{x}\| \leq 1$ and that $(\xi^*, y^*) = (0, 0)$. Then $\|Lx\| = \|\mathbf{S}\mathbf{x}\| \leq \|\mathbf{S}\|$ and, therefore, $\|L\| \leq \|\mathbf{S}\|$.

(iv): We derive from (ii), the skewness of \mathbf{S} , and [28, Theorem 42.1] that $\mathbf{M} + \mathbf{S}$ is maximally monotone. This implies that its inverse is maximally monotone and therefore that

$$\text{zer}(\mathbf{M} + \mathbf{S}) = (\mathbf{M} + \mathbf{S})^{-1}\mathbf{0} = \bigcap_{(y, y^*) \in \text{gra}(\mathbf{M} + \mathbf{S})} \{\mathbf{x} \in \mathcal{X} \mid \langle \mathbf{x} - \mathbf{y}, \mathbf{y}^* \rangle \leq 0\} \quad (3.27)$$

is closed and convex as an intersection of such sets. Now, let $(x, \xi^*, y^*) \in \text{zer}(\mathbf{M} + \mathbf{S})$, that is,

$$\begin{cases} 0 \in \partial\phi^*(\xi^*) - f(x) \\ 0 \in \xi^*\partial f(x) + L^*y^* \\ 0 \in \partial g^*(y^*) - Lx. \end{cases} \quad (3.28)$$

Then, in view of Proposition 3.2(iv),

$$\begin{cases} 0 \in \partial(\phi \circ f)(x) + L^*y^* \\ 0 \in \partial g^*(y^*) - Lx. \end{cases} \quad (3.29)$$

Hence, it follows from Proposition 3.2(ii) and [32, Corollary 2.8.5 and Theorem 2.4.4(iv)] that $x \in \mathcal{P}$ and y^* solves (3.24). Next, we derive from (3.28) and Proposition 3.2(v) that

$$-L^*y^* \in \partial(\phi \circ f)(x) \quad \text{and} \quad \xi^* \in \text{Argmin}(\phi^* + \widetilde{f}^*(-L^*y^*, \cdot)). \quad (3.30)$$

In addition, it results from Proposition 3.2(iii)

$$\phi^*(\xi^*) + \widetilde{f}^*(-L^*y^*, \xi^*) = \min(\phi^* + \widetilde{f}^*(-L^*y^*, \cdot))(\mathbb{R}) = (\phi \circ f)^*(-L^*y^*). \quad (3.31)$$

Therefore, (y^*, ξ^*) solves

$$\underset{y^* \in \mathcal{Y}^*}{\text{minimize}} \left(g^*(y^*) + \underset{\xi^* \in \mathbb{R}}{\text{minimize}} \phi^*(\xi^*) + \widetilde{f}^*(-L^*y^*, \xi^*) \right), \quad (3.32)$$

which shows that $(y^*, \xi^*) \in \mathcal{D}$.

(v): Suppose that $x \in \mathcal{P}$. Then it follows from Proposition 3.4(iii) that there exists $\xi^* \in \mathbb{R}$ such that

$$\begin{cases} 0 \in \partial\phi^*(\xi^*) - f(x) \\ 0 \in \xi^*\partial f(x) + L^*(\partial g(Lx)). \end{cases} \quad (3.33)$$

Thus, there exists $(\xi^*, y^*) \in \mathbb{R} \times \mathcal{Y}^*$ such that

$$\begin{cases} 0 \in \partial\phi^*(\xi^*) - f(x) \\ 0 \in \xi^*\partial f(x) + L^*y^* \\ 0 \in \partial g^*(y^*) - Lx, \end{cases} \quad (3.34)$$

which shows that

$$(x, \xi^*, y^*) \in \text{zer}(M + S). \quad (3.35)$$

To show the reverse inclusion, suppose that $(x, \xi^*, y^*) \in \text{zer}(M + S)$. Then, by (iv), $x \in \mathcal{P}$. \square

Remark 3.8 By analogy with the standard Fenchel–Rockafellar theory [26], where $\phi: t \mapsto t$, we call the operator $M + S$ in Theorem 3.7 the Kuhn–Tucker operator associated with Problem 1.1.

4 Proximal analysis and solution methods

Theorem 3.7 opens a path for solving Problem 1.1 and the dual Problem 3.5 by finding a zero of the Kuhn–Tucker operator of Remark 3.8, i.e., of the sum of a maximally monotone operator M and a monotone Lipschitzian operator S . In a Hilbertian setting, this can be achieved by splitting methods that involve the resolvent of M , which is computed below.

Proposition 4.1 *Suppose that \mathcal{X} and \mathcal{Y} are real Hilbert spaces. Let $x \in \mathcal{X}$, $\xi^* \in \mathbb{R}$, $y^* \in \mathcal{Y}$, and $\gamma \in]0, +\infty[$. Then the following hold:*

- (i) $J_M(x, \xi, y^*) = (J_A(\xi, x), \text{prox}_{g^*}y^*)$.
- (ii) *There exists a unique $\mu \in [0, +\infty[$ such that $\mu = \text{prox}_{\gamma\phi^*}(\xi^* + \gamma f(\text{prox}_{\mu\gamma f}x))$, and $J_{\gamma A}(x, \xi^*) = (\text{prox}_{\mu\gamma f}x, \mu)$.*

Proof. (i): This follows from Theorem 3.7(ii), Proposition 3.4(ii), and [2, Propositions 23.8 and 23.18].

(ii): Set $(p, \mu) \in \mathcal{X} \times \mathbb{R}$. Then

$$(p, \mu) = J_{\gamma A}(x, \xi^*) \Leftrightarrow (x, \xi^*) \in (p, \mu) + \gamma A(p, \mu) \quad (4.1)$$

$$\begin{aligned} &\Leftrightarrow \begin{cases} x \in p + \mu\gamma\partial f(p) \\ \xi^* \in \mu + \gamma\partial\phi^*(\mu) - \gamma f(p) \end{cases} \\ &\Leftrightarrow \begin{cases} p = \text{prox}_{\mu\gamma f}x \\ \mu = \text{prox}_{\gamma\phi^*}(\xi^* + \gamma f(p)) \end{cases} \\ &\Leftrightarrow \begin{cases} p = \text{prox}_{\mu\gamma f}x \\ \mu = \text{prox}_{\gamma\phi^*}(\xi^* + \gamma f(\text{prox}_{\mu\gamma f}x)). \end{cases} \end{aligned} \quad (4.2)$$

We observe that Proposition 3.4(i) implies that μ in (4.1) lies in $[0, +\infty[$. \square

Remark 4.2 Let $\gamma \in]0, +\infty[$, let $(x, \xi^*) \in \mathcal{X} \times \mathbb{R}$, and denote by $m(\gamma, x, \xi^*)$ the unique $\mu \in [0, +\infty[$ identified in Proposition 4.1. Then

$$m(\gamma, x, \xi^*) \in [0, +\infty[\text{ is the unique fixed point of the operator } \begin{aligned} T: \quad \mathbb{R} &\rightarrow \mathbb{R} \\ \mu &\mapsto \text{prox}_{\gamma\phi^*}(\xi^* + \gamma f(\text{prox}_{\mu\gamma f}x)). \end{aligned} \quad (4.3)$$

Note that, in view of [2, Propositions 12.27 and 24.31], T is decreasing.

In the following examples we provide the computation of m in (4.3) and $J_{\gamma\mathbf{A}}$, where $\gamma \in]0, +\infty[$ and \mathbf{A} is defined in (3.12).

Example 4.3 In the context of Example 1.2, suppose that \mathcal{X} is a real Hilbert space, let $\gamma \in]0, +\infty[$, let $(x, \xi^*) \in \mathcal{X} \times \mathbb{R}$, and let $\mu \in [0, +\infty[$ be the unique solution to

$$\mu = \begin{cases} \xi^* + \gamma f(\text{prox}_{\mu\gamma f}x), & \text{if } \xi^* + \gamma f(x) > 0; \\ 0, & \text{if } \xi^* + \gamma f(x) \leq 0. \end{cases} \quad (4.4)$$

Since $\phi^* = \iota_{[0, +\infty[}$, Proposition 4.1 and (4.3) yield

$$m(\gamma, x, \xi^*) = \mu \quad \text{and} \quad J_{\gamma\mathbf{A}}(x, \xi^*) = \begin{cases} (x, 0), & \text{if } \xi^* + \gamma f(x) \leq 0; \\ (\text{prox}_{\mu\gamma f}x, \mu), & \text{if } \xi^* + \gamma f(x) > 0. \end{cases} \quad (4.5)$$

Example 4.4 In Example 1.3, suppose that \mathcal{X} is a real Hilbert space, set $\alpha: \rho \mapsto \rho$, let $\gamma \in]0, +\infty[$, let $(x, \xi^*) \in \mathcal{X} \times \mathbb{R}$, and let $\mu \in [0, +\infty[$ be the unique solution to

$$\mu = \text{prox}_{\gamma\rho\theta^*}(\xi^* + \gamma f(\text{prox}_{\mu\gamma f}x)). \quad (4.6)$$

Then, by [2, Proposition 13.23(ii)], $\phi^* = \rho\theta^*$ and it follows from Proposition 4.1 and (4.3) that $m(\gamma, x, \xi^*) = \mu$ and that $J_{\gamma\mathbf{A}}(x, \xi^*) = (\text{prox}_{\mu\gamma f}x, \mu)$.

Example 4.5 In Example 1.4, let $\gamma \in]0, +\infty[$, let $(x, \xi^*) \in \mathcal{X} \times \mathbb{R}$ and let $\mu \in]0, +\infty[$ is the unique solution to

$$\mu = \begin{cases} \text{prox}_{\gamma\theta^*}(\xi^* + \gamma(f(\text{prox}_{\mu\gamma f}x) - \rho)), & \text{if } \xi^* + \gamma f(x) > \gamma\rho; \\ 0. & \text{if } \xi^* + \gamma f(x) \leq \gamma\rho. \end{cases} \quad (4.7)$$

We deduce from [14, Example 2.4] and [2, Proposition 24.8(ix)] that

$$\text{prox}_{\gamma\phi^*}: \xi^* \mapsto \begin{cases} \text{prox}_{\gamma\theta^*}(\xi^* - \gamma\rho), & \text{if } \xi^* > \gamma\rho; \\ 0, & \text{if } \xi^* \leq \gamma\rho. \end{cases} \quad (4.8)$$

Hence, Proposition 4.1 and (4.3) yield

$$m(\gamma, x, \xi^*) = \mu \quad \text{and} \quad J_{\gamma\mathbf{A}}(x, \xi^*) = \begin{cases} (x, 0), & \text{if } \xi^* + \gamma f(x) \leq \gamma\rho; \\ (\text{prox}_{\mu\gamma f}x, \mu), & \text{if } \xi^* + \gamma f(x) > \gamma\rho. \end{cases} \quad (4.9)$$

We now turn to the design of algorithms for solving Problems 1.1 and 3.5 using Theorem 3.7 and Proposition 4.1. The following approach is based on Tseng's splitting method [29].

Theorem 4.6 Consider the framework of Problem 1.1 under Assumption 3.3, and define m as in (4.3). Suppose that \mathcal{X} and \mathcal{Y} are real Hilbert spaces, and that $\mathcal{P} \neq \emptyset$. Let $x_0 \in \mathcal{X}$, $\xi_0 \in \mathbb{R}$, $y_0^* \in \mathcal{Y}$, let $\varepsilon \in]0, 1/(\|L\| + 1)[$, let $(\gamma_n)_{n \in \mathbb{N}}$ be a sequence in $[\varepsilon, (1 - \varepsilon)/\|L\|]$, and set

$$\begin{aligned} & \text{for } n = 0, 1, \dots \\ & \begin{cases} z_n = x_n - \gamma_n L^* y_n^* \\ z_n^* = y_n^* + \gamma_n L x_n \\ \xi_{n+1} = m(\gamma_n, z_n, \xi_n) \\ p_n = \text{prox}_{\xi_{n+1} \gamma_n f} z_n \\ p_n^* = \text{prox}_{\gamma_n g^*} z_n^* \\ q_n = p_n - \gamma_n L^* p_n^* \\ q_n^* = p_n^* + \gamma_n L p_n \\ x_{n+1} = x_n - z_n + q_n \\ y_{n+1}^* = y_n^* - z_n^* + q_n^* \end{cases} \end{aligned} \quad (4.10)$$

Then $x_n \rightarrow \bar{x}$, $y_n^* \rightarrow \bar{y}^*$, and $\xi_n \rightarrow \bar{\xi}$, where $\bar{x} \in \mathcal{P}$ and $(\bar{\xi}, \bar{y}^*) \in \mathcal{D}$.

Proof. Set $\mathcal{X} = \mathcal{X} \oplus \mathbb{R} \oplus \mathcal{Y}$. It follows from Theorem 3.7(v) that $\text{zer}(M + S) \neq \emptyset$. In addition, Theorem 3.7(ii)-(iii) implies that M is maximally monotone and S is monotone and $\|L\|$ -Lipschitzian. Now set, for every $n \in \mathbb{N}$, $\mathbf{x}_n = (x_n, \xi_n, y_n^*)$, $\mathbf{z}_n = (z_n, \xi_n, z_n^*)$, $\mathbf{p}_n = (p_n, \xi_{n+1}, p_n^*)$, and $\mathbf{q}_n = (q_n, \xi_{n+1}, q_n^*)$. Then, in view of (3.23) and Proposition 4.1, we can express (4.10) as

$$\begin{aligned} & \text{for } n = 0, 1, \dots \\ & \begin{cases} \mathbf{z}_n = \mathbf{x}_n - \gamma_n \mathbf{S} \mathbf{x}_n \\ \mathbf{p}_n = J_{\gamma_n M} \mathbf{z}_n \\ \mathbf{q}_n = \mathbf{p}_n - \gamma_n \mathbf{S} \mathbf{p}_n \\ \mathbf{x}_{n+1} = \mathbf{x}_n - \mathbf{z}_n + \mathbf{q}_n \end{cases} \end{aligned} \quad (4.11)$$

In turn, we derive from [29, Theorem 3.4] that $(\mathbf{x}_n)_{n \in \mathbb{N}}$ converges weakly to a point $\bar{\mathbf{x}} \in \text{zer}(M + S)$. In view of Theorem 3.7(iv), the proof is complete. \square

Remark 4.7

- (i) In particular, Examples 1.2-1.4 can be solved via Theorem 4.6. Note that in these settings $m(\gamma_n, z_n, \xi_n)$ in (4.10) is computed in Examples 4.3-4.5.
- (ii) In the particular case of Example 1.2, (1.3) is solved by (4.10). The algorithm activates the inequality constraint $f(x) \leq 0$ through $\text{prox}_{\xi_{n+1} \gamma_n f}$, where ξ_{n+1} is the solution of the scalar equation in (4.4). Note that general convex inequalities are hard to treat directly in the context of standard proximal methods since they involve the projection onto the 0-sublevel set of f , which is typically not explicit (see [6] for an alternative approach in the smooth case). We circumvent this problem by requiring only the proximity operator of f . For instance, if $\mathcal{X} = \mathbb{R}^N$, $p \in [1, +\infty[$, $\eta \in]0, +\infty[$, and $f = \|\cdot\|_p^p - \eta^p$, (1.3) reduces to

$$\begin{aligned} & \underset{\substack{x \in \mathbb{R}^N \\ \|x\|_p \leq \eta}}{\text{minimize}} \quad g(Lx), \end{aligned} \quad (4.12)$$

a formulation which arises in machine learning [16] and for which the projection is expensive to compute [13]. Within our framework, (4.12) can be solved by (4.10) for a general nonsmooth g , where $\xi_{n+1} = m(\gamma_n, z_n, \xi_n)$ is computed as in Example 4.3 and $\text{prox}_{\xi_{n+1} \gamma_n f} = (\text{prox}_{\xi_{n+1} \gamma_n |\cdot|_p})_{1 \leq i \leq N}$, which is computable in view of [2, Example 24.38].

Remark 4.8 In view of Theorem 3.7(iv), we have reduced Problem 1.1 and Problem 3.5 to finding a zero of the sum of a maximally monotone operator and a monotone Lipschitzian operator. Theorem 4.6 addresses this inclusion problem via the basic form of Tseng’s splitting method [29]. Let us add a few comments.

- (i) Errors can be incorporated in (4.11) by using the error-tolerant version of Tseng’s method [5, Theorem 2.5].
- (ii) Another method tailored to inclusions involving the sum of a maximally monotone operator and a monotone Lipschitzian operator is that of [7, Corollary 5.2], which can also incorporate inertial effects. Another advantage of this framework is that it features, through [7, Theorem 4.8], a strongly convergent variant which does not require any additional assumptions on the operators.
- (iii) A zero of $M+S$ can also be found by generic splitting methods which do not exploit specifically the Lipschitz continuity of S . For instance, the Douglas–Rachford algorithm can be employed; see [5, Remark 2.9] for an implementation with a skew operator similar to S in (3.23).

A noteworthy special case of Problem 1.1 is when $\mathcal{Y} = \mathcal{X}$ and $L = \text{Id}$. In this setting, Proposition 3.4 asserts that Problem 1.1 can be solved by finding a zero of the sum of the monotone operators \mathbf{A} and \mathbf{B} defined in (3.12). We can for instance use the Douglas–Rachford algorithm for this task, which leads to the following implementation.

Proposition 4.9 Consider the framework of Problem 1.1 under Assumption 3.3, and define m as in (4.3). Suppose that $\mathcal{X} = \mathcal{Y}$ is a real Hilbert space, that $L = \text{Id}$, and that $\mathcal{P} \neq \emptyset$. Let $z_0 \in \mathcal{X}$, $\eta_0 \in \mathbb{R}$, let $\gamma \in]0, +\infty[$, let $(\lambda_n)_{n \in \mathbb{N}}$ be a sequence on $]0, +\infty[$ such that $\sum_{n \in \mathbb{N}} \lambda_n(2 - \lambda_n) = +\infty$, and set

$$\begin{aligned} & \text{for } n = 0, 1, \dots \\ & \left\{ \begin{array}{l} \xi_n = m(\gamma, z_n, \eta_n) \\ x_n = \text{prox}_{\xi_n \gamma f} z_n \\ z_{n+1} = z_n + \lambda_n (\text{prox}_{\gamma g}(2x_n - z_n) - x_n) \\ \eta_{n+1} = \eta_n + \lambda_n (\xi_n - \eta_n). \end{array} \right. \end{aligned} \quad (4.13)$$

Then $x_n \rightharpoonup \bar{x}$, where $\bar{x} \in \mathcal{P}$.

Proof. It follows from Proposition 3.4(iii) that $\text{zer}(\mathbf{A} + \mathbf{B}) \neq \emptyset$, where \mathbf{A} and \mathbf{B} are defined in (3.12). Now set, for every $n \in \mathbb{N}$, $\mathbf{x}_n = (x_n, \xi_n) \in \mathcal{X} \oplus \mathbb{R}$ and $\mathbf{y}_n = (z_n, \eta_n) \in \mathcal{X} \oplus \mathbb{R}$. Note that [2, Proposition 23.18] yields $J_{\gamma \mathbf{B}}: (x, \xi) \mapsto (\text{prox}_{\gamma g} x, \xi)$. Hence, it follows from (3.12) and Proposition 4.1 that (4.13) can be written as

$$(\forall n \in \mathbb{N}) \quad \left\{ \begin{array}{l} \mathbf{x}_n = J_{\gamma \mathbf{A}} \mathbf{y}_n \\ \mathbf{y}_{n+1} = \mathbf{y}_n + \lambda_n (J_{\gamma \mathbf{B}}(2\mathbf{x}_n - \mathbf{y}_n) - \mathbf{x}_n). \end{array} \right. \quad (4.14)$$

In turn, [2, Theorem 26.11] yields $\mathbf{x}_n \rightharpoonup \bar{\mathbf{x}}$ for some $\bar{\mathbf{x}} \in \text{zer}(\mathbf{A} + \mathbf{B})$ and the result follows from Proposition 3.4(iii). \square

Remark 4.10 Although we have investigated Problem 1.1 with a single linear composite term, it also models formulations with several such terms. Indeed, let \mathcal{X} be a real reflexive Banach space, let $f \in \Gamma_0(\mathcal{X})$, let $\phi \in \Gamma_0(\mathbb{R})$ be increasing, and let $(\mathcal{Y}_k)_{k \in K}$ be a finite family of reflexive real Banach

spaces. For every $k \in K$, let $g_k \in \Gamma_0(\mathcal{Y}_k)$, let $L_k: \mathcal{X} \rightarrow \mathcal{Y}_k$ be linear and bounded. Set $\mathcal{Y} = \bigoplus_{k \in K} \mathcal{Y}_k$, $g: \mathcal{Y} \rightarrow]-\infty, +\infty]: (y_k)_{k \in K} \mapsto \sum_{k \in K} g_k(y_k)$, and $L: \mathcal{X} \rightarrow \mathcal{Y}: x \mapsto (L_k x)_{k \in K}$. Then (1.2) becomes

$$\underset{x \in \mathcal{X}}{\text{minimize}} \quad \phi(f(x)) + \sum_{k \in K} g_k(L_k x). \quad (4.15)$$

Furthermore, in the Hilbertian setting, the implementation of (4.10) is

$$\begin{array}{l} \text{for } n = 0, 1, \dots \\ \left[\begin{array}{l} z_n = x_n - \gamma_n \sum_{k \in K} L_k^* y_{k,n}^* \\ \xi_{n+1} = m(\gamma_n, z_n, \xi_n) \\ p_n = \text{prox}_{\xi_{n+1} \gamma_n f} z_n \\ \text{for every } k \in K \\ \left[\begin{array}{l} z_{k,n}^* = y_{k,n}^* + \gamma_n L_k x_n \\ p_{k,n}^* = \text{prox}_{\gamma_n g_k} z_{k,n}^* \\ q_{k,n}^* = p_{k,n}^* + \gamma_n L_k p_n \end{array} \right. \\ q_n = p_n - \gamma_n \sum_{k \in K} L_k^* p_{k,n}^* \\ x_{n+1} = x_n - z_n + q_n \\ \text{for every } k \in K \\ \left[\begin{array}{l} y_{k,n+1}^* = y_{k,n}^* - z_{k,n}^* + q_{k,n}^* \end{array} \right. \end{array} \right. \end{array} \quad (4.16)$$

and Theorem 4.6 provides conditions for the weak convergence of $(x_n)_{n \in \mathbb{N}}$ to a solution to (4.15).

References

- [1] A. Auslender, R. Cominetti, and M. Haddou Asymptotic analysis for penalty and barrier methods in convex and linear programming, *Math. Oper. Res.*, vol. 22, pp. 43–62, 1997.
- [2] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed. Springer-Verlag, New York, 2017.
- [3] J. Bolte, Z. Chen, and E. Pauwels, The multiproximal linearization method for convex composite problems, *Math. Program.*, vol. A182, pp. 1–36, 2020.
- [4] L. M. Briceño-Arias and P. L. Combettes, Convex variational formulation with smooth coupling for multicomponent signal decomposition and recovery, *Numer. Math. Theory Methods Appl.*, vol. 2, pp. 485–508, 2009.
- [5] L. M. Briceño-Arias and P. L. Combettes, A monotone+skew splitting model for composite monotone inclusions in duality, *SIAM J. Optim.*, vol. 21, pp. 1230–1250, 2011.
- [6] L. M. Briceño-Arias and D. Davis, Forward-backward-half forward algorithm for solving monotone inclusions, *SIAM J. Optim.*, vol. 28, pp. 2839–2871, 2018.
- [7] M. N. Bui and P. L. Combettes, Warped proximal iterations for monotone inclusions, *J. Math. Anal. Appl.*, vol. 491, art. 124315, 2020.
- [8] J. V. Burke, Descent methods for composite nondifferentiable optimization problems, *Math. Program.*, vol. 33, pp. 260–279, 1985.
- [9] G. Chen and M. Teboulle, A proximal-based decomposition method for convex minimization problems, *Math. Program.*, vol. 64, pp. 81–101, 1994.

- [10] G. Choquet, *Topologie*. Masson, Paris, 1964 (English translation: *Topology*. Academic Press, New York, 1966).
- [11] C. Combari, M. Laghdir, and L. Thibault, Sous-différentiels de fonctions convexes composées, *Ann. Sci. Math. Québec*, vol. 18, pp. 119–148, 1994.
- [12] C. Combari, M. Laghdir, and L. Thibault, A note on subdifferentials of convex composite functionals, *Arch. Math. (Basel)*, vol. 67, pp. 239–252, 1996.
- [13] C. W. Combettes and S. Pokutta, Complexity of linear minimization and projection on some sets, *Oper. Res. Lett.*, vol. 49, pp. 565–571, 2021.
- [14] P. L. Combettes and L. E. Glaudin, Proximal activation of smooth functions in splitting algorithms for convex image recovery, *SIAM J. Imaging Sci.*, vol. 12, pp. 1905–1935, 2019.
- [15] L. Condat, A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms, *J. Optim. Theory Appl.*, vol. 158, pp. 460–479, 2013.
- [16] M. Jaggi, Revisiting Frank-Wolfe: Projection-free sparse convex optimization, *Proc. 30th Int. Conf. Machine Learn.*, vol. 28, pp. 427–435, 2013.
- [17] S. S. Kutateladze, Formulas for computing subdifferentials, *Soviet. Math. Dokl.*, vol. 18, pp. 146–148, 1977.
- [18] B. Lemaire, Application of a subdifferential of a convex composite functional to optimal control in variational inequalities, *Lecture Notes in Economics and Mathematical Systems*, vol. 255, pp. 103–117, 1985.
- [19] C. Lescarret, Sous-différentialité de fonctions composées, *Séminaire d'Analyse Unilatérale*, exposé no. 2, Montpellier, France, 1968.
- [20] A. S. Lewis and S. J. Wright, A proximal method for composite minimization, *Math. Program.*, vol. 158, pp. 501–546, 2016.
- [21] B. S. Mordukhovich, N. M. Nam, and J. Salinas, Solving a generalized Heron problem by means of convex analysis, *Amer. Math. Monthly*, vol. 119, pp. 87–99, 2012.
- [22] E. Pauwels, The value function approach to convergence analysis in composite optimization, *Oper. Res. Lett.*, vol. 44, pp. 790–795, 2016.
- [23] R. T. Rockafellar, Duality and stability in extremum problems involving convex functions, *Pacific J. Math.*, vol. 21, pp. 167–187, 1967.
- [24] R. T. Rockafellar, Monotone operators associated with saddle-functions and minimax problems, in: *Nonlinear Functional Analysis, Part 1*, (F. E. Browder, ed.), pp. 241–250. AMS, Providence, RI, 1970.
- [25] R. T. Rockafellar, Saddle points and convex analysis, in: *Differential Games and Related Topics*, H. W. Kuhn and G. P. Szegö (eds.), pp. 109–127. North-Holland, New York, 1971.
- [26] R. T. Rockafellar, *Conjugate Duality and Optimization*. SIAM, Philadelphia, PA, 1974.
- [27] A. M. Rubinov, Sublinear operators and their applications, *Russian Math. Surveys*, vol. 32, pp. 115–175, 1977.
- [28] S. Simons, *Minimax and Monotonicity*. Springer-Verlag, New York, 1998.
- [29] P. Tseng, A modified forward-backward splitting method for maximal monotone mappings, *SIAM J. Control Optim.*, vol. 38, pp. 431–446, 2000.
- [30] B. C. Vũ, A splitting algorithm for dual monotone inclusions involving cocoercive operators, *Adv. Comput. Math.*, vol. 38, pp. 667–681, 2013.
- [31] S. J. Wright, Convergence of an inexact algorithm for composite nonsmooth optimization, *IMA J. Numer. Anal.*, vol. 10, pp. 299–321, 1990.
- [32] C. Zălinescu, *Convex Analysis in General Vector Spaces*. World Scientific Publishing, River Edge, NJ, 2002.