Inclusions Monotones en Dualité et Applications

Bang Cong Vu

THÈSE DE DOCTORAT DE
L'UNIVERSITÉ PIERRE ET MARIE CURIE – PARIS VI

Spécialité :

MATHÉMATIQUES APPLIQUÉES

Présentée par :

Bằng Công Vũ

pour l'obtention du grade de

DOCTEUR DE L'UNIVERSITÉ PIERRE ET MARIE CURIE – PARIS VI

Sujet de la thèse :

Inclusions Monotones en Dualité et Applications

Soutenue le 15 avril 2013 devant le jury composé de :

Alexandre d'ASPREMONT	*examinateur*
Hédy ATTOUCH	*rapporteur*
Jean-Bernard BAILLON	*examinateur*
Frédéric BONNANS	*examinateur*
Patrick L. COMBETTES	*Directeur de thèse*
Roberto COMINETTI	*rapporteur*
Sylvain SORIN	*président*

Laboratoire Jacques-Louis Lions – UMR 7598

Remerciements

Premièrement, je voudrais remercier mon directeur de thèse, Professeur Patrick Louis Combettes, pour son soutien, ses précieux conseils, sa disponibilité et sa patience au cours de ces dernières années. Professeur P. L. Combettes est pour moi une grande source d'inspiration pour que je puisse continuer à cultiver et enrichir mes connaissances. Je tiens par conséquent à exprimer ma profonde gratitude et ma sincère reconnaissance au Professeur P. L. Combettes.

Je suis très reconnaissant envers MM. Attouch et Cominetti d'avoir accepté d'être rapporteurs de ma thèse. Je tiens aussi à remercier tous les membres du jurys de me faire l'honneur d'accepter de participaer à la soutenance et pour l'intérêt qu'ils ont porté à ce travail.

En outre, j'ai reçu, au cours de ma thèse au Laboratoire Jacques-Louis Lions, Université Pierre et Marie Curie – Paris VI, une aide financière du Ministère de l'Éducation et de la Formation du Vietnam par le biais du Fonds de bourses d'études 322. Par conséquent, je tiens à remercier le Ministère de l'Éducation et de la Formation, en particulier au Ministre de la Formation à l'Étranger, Monsieur Nguyen Xuan Vang, et à des experts Bui Anh Ngoc, Bui Thi Hong Thuy. J'aimerais également remercier les Professeurs P. L. Combettes, Đinh Dũng et Pham Ky Anh pour leurs lettres de recommandation pour que je puisse obtenir cette bourse.

D'ailleurs, je tiens à remercier les membres du projet GdR MOA et du Laboratoire Jacques-Louis Lions pour leur assistance financière qui m'a permis d'assister à des conférences scientifiques et à des cours de langue française pendant les années 2011 et 2012.

De plus, je tiens également à exprimer ma gratitude à la Faculté des Mathématiques Computationnelles et des Mathématiques Appliquées, Département de Mathématiques, Université des sciences naturelles de l'Université Nationale de Hanoï pour son autorisation et son encouragement pour que je puisse faire mes études en France.

Enfin, j'aimerais remercier ma famille et mes amis qui m'ont soutenu et encouragé à poursuivre la recherche scientifique.

Paris, le 28 mars 2013

Bằng Công Vũ

Kính tặng Bố Mẹ !

Table des matières

Résumé

Inclusions Monotones en Dualité et Applications

Le but de cette thèse est de développer de nouvelles techniques d'éclatement d'opérateurs multivoques pour résoudre des problèmes d'inclusion monotone structurés dans des espaces hilbertiens. La dualité au sens des inclusions monotones tient une place essentielle dans ce travail et nous permet d'obtenir des décompositions qui ne seraient pas disponibles via une approche purement primale. Nous développons plusieurs algorithmes à métrique fixe ou variable dans un cadre unifié, et montrons en particulier que de nombreuses méthodes existantes sont des cas particuliers de la méthode explicite–implicite formulée dans des espaces produits adéquats. Les méthodes proposées sont appliquées aux problèmes d'inéquations variationnelles, aux problèmes de minimisation, aux problèmes inverses, aux problèmes de traitement du signal, aux problèmes d'admissibilité et aux problèmes de meilleure approximation. Dans un second temps, nous introduisons une notion de suite quasi-fejérienne à métrique variable et analysons ses propriétés asymptotiques. Ces résultats nous permettent d'obtenir des extensions de méthodes d'éclatement aux problèmes où la métrique varie à chaque itération.

Mots-clés : algorithme primal-dual, algorithme proximal, analyse convexe, cocoercivité, meilleure approximation, méthode explicite-implicite, méthode explicite-implicite-explicite, métrique variable, dualité, inclusions monotones, opérateur monotone, restauration d'images.

Abstract

Monotone Inclusions in Duality and Applications

The goal of this thesis is to develop new splitting techniques for set-valued operators to solve structured monotone inclusion problems in Hilbert spaces. Duality plays a central role in this work. It allows us to obtain decompositions which would not be available through a purely primal approach. We develop several fixed and variable metric algorithms in a unified framework, and show in particular that many existing methods are special cases of the forward-backward method formulated in a suitable product space. The proposed methods are applied to variational inequalities, minimization problems, inverse problems, signal processing problems, feasibility problems, and best approximation problems. Next, we introduce the notion of a variable metric quasi-Fejér sequence and analyze its asymptotic properties. These results allow us to obtain extensions of splitting schemes to problems in which the metric varies at each iteration.

Key words : best approximation, forward-backward method, forward-backward-forward method, image recovery, monotone operators, operator splitting, primal-dual algorithm, proximal algorithm, signal theory, variable metric.

Notations et Glossaire

Les notations suivantes seront utilisées dans toute la thèse. De plus, nous rappelons certaines définitions de base en analyse convexe.

Notations générales

- $\mathcal{H}, \mathcal{G}, \mathcal{G}_1, \ldots, \mathcal{G}_m$: Espaces de Hilbert réels.
- $\langle \cdot \mid \cdot \rangle$: Produit scalaire et norme de l'espace $\mathcal{H}$.
- $\|\cdot\|$: Norme de l'espace $\mathcal{H}$.
- Id : Opérateur identité sur $\mathcal{H}$.
- $2^{\mathcal{H}}$: Ensemble des parties de $\mathcal{H}$.
- $\boldsymbol{\mathcal{G}} = \mathcal{G}_1 \oplus \cdots \oplus \mathcal{G}_m$: Somme hilbertienne directe.
- $\Gamma_0(\mathcal{H})$: Famille des fonctions convexes, propres et semi-continues inférieurement de $\mathcal{H}$ dans $]-\infty, +\infty]$.
- $\mathcal{B}(\mathcal{H}, \mathcal{G})$: Espace des opérateurs linéaires et bornés de $\mathcal{H}$ dans $\mathcal{G}$.
- $\mathcal{B}(\mathcal{H}) = \mathcal{B}(\mathcal{H}, \mathcal{H})$.
- L^* : Adjoint de l'opérateur $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.
- $\mathcal{S}(\mathcal{H}) = \big\{L \in \mathcal{B}(\mathcal{H}) \;\big|\; L^* = L\big\}$.
- $U \succcurlyeq V$: $(\forall x \in \mathcal{H})\ \langle Ux \mid x\rangle \geqslant \langle Vx \mid x\rangle$, où $U \in \mathcal{S}(\mathcal{H}), V \in \mathcal{S}(\mathcal{H})$.
- $\mathcal{P}_\alpha(\mathcal{H}) = \big\{U \in \mathcal{B}(\mathcal{H}) \;\big|\; (\forall x \in \mathcal{H})\ \langle Ux \mid x\rangle \geqslant \alpha\|x\|^2\big\}$, où $\alpha \in \,]0, +\infty[$.
- $(\forall x \in \mathcal{H})(\forall y \in \mathcal{H})\ \langle x \mid y\rangle_U = \langle x \mid Uy\rangle$, où $U \in \mathcal{P}_\alpha(\mathcal{H})$.
- $(\forall x \in \mathcal{H})\ \|x\|_U = \sqrt{\langle Ux \mid x\rangle}$, où $U \in \mathcal{P}_\alpha(\mathcal{H})$.
- $\to$: Convergence forte.
- $\rightharpoonup$: Convergence faible.
- $\overline{\lim}\, \alpha_n$: Limite supérieure de la suite $(\alpha_n)_{n\in\mathbb{N}}$ de $\mathbb{R}$.
- $\underline{\lim}\, \alpha_n$: Limite inférieure de la suite $(\alpha_n)_{n\in\mathbb{N}}$ de $\mathbb{R}$.
- $\ell_+^1(\mathbb{N})$: L'ensemble des suites absolument sommables dans $[0, +\infty[$.

Soit C un sous-ensemble non vide de $\mathcal{H}$.

- $\iota_C \colon x \mapsto \begin{cases} 0, & \text{si } x \in C\,; \\ +\infty, & \text{si } x \notin C \end{cases}$: Fonction indicatrice de C.
- $d_C \colon x \mapsto \inf_{y\in C} \|x - y\|$: Fonction distance à C associée à la norme $\|\cdot\| = \sqrt{\langle \cdot \mid \cdot\rangle}$.

- $\sigma_C\colon x \mapsto \sup_{y\in C} \langle x \mid y\rangle$: Fonction d'appui de C.
- P_C : Projecteur sur le sous-ensemble convexe fermé non vide C de $\mathcal{H}$.
- P_C^U : Projecteur sur le sous-ensemble convexe fermé non vide C de $\mathcal{H}$ relativement à la norme $\|\cdot\|_U$, où $U \in \mathcal{P}_\alpha(\mathcal{H})$.
- $N_C\colon x \mapsto \begin{cases} \{u \in \mathcal{H} \mid (\forall y \in C)\ \langle y-x \mid u\rangle \leqslant 0\} & \text{si } x \in C \\ \varnothing & \text{sinon} \end{cases}$: Opérateur cône normal à C.
- $\operatorname{int} C$: Intérieur de C.
- $\operatorname{cone} C = \cup_{\lambda>0} \lambda C$.
- $\operatorname{sri} C = \big\{x \in C \bigm| \operatorname{cone}(C-x) = \overline{\operatorname{span}}(C-x)\big\}$: Intérieur relatif fort de C.
- $\operatorname{ri} C = \big\{x \in C \bigm| \operatorname{cone}(C-x) = \operatorname{span}(C-x)\big\}$: Intérieur relatif de C.

Notations et définitions relatives à un opérateur multivoque $A\colon \mathcal{H} \to 2^{\mathcal{H}}$

- $\operatorname{dom} A = \big\{x \in \mathcal{H} \bigm| Ax \neq \varnothing\big\}$: Domaine de A.
- $\operatorname{gra} A = \big\{(x,u) \in \mathcal{H}^2 \bigm| u \in Ax\big\}$: Graphe de A.
- $A^{-1}\colon \mathcal{H} \to 2^{\mathcal{H}}\colon u \mapsto \big\{x \in \mathcal{H} \bigm| u \in Ax\big\}$: Inverse de A.
- $\operatorname{Fix} A = \big\{x \in \mathcal{H} \bigm| x \in Ax\big\}$: Points fixes de A.
- $\operatorname{zer} A = \big\{x \in \mathcal{H} \bigm| 0 \in Ax\big\}$: Zéros de A.
- $\operatorname{ran} A = \big\{u \in \mathcal{H} \bigm| (\exists x \in \mathcal{H})\ u \in Ax\big\}$: Image de A.
- $J_A = (\operatorname{Id} + A)^{-1}$: Résolvante de A.
- $R_A = 2J_A - \operatorname{Id}$: Opérateur de réflexion de A.
- A est monotone :

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

- A est maximalement monotone :

$$(\forall (x,u) \in \mathcal{H} \oplus \mathcal{H}) \quad \Big((x,u) \in \operatorname{gra} A \Leftrightarrow (\forall (y,v) \in \operatorname{gra} A) \quad \langle x-y \mid u-v\rangle \geqslant 0\Big).$$

- A est γ-fortement monotone :

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

- A est demirégulier en $x \in \operatorname{dom} A$:

$$(\forall ((x_n,u_n))_{n\in\mathbb{N}} \in (\operatorname{gra} A)^{\mathbb{N}})(\forall u \in Ax) \quad \begin{cases} x_n \rightharpoonup x \\ u_n \to u \end{cases} \Rightarrow \quad x_n \to x.$$

Définitions relatives à un opérateur univoque $T\colon \mathcal{H} \to \mathcal{H}$

- L'ensemble des points fixes de T :

$$\operatorname{Fix} T = \big\{x \in \mathcal{H} \;\big|\; Tx = x\big\}.$$

- T est lipschitzien de constante $\chi \in\,]0,+\infty[$ (ou T est χ–lipschitzien) :

$$(\forall (x,y) \in \mathcal{H}^2) \quad \|Tx - Ty\| \leqslant \chi\|x-y\|.$$

- T est β–cocoercif, où $\beta \in\,]0,+\infty[$: βT est une contraction ferme,

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \langle x - y \mid Tx - Ty\rangle \geqslant \beta\|Tx - Ty\|^2.$$

- $\mathfrak{T}(W) = \big\{T\colon \mathcal{H} \to \mathcal{H} \;\big|\; (\forall x \in \mathcal{H})(\forall y \in \operatorname{Fix} T) \quad \langle y - Tx \mid x - Tx\rangle_W \leqslant 0\big\}.$

Notations relatives à une fonction $f \in \Gamma_0(\mathcal{H})$

- Domaine de f :

$$\operatorname{dom} f = \big\{x \in \mathcal{H} \;\big|\; f(x) < +\infty\big\}.$$

- Ensemble des minimiseurs de f :

$$\operatorname{Argmin} f.$$

- Le minimiseur de f en cas d'unicité :

$$\operatorname{argmin} f(\mathcal{H}) \quad \text{ou} \quad \underset{y\in\mathcal{H}}{\operatorname{argmin}}\; f(y).$$

- Conjuguée de f :

$$f^*\colon u \mapsto \sup_{x\in\mathcal{H}} \big(\langle x \mid u\rangle - f(x)\big).$$

- Enveloppe de Moreau d'indice $\gamma \in\,]0,+\infty[$ de f :

$$^{\gamma}f\colon x \mapsto \inf_{y\in\mathcal{H}} \left(f(y) + \frac{1}{2\gamma}\|x-y\|^2\right).$$

  Si $\gamma = 1$, on note $\widetilde{f} = {}^1 f$.

- Le sous-différentiel de f en $x \in \operatorname{dom} f$:

$$\partial f(x) = \big\{u \in \mathcal{H} \;\big|\; (\forall y \in \mathcal{H})\, \langle y - x \mid u\rangle + f(x) \leqslant f(y)\big\}.$$

- L'opérateur de proximité de f :

$$\mathrm{prox}_f \colon \mathcal{H} \to \mathcal{H} \colon x \mapsto \underset{y \in \mathcal{H}}{\mathrm{argmin}} \left(f(y) + \frac{1}{2}\|x - y\|^2 \right).$$

- L'opérateur de proximité de f relativement à la norme $\|\cdot\|_U$:

$$\mathrm{prox}_f^U \colon \mathcal{H} \to \mathcal{H} \colon x \mapsto \underset{y \in \mathcal{H}}{\mathrm{argmin}} \left(f(y) + \frac{1}{2}\|x - y\|_U^2 \right).$$

- La section inférieure de f à hauteur $\eta \in \mathbb{R}$:

$$\mathrm{lev}_{\leqslant \eta} f = \left\{ x \in \mathcal{H} \;\middle|\; f(x) \leqslant \eta \right\}.$$

Chapitre 1

Introduction

1.1 Inclusions monotones

Rappelons un problème classique de la théorie des opérateurs monotones et de ses applications.

Problème 1.1 Soit $\mathcal{H}$ un espace hilbertien réel, soit $C\colon \mathcal{H} \to 2^{\mathcal{H}}$ un opérateur maximalement monotone. Le problème est de

$$\text{trouver } \overline{x} \text{ dans } \mathcal{H} \text{ tel que } 0 \in C\overline{x}. \tag{1.1}$$

Ce problème a été étudié extensivement dans la littérature (voir [4, 31, 38] et leur bibliographies). La méthode proximale a été proposée dans [6, 31] pour résoudre le Problème 1.1. On rappelle le résultat suivant.

Théorème 1.2 [2, Theorem 23.41] *Dans le Problème 1.1, supposons que* $\operatorname{zer} C \neq \varnothing$. *Soient* $x_0 \in \mathcal{H}$ *et* $(\gamma_n)_{n\in\mathbb{N}}$ *une suite dans* $]0, +\infty[$ *telle que* $\sum_{n\in\mathbb{N}} \gamma_n^2 = +\infty$. *Alors, la suite* $(x_n)_{n\in\mathbb{N}}$ *engendrée par l'algorithme*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = J_{\gamma_n C} x_n \tag{1.2}$$

converge faiblement vers une solution $\overline{x}$ *du problème* (1.1).

En général, les résolvantes $(J_{\gamma_n C})_{n\in\mathbb{N}}$ sont difficile à mettre en œuvre numériquement. On s'oriente alors vers des stratégies d'éclatement sous forme de sommes d'opérateurs. Ainsi le Problème 1.1 a ensuite été étendu dans [27] au problème de trouver un zéro de la somme $C = A + B$ de deux opérateurs maximalement monotone, où l'un d'entre eux est cocoercif, i.e., son inverse est fortement monotone (voir aussi [1, 4, 22, 23, 24, 34, 35, 39] pour des travaux concernant les opérateurs cocoercifs).

Problème 1.3 Soient $\beta \in]0, +\infty[$, $\mathcal{H}$ un espace hilbertien réel, $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ un opérateur maximalement monotone, et $B\colon \mathcal{H} \to \mathcal{H}$ un opérateur β-cocoercif. Le problème est de

$$\text{trouver } \overline{x} \text{ dans } \mathcal{H} \text{ tel que } 0 \in A\overline{x} + B\overline{x}. \tag{1.3}$$

La méthode explicite-implicite a été proposée dans [27] pour résoudre ce problème. Cette méthode trouve ses origines dans la méthode du gradient projeté en optimisation convexe (voir aussi [1, 4, 13, 16, 17, 19] et leur bibliographies). On présente le résultat plus général sur cette méthode dans le théorème suivant [1, 13].

Théorème 1.4 (**Méthode explicite-implicite** [1, Theorem 2.8], [13, Section 6.2]) *Considérons le Problème 1.3. Soient $\varepsilon \in]0, \beta/2[$, $(\lambda_n)_{n\in\mathbb{N}}$ une suite dans $[\varepsilon, 1]$, $(\gamma_n)_{n\in\mathbb{N}}$ une suite dans $[\varepsilon, 2\beta - \varepsilon]$, $x_0 \in \mathcal{H}$, $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ des suites absolument sommables dans $\mathcal{H}$. On engendre une suite $(x_n)_{n\in\mathbb{N}}$ comme suit.*

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} y_n = x_n - \gamma_n(Bx_n + b_n) \\ x_{n+1} = x_n + \lambda_n\big(J_{\gamma_n A} y_n + a_n - x_n\big). \end{array} \right. \tag{1.4}$$

Supposons que $\operatorname{zer}(A + B) \neq \varnothing$. *Alors, on a les résultats suivants pour une solution $\overline{x}$ du problème* (1.3).

(i) *$(x_n)_{n\in\mathbb{N}}$ converge faiblement vers $\overline{x}$.*

(ii) *Supposons que l'une de conditions suivante soit satisfaite :*

  (a) *A ou B est demirégulier en $\overline{x}$.*

  (b) $\operatorname{int}\operatorname{zer}(A + B) \neq \varnothing$.

  Alors $(x_n)_{n\in\mathbb{N}}$ converge fortement vers $\overline{x}$.

Dans le cas où l'opérateur B dans le Problème 1.3 est seulement lipschitzien et monotone, on arrive au problème suivant qui est plus général que Problème 1.3.

Problème 1.5 Soient $\beta \in]0, +\infty[$, $\mathcal{H}$ un espace hilbertien réel, $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ un opérateur maximalement monotone, $B\colon \mathcal{H} \to \mathcal{H}$ un opérateur monotone et β-lipschitzien. Le problème est de

$$\text{trouver } \overline{x} \text{ dans } \mathcal{H} \text{ tel que } 0 \in A\overline{x} + B\overline{x}. \tag{1.5}$$

On peut utiliser la méthode explicite-implicite-explicite proposée initialement dans [36] pour résoudre ce problème. On rappelle le résultat suivant de [7] qui incorpore des termes d'erreur.

Théorème 1.6 (**Méthode explicite-implicite-explicite** [7, Theorem 2.5]) *Considérons le Problème 1.5. Soient $x_0 \in \mathcal{H}$, $\varepsilon \in]0, \beta^{-1}/2[$, $(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}}$, et $(c_n)_{n\in\mathbb{N}}$ des suites* *absolument sommables dans $\mathcal{H}$. Posons*

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} \gamma_n \in [\varepsilon, \beta^{-1} - \varepsilon] \\ y_n = x_n - \gamma_n(Bx_n + b_n) \\ p_n = J_{\gamma_n A} y_n + a_n \\ q_n = p_n - \gamma_n(Bp_n + c_n) \\ x_{n+1} = x_n - y_n + q_n. \end{array} \right. \tag{1.6}$$

Supposons que $\operatorname{zer}(A+B) \neq \varnothing$. *Alors, on a les résultats suivants pour une solution $\overline{x}$ du problème* (1.5).

(i) $(x_n)_{n\in\mathbb{N}}$ *converge faiblement vers $\overline{x}$.*

(ii) *Supposons que l'une des conditions suivantes soit satisfaite.*

  (a) $A + B$ *est demirégulier en $\overline{x}$.*

  (b) A *ou* B *est uniformément monotone en $\overline{x}$.*

  (c) $\operatorname{int} \operatorname{zer}(A+B) \neq \varnothing$.

  Alors $(x_n)_{n\in\mathbb{N}}$ *converge fortement vers $\overline{x}$.*

On a vu que l'opérateur B dans les Problèmes 1.3 et 1.5 est univoque. Dans le cas où il est multivoque, on a le problème suivant.

Problème 1.7 Soient $\mathcal{H}$ un espace hilbertien réel, $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ et $B\colon \mathcal{H} \to 2^{\mathcal{H}}$ des opérateurs maximalement monotones. Le problème est de

$$\text{trouver } \overline{x} \text{ dans } \mathcal{H} \text{ tel que } 0 \in A\overline{x} + B\overline{x}. \tag{1.7}$$

On peut utiliser la méthode de Douglas-Rachford proposée initialement dans [26] pour résoudre le problème (1.7) (voir aussi [4, 13, 15, 21, 33] et leur bibliographies).

Théorème 1.8 ([4, Theorem 25.6]) *Dans le Problème 1.7, supposons que* $\operatorname{zer}(A+B) \neq \varnothing$. *Soient $\gamma \in \left]0, +\infty\right[$, $x_0 \in \mathcal{H}$, $(\lambda_n)_{n\in\mathbb{N}}$ une suite dans $[0, 2]$ telle que $\sum_{n\in\mathbb{N}} \lambda_n(2 - \lambda_n) = +\infty$. On engendre des suites $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$ et $(z_n)_{n\in\mathbb{N}}$ comme suit.*

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} y_n = J_{\gamma B} x_n \\ z_n = J_{\gamma A}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{array} \right. \tag{1.8}$$

Alors on a les résultats suivants pour un point $\overline{x} \in \operatorname{Fix} R_{\gamma A} R_{\gamma B}$.

(i) $J_{\gamma B}\overline{x} \in \operatorname{zer}(A+B)$ *et* $(x_n)_{n\in\mathbb{N}}$ *converge faiblement vers $\overline{x}$.*

(ii) $(y_n)_{n\in\mathbb{N}}$ *converge faiblement vers $J_{\gamma B}\overline{x}$.*

(iii) $(z_n)_{n\in\mathbb{N}}$ *converge faiblement vers $J_{\gamma B}\overline{x}$.*

Dans le cas où l'opérateur C dans le problème (1.1) est une somme quelconque d'opérateurs maximalement monotones, on arrive au problème suivant [14, 32].

Problème 1.9 Soient m un entier strictement positif et $\mathcal{H}$ un espace hilbertien réel. Pour tout $i \in \{1, \dots, m\}$, soit $C_i \colon \mathcal{H} \to 2^{\mathcal{H}}$ un opérateur maximalement monotone. Le problème est de

$$\text{trouver } \overline{x} \text{ dans } \mathcal{H} \text{ tel que } 0 \in \sum_{i=1}^{m} C_i \overline{x}. \tag{1.9}$$

On peut le résoudre par la méthode parallèle basée sur Douglas-Rachford proposée dans [14] où celle des inverses partiels proposée dans [32]. Dans le cas où l'un des opérateurs $(C_i)_{1 \leqslant i \leqslant m}$ est fortement monotone, on dispose également de la méthode parallèle de type Dykstra de [14].

1.2 Objectifs

L'objectif principal de cette thèse est de développer des méthodes d'éclatement d'opérateurs pour résoudre des problèmes plus généraux par leur structure que le Problème 1.9. Le problème générique que nous considérons est le suivant.

Problème 1.10 Soient $\boldsymbol{\mathcal{H}}$ un espace hilbertien réel, $\boldsymbol{z} \in \boldsymbol{\mathcal{H}}$, $\boldsymbol{A} \colon \boldsymbol{\mathcal{H}} \to 2^{\boldsymbol{\mathcal{H}}}$ un opérateur maximalement monotone, et $\boldsymbol{C} \colon \boldsymbol{\mathcal{H}} \to \boldsymbol{\mathcal{H}}$ un opérateur maximalement monotone. Soient $\boldsymbol{\mathcal{G}}$ un espace hilbertien réel, $\boldsymbol{r} \in \boldsymbol{\mathcal{G}}$, $\boldsymbol{B} \colon \boldsymbol{\mathcal{G}} \to 2^{\boldsymbol{\mathcal{G}}}$ et $\boldsymbol{D} \colon \boldsymbol{\mathcal{G}} \to 2^{\boldsymbol{\mathcal{G}}}$ deux opérateurs maximalement monotones, et $0 \neq \boldsymbol{L} \in \mathcal{B}(\boldsymbol{\mathcal{H}}, \boldsymbol{\mathcal{G}})$. Le problème est de résoudre l'inclusion primale

$$\text{trouver } \overline{\boldsymbol{x}} \in \boldsymbol{\mathcal{H}} \text{ tel que } \boldsymbol{z} \in (\boldsymbol{A} + \boldsymbol{C})\overline{\boldsymbol{x}} + \boldsymbol{L}^*\big((\boldsymbol{B} \,\square\, \boldsymbol{D})(\boldsymbol{L}\overline{\boldsymbol{x}} - \boldsymbol{r})\big) \tag{1.10}$$

et l'inclusion duale

$$\begin{aligned} &\text{trouver } \overline{\boldsymbol{v}} \in \boldsymbol{\mathcal{G}} \text{ tel que} \\ &\qquad -\boldsymbol{r} \in -\boldsymbol{L}\big((\boldsymbol{A} + \boldsymbol{C})^{-1}(\boldsymbol{z} - \boldsymbol{L}^*\overline{\boldsymbol{v}})\big) + \boldsymbol{B}^{-1}\overline{\boldsymbol{v}} + \boldsymbol{D}^{-1}\overline{\boldsymbol{v}}. \end{aligned} \tag{1.11}$$

Cette dualité générale a été introduite dans [18] (on trouve des cas particuliers dans [2, 7, 10, 11, 28, 29]).

La principale motivation de cette thèse est d'unifier un grand nombre de méthodes existantes pour résoudre certains cas particuliers de Problème 1.10 et de développer de nouveaux algorithmes à métriques constante et variable. Ces algorithmes seront appliqués à plusieurs problèmes concrets.

1.3 Organisation

Au Chapitre 2, nous élaborons une méthode primale-duale pour résoudre des problèmes d'optimisation fortement convexe composites. L'algorithme proposé est une application de la méthode explicite-implicite (1.4) au problème dual, et il est appliqué ensuite à divers problèmes en mathématiques appliquées. On montre que plusieurs algorithmes connus [3, 5, 8, 30, 37] sont des cas particuliers de cet algorithme. Des comparaisons numériques avec l'algorithme proposé récemment dans [9] et avec celui de [11] sont présentées dans le contexte du débruitage d'image.

Au Chapitre 3, nous proposons un algorithme pour calculer l'opérateur proximal d'une fonction composite de la forme

$$h\colon \mathcal{H} \to \left]-\infty,+\infty\right] : x \mapsto \sum_{i=1}^{m} g_i(L_i x - r_i), \tag{1.12}$$

où $(\forall i \in \{1,\ldots,m\})$ $r_i \in \mathcal{G}_i, g_i \in \Gamma_0(\mathcal{G}_i)$, et $0 \neq L_i \in \mathcal{B}(\mathcal{H},\mathcal{G}_i)$. Ensuite, nous présentons des applications aux problèmes de meilleure approximation relativement à l'intersection de sous-ensembles convexes composites, et de traitement du signal.

Au Chapitre 4, nous nous intéressons aux inclusions monotones composites impliquant des opérateurs cocoercifs. Nous proposons un algorithme qui admet une structure de la méthode explicite-implicite pour résoudre ce problème. Des liens avec les méthodes de [9, 13, 20, 25] sont présentés.

Au Chapitre 5, nous introduisons la notion de suite quasi-fejérienne à métrique variable et analysons son comportement asymptotique. Dans le cas d'une métrique constante, ces résultats se réduisent aux résultats connus de [12]. Les résultats obtenus sont utilisés pour montrer la convergence faible et forte d'algorithmes pour résoudre des problèmes de point fixe et d'admissibilité convexe dans le Chapitre 5, et des inclusions monotones en dualité dans les Chapitres 6 et 7.

Au Chapitre 6 nous proposons tout d'abord une méthode explicite-implicite à métrique variable pour résoudre le Problème 1.3. Les résultats se réduisent aux résultats connus de [7, 13, 19, 27] dans le cas d'une métrique constante. Ensuite, nous appliquons cet algorithme à la résolution d'inclusions monotones en dualité. De plus, des nouvelles applications sont présentées.

Au Chapitre 7, nous développons une méthode explicite-implicite-explicite à métrique variable pour résoudre le Problème 1.5. De plus, nous proposons un algorithme primal-dual à métrique variable pour résoudre des inclusions monotones composites impliquant des opérateurs lipschitziens et monotones. Des liens avec les méthodes de [7, 18, 36] sont établis.

Au Chapitre 8, nous présentons quelques conclusions et des problèmes ouverts.

1.4 Contributions principales

- Unification de nombreuses méthodes d'éclatement d'opérateurs. Plusieurs méthodes en apparence sans lien sont regroupées et étendues dans un cadre commun.
- Étude primale-duale de la méthode explicite-implicite pour résoudre les problèmes d'optimisation composites fortement convexes et d'inclusions composites fortement monotones. Cette approche nous permet de développer de nouveaux algorithmes et de résoudre des problèmes pour lesquels aucune méthode déclatement existait jusqu'alors.
- Conception et étude asymptotique de la méthode explicite-implicite à métrique variable pour résoudre le Problème 1.3 et de la méthode explicite-implicite à métrique variable pour résoudre le Problème 1.5.
- Introduction de la notion de suite quasi-fejérienne à métrique variable. Les résultats obtenus sont des outils fondamentaux pour démontrer la convergence faible et forte de schémas numériques à métrique variable en analyse non-linéaire.
- Conception et étude asymptotique d'une méthode à métrique variable pour résoudre le problème de point fixe commun. En particulier, nous obtenons une nouvelle méthode proximale à métrique variable.
- Étude systématique de la méthode explicite-implicite (à métrique variable) et de la méthode explicite-implicite-explicite (à métrique variable) pour résoudre des inclusions monotones associées à des opérateurs cocoercifs et lipschitziens, respectivement.
- Développement de nouvelles méthodes de résolution de problèmes inverses, de théorie du signal, de traitement de l'image, et de meilleure approximation.

1.5 Publications issues de la thèse

- P. L. Combettes, Đinh Dũng, and B. C. Vũ, Dualization of signal recovery problems, *Set-Valued Var. Anal.*, vol. 18, pp. 373–404, 2010.
- P. L. Combettes, Đinh Dũng, and B. C. Vũ, Proximity for sums of composite functions, *J. Math. Anal. Appl.*, vol. 380, pp. 680–688, 2011.
- B. C. Vũ, A splitting algorithm for dual monotone inclusions involving cocoercive operators, *Adv. Comput. Math.*, vol. 38, pp. 667–681, 2013.
- P. L. Combettes and B. C. Vũ, Variable metric quasi-Fejér monotonicity, *Nonlinear Anal.,* vol. 78, pp. 17–31, 2013.
- P. L. Combettes and B. C. Vũ, Variable metric forward-backward splitting with applications to monotone inclusions in duality, *Optimization*, à paraître, 2013. http://www.tandfonline.com/doi/full/10.1080/02331934.2012.733883
- B. C. Vũ, A variable metric extension of the forward–backward–forward algorithm for monotone operators, *Numer. Funct. Anal. Optim.*, à paraître, 2013.

1.6 Bibliographie

[1] H. Attouch, L. M. Briceño-Arias, and P. L. Combettes, A parallel splitting method for coupled monotone inclusions, *SIAM J. Control Optim.*, vol. 48, pp. 3246–3270, 2010.

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

[3] H. H. Bauschke and P. L. Combettes, A Dykstra-like algorithm for two monotone operators, *Pacific J. Optim.,* vol. 4, pp. 383–391, 2008.

[4] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York, 2011.

[5] A. Bermùdez and C. Moreno, Duality methods for solving variational inequalities, *Comp. Math. Appl.*, vol. 7, pp. 43–58, 1981.

[6] H. Brézis and P. L. Lions, Produits infinis de résolvantes, *Israel J. Math.*, vol. 29, pp. 329–345, 1978.

[7] 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.

[8] A. Chambolle, Total variation minimization and a class of binary MRF model, *Lecture Notes in Comput. Sci.,* vol. 3757, pp. 136–152, 2005.

[9] A. Chambolle and T. Pock, A first order primal dual algorithm for convex problems with applications to imaging, *J. Math. Imaging Vision*, vol. 40, pp. 120–145, 2011.

[10] G. H-G. Chen and R. T. Rockafellar, Convergence rates in forward-backward splitting, *SIAM J. Optim.,* vol. 7, pp. 421–444, 1997.

[11] G. Chen and M. Teboulle, A proximal based decomposition method for minimization problems, *Math. Program.*, vol. 64, pp. 81–101, 1994.

[12] P. L. Combettes, Quasi-Fejérian analysis of some optimization algorithms, in : *Inherently Parallel Algorithms for Feasibility and Optimization*, (D. Butnariu, Y. Censor, and S. Reich, eds.), pp. 115–152. Elsevier, New York, 2001.

[13] P. L. Combettes, Solving monotone inclusions via compositions of nonexpansive averaged operators, *Optimization*, vol. 53, pp. 475–504, 2004.

[14] P. L. Combettes, Iterative construction of the resolvent of a sum of maximal monotone operators, *J. Convex Anal.*, vol. 16, pp. 727–748, 2009.

[15] P. L. Combettes and J.-C. Pesquet, A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery, *IEEE J. Selected Topics Signal Process.,* vol. 1, pp. 564–574, 2007.

[16] P. L. Combettes and J.-C. Pesquet, Proximal thresholding algorithm for minimization over orthonormal bases, *SIAM J. Optim*, vol. 18, pp. 1351–1376, 2007.

[17] P. L. Combettes and J.-C. Pesquet, Proximal splitting methods in signal processing, in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering,* (H. H. Bauschke et al., eds), pp. 185–212. Springer, New York, 2011.

[18] P. L. Combettes and J.-C. Pesquet, Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators, *Set-Valued Var. Anal.*, vol. 20, pp. 307–330, 2012.

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

[20] L. Condat, A generic first-order primal-dual method for convex optimization involving Lipschitzian, proximable and linear composite terms, *J. Optim. Theory Appl.*, to appear, 2012.

[21] J. Eckstein and D. Bertsekas, On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators, *Math. Programming*, vol. 55, pp. 293–318, 1992.

[22] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer-Verlag, New York, 2003.

[23] D. Gabay, Applications of the method of multipliers to variational inequalities, in : M. Fortin and R. Glowinski (eds.), *Augmented Lagrangian Methods : Applications to the Numerical Solution of Boundary Value Problems,* pp. 299–331. North-Holland, Amsterdam, 1983.

[24] R. Glowinski and P. Le Tallec, *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*, SIAM, Philadelphia, 1989.

[25] B. He and X. Yuan, Convergence analysis of primal-dual algorithms for a saddle-point problem : From contraction perspective, *SIAM J. Imaging Sci.,* vol. 5, pp. 119–149, 2012.

[26] P.-L. Lions and B. Mercier, Splitting algorithms for the sum of two nonlinear operators, *SIAM J. Numer. Anal.*, vol. 16, pp. 964–979, 1979.

[27] B. Mercier, *Topics in Finite Element Solution of Elliptic Problems* (Lectures on Mathematics, no. 63). Tata Institute of Fundamental Research, Bombay, 1979.

[28] U. Mosco, Dual variational inequalities, *J. Math. Anal. Appl.,* vol. 40, pp. 202–206, 1972.

[29] T. Pennanen, Dualization of generalized equations of maximal monotone type, *SIAM J. Optim.,* vol. 10, pp. 809–835, 2000.

[30] L. C. Potter and K. S. Arun, A dual approach to linear inverse problems with convex constraints, *SIAM J. Control Optim.,* vol. 31, pp. 1080–1092, 1993.

[31] R. T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM J. Control Optimization*, vol. 14, pp. 877–898, 1976.

[32] J. E. Spingarn, Partial inverse of a monotone operator, *Appl. Math. Optim.*, vol. 10, pp. 247–265, 1983.

[33] B. F. Svaiter, On weak convergence of the Douglas-Rachford method, *SIAM J. Control Optim.*, vol. 49, pp. 280–287, 2011.

[34] P. Tseng, Further applications of a splitting algorithm to decomposition in variational inequalities and convex programming, *Math. Programming*, vol. 48, pp. 249–263, 1990.

[35] P. Tseng, Applications of a splitting algorithm to decomposition in convex programming and variational inequalities, *SIAM J. Control Optim.*, vol. 29, pp. 119–138, 1991.

[36] P. Tseng, A modified forward-backward splitting method for maximal monotone mappings, *SIAM J. Control Optim.*, vol. 38, pp. 431–446, 2000.

[37] D. C. Youla, Generalized image restoration by the method of alternating orthogonal projections, *IEEE Trans. Circuits Syst.,* vol. 25, pp. 694–702, 1978.

[38] E. Zeidler, *Nonlinear Functional Analysis and Its Applications II/B– Nonlinear Monotone Operators*, Springer-Verlag, New York, 1990.

[39] D. L. Zhu and P. Marcotte, Co-coercivity and its role in the convergence of iterative schemes for solving variational inequalities, *SIAM J. Optim.*, vol. 6, pp. 714–726, 1996.

Chapitre 2

Dualisation de problèmes inverses en théorie du signal

Nous proposons un algorithme pour minimiser la somme d'une fonction fortement convexe et d'une fonction composite. L'algorithme résulte de l'application de la méthode explicite-implicite au problème dual. Nous obtenons la convergence forte de la suite primale et faible de la suite duale dans des espaces hilbertiens réels.

2.1 Description et résultats principaux

Nous nous intéressons au problème suivant qui permet la modélisation d'une grande classe de problèmes [23, 24, 50, 56, 66, 68, 72, 79, 80].

Problème 2.1 Soient $\mathcal{H}$ et $\mathcal{G}$ deux espaces hilbertiens réels, $z \in \mathcal{H}$, $r \in \mathcal{G}$, $f \in \Gamma_0(\mathcal{H})$, $g \in \Gamma_0(\mathcal{G})$, et $0 \neq L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ tels que

$$r \in \operatorname{sri}\big(L(\operatorname{dom} f) - \operatorname{dom} g\big). \tag{2.1}$$

Le problème primal est de

$$\underset{x \in \mathcal{H}}{\text{minimiser}}\; f(x) + g(Lx - r) + \frac{1}{2}\|x - z\|^2, \tag{2.2}$$

et le problème dual est de

$$\underset{v \in \mathcal{G}}{\text{minimiser}}\; \widetilde{f}^*(z - L^*v) + g^*(v) + \langle v \mid r \rangle. \tag{2.3}$$

Le premier résultat établit des liens entre le problème primal et le problème dual.

Proposition 2.2 *Soit $\gamma \in]0,+\infty[$. Sous les hypothèses du Problème 2.1, le problème (2.2) et le problème (2.3) sont en dualité forte, c'est à dire,*

$$\inf_{x\in\mathcal{H}} f(x) + g(Lx-r) + \frac{1}{2}\|x-z\|^2 = -\min_{v\in\mathcal{G}} \widetilde{f}^*(z-L^*v) + g^*(v) + \langle v \mid r\rangle, \tag{2.4}$$

le problème (2.3) possède au moins une solution $\overline{v}$, le problème (2.2) possède une solution unique $\overline{x}$, et ces solutions sont liées par les relations

$$\overline{x} = \operatorname{prox}_f(z - L^*\overline{v}) \quad \text{et} \quad \overline{v} = \operatorname{prox}_{\gamma g^*}(\overline{v} + \gamma L\overline{x}). \tag{2.5}$$

Observons que la fonction $\widetilde{f}^*$ est différentiable sur $\mathcal{G}$ avec un gradient lipschitzien [7]. Donc, pour résoudre le Problème 2.1, nous appliquons la méthode explicite-implicite (1.4) au problème dual, et allons par ce biais récupérer la solution primale.

Algorithme 2.3 Soit $(a_n)_{n\in\mathbb{N}}$ une suite absolument sommable dans $\mathcal{G}$, et soit $(b_n)_{n\in\mathbb{N}}$ une suite absolument sommable dans $\mathcal{H}$. Des suites $(x_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ sont engendrées comme suit.

$$\begin{array}{l} \text{Initialisation} \\ \left\lfloor \begin{array}{l} \varepsilon \in]0, \min\{1, \|L\|^{-2}\}[\\ v_0 \in \mathcal{G} \end{array} \right. \\ \text{Pour } n = 0,1,\ldots \\ \left\lfloor \begin{array}{l} x_n = \operatorname{prox}_f(z - L^*v_n) + b_n \\ \gamma_n \in [\varepsilon, 2\|L\|^{-2} - \varepsilon] \\ \lambda_n \in [\varepsilon, 1] \\ v_{n+1} = v_n + \lambda_n\big(\operatorname{prox}_{\gamma_n g^*}(v_n + \gamma_n(Lx_n - r)) + a_n - v_n\big). \end{array} \right. \end{array} \tag{2.6}$$

Dans cette méthode, on obtient un éclatement de tous les opérateurs puisque L, prox_f et $\operatorname{prox}_{g^*}$ sont utilisés individuellement à chaque itération. De plus, la méthode tolère des erreurs dans l'évaluation de chaque opérateur impliqué. Nous obtenons le résultat de convergence suivant.

Théorème 2.4 *Soient $(x_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ des suites engendrées par l'Algorithme 2.3 et soit $\overline{x}$ la solution du problème (2.2). Alors nous avons les résultats suivants.*

(i) *$(x_n)_{n\in\mathbb{N}}$ converge fortement vers $\overline{x}$.*

(ii) *$(v_n)_{n\in\mathbb{N}}$ converge faiblement vers une solution $\overline{v}$ du problème (2.3) et $\overline{x} = \operatorname{prox}_f(z - L^*\overline{v})$.*

Nous illustrons à présent des applications du Problème 2.1 aux problèmes de meilleure approximation, de débruitage de signaux à l'aide de dictionnaires et de restauration de signaux avec des fonctions d'appui. Nous listons ci-dessous quelques cas particuliers du Problème 2.1.

Exemple 2.5 Soient $z \in \mathcal{H}$, $r \in \mathcal{G}$, $C \subset \mathcal{H}$ et $D \subset \mathcal{G}$ deux sous-ensembles convexes fermés, et $0 \neq L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ tels que

$$r \in \operatorname{sri}\big(L(C) - D\big). \tag{2.7}$$

Le problème primal est de

$$\underset{\substack{x \in C \\ Lx - r \in D}}{\text{minimiser}} \ \frac{1}{2}\|x - z\|^2, \tag{2.8}$$

et le problème dual est de

$$\underset{v \in \mathcal{G}}{\text{minimiser}} \ \frac{1}{2}\|z - L^*v\|^2 - \frac{1}{2}d_C^2(z - L^*v) + \sigma_D(v) + \langle v \mid r\rangle. \tag{2.9}$$

Cet exemple est un cas particulier du Problème 2.1 où $f = \iota_C$ avec $\operatorname{dom} f = C$, et $g = \iota_D$ avec $\operatorname{dom} g = D$. Donc, nous pouvons utiliser l'Algorithme 2.3 pour résoudre le problème (2.8) et le problème (2.9).

La condition (2.7) implique que l'intersection $C \cap L^{-1}(r + D)$ dans l'Exemple 2.5 est non vide. Pourtant, dans certaines situations (voir [31, 81]), l'intersection peut être vide. Dans ce cas, nous proposons le problème suivant.

Exemple 2.6 Soient $z \in \mathcal{H}$, $r \in \mathcal{G}$, $C \subset \mathcal{H}$ et $D \subset \mathcal{G}$ deux sous-ensembles convexes fermés non vides, $0 \neq L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$, ϕ et ψ deux fonctions paires dans $\Gamma_0(\mathbb{R}) \smallsetminus \{\iota_{\{0\}}\}$ tels que

$$r \in \operatorname{sri}\Big(L\big(\{x \in \mathcal{H} \mid d_C(x) \in \operatorname{dom}\phi\}\big) - \{y \in \mathcal{G} \mid d_D(y) \in \operatorname{dom}\psi\}\Big). \tag{2.10}$$

Le problème primal est de

$$\underset{x \in \mathcal{H}}{\text{minimiser}} \ \phi\big(d_C(x)\big) + \psi\big(d_D(Lx - r)\big) + \frac{1}{2}\|x - z\|^2, \tag{2.11}$$

et le problème dual est de

$$\underset{v \in \mathcal{G}}{\text{minimiser}} \ \frac{1}{2}\|z - L^*v\|^2 - (\phi \circ d_C)^{\sim}(z - L^*v) + \sigma_D(v) + \psi^*(\|v\|) + \langle v \mid r\rangle. \tag{2.12}$$

Cet exemple est encore un cas particulier du Problème 2.1 où $f = \phi \circ d_C$ avec $\operatorname{dom} f = \big\{x \in \mathcal{H} \mid d_C(x) \in \operatorname{dom}\phi\big\}$, et $g = \psi \circ d_D$ avec $\operatorname{dom} g = \big\{x \in \mathcal{H} \bigm| d_D(x) \in \operatorname{dom}\psi\big\}$.

D'autres cas particuliers du Problème 2.1 sont des problèmes de restauration de signal. Nous nous intéressons à la restauration d'un signal original $\widetilde{x}$ à partir d'un signal bruité z dans $\mathcal{H}$ selon le modèle

$$z = \widetilde{x} + w, \tag{2.13}$$

où w est un bruit additif. Des méthodes variationnelles ont été proposées dans [2, 23, 24, 27, 32, 36, 44, 45, 54, 72, 77, 78] pour résoudre le problème (2.13). Une approche commune de résoudre ce problème est de minimiser la fonction $x \mapsto \|x-z\|^2/2$ sous des contraintes sur x qui représentent les information a priori sur $\widetilde{x}$, et quelques transformations affines $L\widetilde{x}-z$ de celles-ci. Dans ce contexte, L peut être un gradient [23, 24, 45, 54, 72], un filtre de basse-fréquence [2, 77], un opérateur de décomposition sur une base d'ondellette [36, 44, 78]. Nous proposons la formulation variationnelle suivante où les informations sur $\widetilde{x}$ portent sur les produits scalaires $(\langle \widetilde{x} \mid e_k \rangle)_{k\in\mathbb{K}}$, où $(e_k)_{k\in\mathbb{K}}$ est une suite finie ou infinie de vecteurs de références dans $\mathcal{H}$, et la fonction f modélise les autres propriétés connues de x.

Exemple 2.7 Soient $z \in \mathcal{H}$, $f \in \Gamma_0(\mathcal{H})$, et $(e_k)_{k\in\mathbb{K}}$ une suite de vecteurs normés dans $\mathcal{H}$ tels que

$$(\exists\, \delta \in \left]0,+\infty\right[)(\forall x \in \mathcal{H}) \quad \sum_{k\in\mathbb{K}} |\langle x \mid e_k\rangle|^2 \leqslant \delta \|x\|^2, \tag{2.14}$$

et $(\phi_k)_{k\in\mathbb{K}}$ des fonctions dans $\Gamma_0(\mathbb{R})$ telles que

$$(\forall k \in \mathbb{K}) \quad \phi_k \geqslant \phi_k(0) = 0 \tag{2.15}$$

et

$$0 \in \operatorname{sri}\left\{ \big(\langle x \mid e_k\rangle - \xi_k\big)_{k\in\mathbb{K}} \;\middle|\; (\xi_k)_{k\in\mathbb{K}} \in \ell^2(\mathbb{K}),\; \sum_{k\in\mathbb{K}} \phi_k(\xi_k) < +\infty,\; \text{ et } x \in \operatorname{dom} f \right\}. \tag{2.16}$$

Le problème primal est de

$$\underset{x\in\mathcal{H}}{\text{minimiser}} \;\; f(x) + \sum_{k\in\mathbb{K}} \phi_k(\langle x \mid e_k\rangle) + \frac{1}{2}\|x-z\|^2, \tag{2.17}$$

et le problème dual est de

$$\underset{(\nu_k)_{k\in\mathbb{K}}\in\ell^2(\mathbb{K})}{\text{minimiser}} \;\; \widetilde{f}^*\Big(z - \sum_{k\in\mathbb{K}} \nu_{n,k} e_k\Big) + \sum_{k\in\mathbb{K}} \phi_k^*(\nu_k). \tag{2.18}$$

Cet exemple est un cas particulier du Problème 2.1 avec $\mathcal{G} = \ell^2(\mathbb{K})$, $r = 0$, et

$$L\colon \mathcal{H} \to \mathcal{G}\colon x \mapsto (\langle x \mid e_k\rangle)_{k\in\mathbb{K}} \quad \text{et} \quad g\colon \mathcal{G} \to \left]-\infty,+\infty\right] : (\xi_k)_{k\in\mathbb{K}} \mapsto \sum_{k\in\mathbb{K}} \phi_k(\xi_k). \tag{2.19}$$

Enfin, nous présentons une application aux problèmes de débruitage de signaux avec des fonctions d'appui.

Exemple 2.8 Soient $z \in \mathcal{H}$, $r \in \mathcal{G}$, $f \in \Gamma_0(\mathcal{H})$, D un ensemble convexe fermé non vide de $\mathcal{G}$, et $0 \neq L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ tels que

$$r \in \operatorname{sri}\big(L(\operatorname{dom} f) - \big\{y \in \mathcal{G} \,\big|\, \sup_{v \in D} \langle y \mid v \rangle < +\infty\big\}\big). \tag{2.20}$$

Le problème primal est de

$$\underset{x \in \mathcal{H}}{\text{minimiser}} \;\; f(x) + \sigma_D(Lx - r) + \frac{1}{2}\|x - z\|^2, \tag{2.21}$$

et le problème dual est de

$$\underset{v \in D}{\text{minimiser}} \;\; \widetilde{f^*}(z - L^* v) + \langle v \mid r \rangle. \tag{2.22}$$

Cet exemple est un cas particulier du Problème 2.1 où $g = \sigma_D$ avec $\operatorname{dom} g = \big\{y \in \mathcal{G} \,\big|\, \sup_{v \in D} \langle y \mid v \rangle < +\infty\big\}$. On trouve des exemples de telles fonctions en débruitage de signaux dans [1, 8, 9, 24, 35, 39, 42, 65, 72, 79].

Remarque 2.9 En appliquant l'Algorithme 2.3 à ces cas particuliers, nous obtenons des algorithmes pour résoudre les problèmes dans les Exemples 2.5, 2.6, 2.7, et 2.8. Des liens avec des méthodes existantes sont présentés dans les Sections 2.2.4.1 et 2.2.4.3.

2.2 Article en anglais

DUALIZATION OF SIGNAL RECOVERY PROBLEMS [1]

Abstract : In convex optimization, duality theory can sometimes lead to simpler solution methods than those resulting from direct primal analysis. In this paper, this principle is applied to a class of composite variational problems arising in particular in signal recovery. These problems are not easily amenable to solution by current methods but they feature Fenchel-Moreau-Rockafellar dual problems that can be solved by forward-backward splitting. The proposed algorithm produces simultaneously a sequence converging weakly to a dual solution, and a sequence converging strongly to the primal solution. Our framework is shown to capture and extend several existing duality-based signal recovery methods and to be applicable to a variety of new problems beyond their scope.

1. P. L. Combettes, Đinh Dũng, and B. C. Vũ, Dualization of signal recovery problems, *Set-Valued Var. Anal.*, vol. 18, pp. 373–404, 2010.

2.2.1 Introduction

Over the years, several structured frameworks have been proposed to unify the analysis and the numerical solution methods of classes of signal (including image) recovery problems. An early contribution was made by Youla in 1978 [80]. He showed that several signal recovery problems, including those of [50, 66], shared a simple common geometrical structure and could be reduced to the following formulation in a Hilbert space $\mathcal{H}$ with scalar product $\langle \cdot \mid \cdot \rangle$ and associated norm $\|\cdot\|$: find the signal in a closed vector subspace C which admits a known projection r onto a closed vector subspace V, and which is at minimum distance from some reference signal z. This amounts to solving the variational problem

$$\underset{\substack{x \in C \\ P_V x = r}}{\text{minimize}} \; \frac{1}{2}\|x - z\|^2, \tag{2.23}$$

where P_V denotes the projector onto V. Abstract Hilbert space signal recovery problems have also been investigated by other authors. For instance, in 1965, Levi [56] considered the problem of finding the minimum energy band-limited signal fitting N linear measurements. In the Hilbert space $\mathcal{H} = L^2(\mathbb{R})$, the underlying variational problem is to

$$\underset{\substack{x \in C \\ \langle x | s_1 \rangle = \rho_1 \\ \vdots \\ \langle x | s_N \rangle = \rho_N}}{\text{minimize}} \; \frac{1}{2}\|x\|^2, \tag{2.24}$$

where C is the subspace of band-limited signals, $(s_i)_{1\leqslant i\leqslant N} \in \mathcal{H}^N$ are the measurement signals, and $(\rho_i)_{1\leqslant i\leqslant N} \in \mathbb{R}^N$ are the measurements. In [68], Potter and Arun observed that, for a general closed convex set C, the formulation (2.24) models a variety of problems, ranging from spectral estimation [10, 74] and tomography [58], to other inverse problems [12]. In addition, they employed an elegant duality framework to solve it, which led to the following result.

Proposition 2.10 [68, Theorems 1 and 3] *Set $r = (\rho_i)_{1\leqslant i\leqslant N}$ and $L\colon \mathcal{H} \to \mathbb{R}^N \colon x \mapsto (\langle x \mid s_i \rangle)_{1\leqslant i\leqslant N}$, and let $\gamma \in \left]0, 2\right[$. Suppose that $\sum_{i=1}^N \|s_i\|^2 \leqslant 1$ and that r lies in the relative interior of $L(C)$. Set*

$$w_0 \in \mathbb{R}^N \quad \textit{and} \quad (\forall n \in \mathbb{N}) \quad w_{n+1} = w_n + \gamma\big(r - LP_C L^* w_n\big), \tag{2.25}$$

where $L^*\colon \mathbb{R}^N \to \mathcal{H}\colon (\nu_i)_{1\leqslant i\leqslant N} \mapsto \sum_{i=1}^N \nu_i s_i$ is the adjoint of L. Then $(w_n)_{n\in\mathbb{N}}$ converges to a point w such that $LP_C L^* w = r$ and $P_C L^* w$ is the solution to (2.24).

Duality theory plays a central role in convex optimization [46, 62, 71, 83] and it has been used, in various forms and with different objectives, in several places in signal recovery, e.g., [10, 14, 23, 26, 39, 43, 47, 51, 53, 55, 79] ; let us add that, since the completion of the present paper [33], other aspects of duality in imaging have been investigated in [15]. For our purposes, the most suitable type of duality is the so-called Fenchel-Moreau-Rockafellar duality, which associates to a composite minimization problem a "dual" minimization problem involving the conjugates of the functions and the adjoint of the linear operator acting in the primal problem. In general, the dual problem sheds a new light on the properties of the primal problem and enriches its analysis. Moreover, in certain specific situations, it is actually possible to solve the dual problem and to recover a solution to the primal problem from any dual solution. Such a scenario underlies Proposition 2.10 : the primal problem (2.24) is difficult to solve but, if C is simple enough, the dual problem can be solved efficiently and, furthermore, a primal solution can be recovered explicitly. This principle is also explicitly or implicitly present in other signal recovery problems. For instance, the variational denoising problem

$$\underset{x\in\mathcal{H}}{\text{minimize}} \;\; g(Lx) + \frac{1}{2}\|x - z\|^2, \tag{2.26}$$

where z is a noisy observation of an ideal signal, L is a bounded linear operator from $\mathcal{H}$ to some Hilbert space $\mathcal{G}$, and $g\colon \mathcal{G} \to \left]-\infty, +\infty\right]$ is a proper lower semicontinuous convex function, can often be approached efficiently using duality arguments [39]. A popular development in this direction is the total variation denoising algorithm proposed in [23] and refined in [24].

The objective of the present paper is to devise a duality framework that captures problems such as (2.23), (2.24), and (2.26) and leads to improved algorithms and convergence results, in an effort to standardize the use of duality techniques in signal recovery and extend their range of potential applications. More specifically, we focus on a class of convex variational problems which satisfy the following.

(a) They cover the above minimization problems.

(b) They are not easy to solve directly, but they admit a Fenchel-Moreau-Rockafellar dual which can be solved reliably in the sense that an implementable algorithm is available with proven weak or strong convergence to a solution of the sequences of iterates it generates. Here "implementable" is taken in the classical sense of [67] : the algorithm does not involve subprograms (e.g., "oracles" or "black-boxes") which are not guaranteed to converge in a finite number of steps.

(c) They allow for the construction of a primal solution from any dual solution.

A problem formulation which complies with these requirements is the following, where we denote by $\operatorname{sri} C$ the strong relative interior of a convex set C (see (2.42) and Remark 2.12).

Problem 2.11 (primal problem) Let $\mathcal{H}$ and $\mathcal{G}$ be real Hilbert spaces, let $z \in \mathcal{H}$, let $r \in \mathcal{G}$, let $f\colon \mathcal{H} \to \left]-\infty, +\infty\right]$ and $g\colon \mathcal{G} \to \left]-\infty, +\infty\right]$ be lower semicontinuous convex functions, and let $L\colon \mathcal{H} \to \mathcal{G}$ be a nonzero linear bounded operator such that the qualification condition

$$r \in \operatorname{sri}\big(L(\operatorname{dom} f) - \operatorname{dom} g\big) \tag{2.27}$$

holds. The problem is to

$$\underset{x\in\mathcal{H}}{\text{minimize}} \;\; f(x) + g(Lx - r) + \frac{1}{2}\|x - z\|^2. \tag{2.28}$$

In connection with (a), it is clear that (2.28) covers (2.26) for $f = 0$. Moreover, if we let f and g be the indicator functions (see (2.38)) of closed convex sets $C \subset \mathcal{H}$ and $D \subset \mathcal{G}$, respectively, then (2.28) reduces to the best approximation problem

$$\underset{\substack{x\in C\\ Lx-r\in D}}{\text{minimize}} \;\; \frac{1}{2}\|x - z\|^2, \tag{2.29}$$

which captures both (2.23) and (2.24) in the case when C is a closed vector subspace and $D = \{0\}$. Indeed, (2.23) corresponds to $\mathcal{G} = \mathcal{H}$ and $L = P_V$, while (2.24) corresponds to $\mathcal{G} = \mathbb{R}^N$, $L\colon \mathcal{H} \to \mathbb{R}^N : x \mapsto (\langle x \mid s_i\rangle)_{1\leqslant i\leqslant N}$, $r = (\rho_i)_{1\leqslant i\leqslant N}$, and $z = 0$. As will be seen in Section 2.2.4, Problem 2.11 models a broad range of additional signal recovery problems.

In connection with (b), it is natural to ask whether the minimization problem (2.28) can be solved reliably by existing algorithms. Let us set

$$h\colon \mathcal{H} \to \,]-\infty, +\infty] : x \mapsto f(x) + g(Lx - r). \tag{2.30}$$

Then it follows from (2.27) that h is a proper lower semicontinuous convex function. Hence its proximity operator prox_h, which maps each $y \in \mathcal{H}$ to the unique minimizer of the function $x \mapsto h(x) + \|y - x\|^2/2$, is well defined (see Section 2.2.2.3). Accordingly, Problem 2.11 possesses a unique solution, which can be concisely written as

$$x = \text{prox}_h z. \tag{2.31}$$

Since no-closed form expression exists for the proximity operator of composite functions such as h, one can contemplate the use of splitting strategies to construct $\text{prox}_h z$ since (2.28) is of the form

$$\underset{x\in\mathcal{H}}{\text{minimize}} \;\; f_1(x) + f_2(x), \tag{2.32}$$

where

$$f_1\colon x \mapsto f(x) + \frac{1}{2}\|x - z\|^2 \quad \text{and} \quad f_2\colon x \mapsto g(Lx - r) \tag{2.33}$$

are lower semicontinuous convex functions from $\mathcal{H}$ to $]-\infty, +\infty]$. To tackle (2.32), a first splitting framework is that described in [39], which requires the additional assumption that f_2 be Lipschitz-differentiable on $\mathcal{H}$ (see also [13, 16, 20, 19, 27, 35, 42, 49] for recent work within this setting). In this case, (2.32) can be solved by the proximal forward-backward algorithm, which is governed by the updating rule

$$\left\lfloor \begin{array}{ll} x_{n+\frac{1}{2}} &= \nabla f_2(x_n) + a_{2,n} \\ x_{n+1} &= x_n + \lambda_n\Big(\text{prox}_{\gamma_n f_1}\big(x_n - \gamma_n x_{n+\frac{1}{2}}\big) + a_{1,n} - x_n\Big), \end{array}\right. \tag{2.34}$$

where $\lambda_n > 0$ and $\gamma_n > 0$, and where $a_{1,n}$ and $a_{2,n}$ model respectively tolerances in the approximate implementation of the proximity operator of f_1 and the gradient of f_2. Precise convergence results for the iterates $(x_n)_{n\in\mathbb{N}}$ can be found in Theorem 2.37. Let us add that there exist variants of this splitting method, which do not guarantee convergence of the iterates but do provide an optimal (in the sense of [63]) $O(1/n^2)$ rate of convergence of the objective values [8]. A limitation of this first framework is that it imposes that g be Lipschitz-differentiable and therefore excludes key problems such as (2.29). An alternative framework, which does not demand any smoothness assumption in (2.32), is investigated in [36]. It employs the Douglas-Rachford splitting algorithm, which revolves around the updating rule

$$\left\lfloor \begin{array}{ll} x_{n+\frac{1}{2}} &= \operatorname{prox}_{\gamma f_2} x_n + a_{2,n} \\ x_{n+1} &= x_n + \lambda_n \Big(\operatorname{prox}_{\gamma f_1} \big(2x_{n+\frac{1}{2}} - x_n\big) + a_{1,n} - x_{n+\frac{1}{2}} \Big), \end{array} \right. \tag{2.35}$$

where $\lambda_n > 0$ and $\gamma > 0$, and where $a_{1,n}$ and $a_{2,n}$ model tolerances in the approximate implementation of the proximity operators of f_1 and f_2, respectively (see [36, Theorem 20] for precise convergence results and [28] for further applications). However, this approach requires that the proximity operator of the composite function f_2 in (2.33) be computable to within some quantifiable error. Unfortunately, this is not possible in general, as explicit expressions of $\operatorname{prox}_{g\circ L}$ in terms of prox_g require stringent assumptions, for instance $L \circ L^* = \kappa \operatorname{Id}$ for some $\kappa > 0$ (see Example 2.19), which does not hold in the case of (2.24) and many other important problems. A third framework that appears to be relevant is that of [5], which is tailored for problems of the form

$$\underset{x\in\mathcal{H}}{\text{minimize}} \;\; h_1(x) + h_2(x) + \frac{1}{2}\|x - z\|^2, \tag{2.36}$$

where h_1 and h_2 are lower semicontinuous convex functions from $\mathcal{H}$ to $]-\infty, +\infty]$ such that $\operatorname{dom} h_1 \cap \operatorname{dom} h_2 \neq \varnothing$. This formulation coincides with our setting for $h_1 = f$ and $h_2 \colon x \mapsto g(Lx - r)$. The Dykstra-like algorithm devised in [5] to solve (2.36) is governed by the iteration

$$\begin{array}{l} \text{Initialization} \\ \left\lfloor \begin{array}{l} y_0 = z \\ q_0 = 0 \\ p_0 = 0 \end{array} \right. \\ \text{For } n = 0, 1, \ldots \\ \left\lfloor \begin{array}{lcl} x_n &=& \operatorname{prox}_{h_2}(y_n + q_n) \\ q_{n+1} &=& y_n + q_n - x_n \\ y_{n+1} &=& \operatorname{prox}_{h_1}(x_n + p_n) \\ p_{n+1} &=& x_n + p_n - y_{n+1} \end{array} \right. \end{array} \tag{2.37}$$

and therefore requires that the proximity operators of h_1 and h_2 be computable explicitly. As just discussed, this is seldom possible in the case of the composite function h_2.

To sum up, existing splitting techniques do not offer satisfactory options to solve Problem 2.11 and alternative routes must be explored. The cornerstone of our paper is that, by contrast, Problem 2.11 can be solved reliably via Fenchel-Moreau-Rockafellar duality so long as the operators prox_f and prox_g can be evaluated to within some quantifiable error, which will be shown to be possible in a wide variety of problems.

The paper is organized as follows. In Section 2.2.2 we provide the convex analytical background required in subsequent sections and, in particular, we review proximity operators. In Section 2.2.3, we show that Problem 2.11 satisfies properties (b) and (c). We then derive the Fenchel-Moreau-Rockafellar dual of Problem 2.11 and then show that it is amenable to solution by forward-backward splitting. The resulting primal-dual algorithm involves the functions f and g, as well as the operator L, separately and therefore achieves full splitting of the constituents of the primal problem. We show that the primal sequence produced by the algorithm converges strongly to the solution to Problem 2.11, and that the dual sequence converges weakly to a solution to the dual problem. Finally, in Section 2.2.4, we highlight applications of the proposed duality framework to best approximation problems, denoising problems using dictionaries, and recovery problems involving support functions. In particular, we extend and provide formal convergence results for the total variation denoising algorithm proposed in [24]. Although signal recovery applications are emphasized in the present paper, the proposed duality framework is applicable to any variational problem conforming to the format described in Problem 2.11.

2.2.2 Convex-analytical tools

2.2.2.1 General notation

Throughout the paper, $\mathcal{H}$ and $\mathcal{G}$ are real Hilbert spaces, and $\mathcal{B}(\mathcal{H},\mathcal{G})$ is the space of bounded linear operators from $\mathcal{H}$ to $\mathcal{G}$. The identity operator is denoted by Id, the adjoint of an operator $T \in \mathcal{B}(\mathcal{H},\mathcal{G})$ by T^*, the scalar products of both $\mathcal{H}$ and $\mathcal{G}$ by $\langle \cdot \mid \cdot \rangle$ and the associated norms by $\|\cdot\|$. Moreover, $\rightharpoonup$ and $\to$ denote respectively weak and strong convergence. Finally, we denote by $\Gamma_0(\mathcal{H})$ the class of lower semicontinuous convex functions $\varphi\colon \mathcal{H} \to \left]-\infty,+\infty\right]$ which are proper in the sense that $\text{dom}\,\varphi = \big\{x \in \mathcal{H} \bigm| \varphi(x) < +\infty\big\} \neq \varnothing$.

2.2.2.2 Convex sets and functions

We provide some background on convex analysis; for a detailed account, see [83] and, for finite-dimensional spaces, [70].

Let C be a nonempty convex subset of $\mathcal{H}$. The indicator function of C is

$$\iota_C : x \mapsto \begin{cases} 0, & \text{if } x \in C; \\ +\infty, & \text{if } x \notin C, \end{cases} \tag{2.38}$$

the distance function of C is

$$d_C : \mathcal{H} \to [0, +\infty[\, : x \mapsto \inf_{y \in C} \|x - y\|, \tag{2.39}$$

the support function of C is

$$\sigma_C : \mathcal{H} \to \,]-\infty, +\infty] : u \mapsto \sup_{x \in C} \langle x \mid u \rangle, \tag{2.40}$$

and the conical hull of C is

$$\operatorname{cone} C = \bigcup_{\lambda > 0} \big\{ \lambda x \bigm| x \in C \big\}. \tag{2.41}$$

If C is also closed, the projection of a point x in $\mathcal{H}$ onto C is the unique point $P_C x$ in C such that $\|x - P_C x\| = d_C(x)$. We denote by $\operatorname{int} C$ the interior of C, by $\operatorname{span} C$ the span of C, and by $\overline{\operatorname{span}}\, C$ the closure of $\operatorname{span} C$. The core of C is $\operatorname{core} C = \big\{ x \in C \bigm| \operatorname{cone}(C - x) = \mathcal{H} \big\}$, the strong relative interior of C is

$$\operatorname{sri} C = \big\{ x \in C \bigm| \operatorname{cone}(C - x) = \overline{\operatorname{span}}\,(C - x) \big\}, \tag{2.42}$$

and the relative interior of C is $\operatorname{ri} C = \big\{ x \in C \bigm| \operatorname{cone}(C - x) = \operatorname{span}\,(C - x) \big\}$. We have

$$\operatorname{int} C \subset \operatorname{core} C \subset \operatorname{sri} C \subset \operatorname{ri} C \subset C. \tag{2.43}$$

The strong relative interior is therefore an extension of the notion of an interior. This extension is particularly important in convex analysis as many useful sets have empty interior infinite-dimensional spaces.

Remark 2.12 The qualification condition (2.27) in Problem 2.11 is rather mild. In view of (2.43), it is satisfied in particular when r belongs to the core and, a fortiori, to the interior of $L(\operatorname{dom} f) - \operatorname{dom} g$; the latter is for instance satisfied when $L(\operatorname{dom} f) \cap (r + \operatorname{int} \operatorname{dom} g) \neq \varnothing$. If f and g are proper, then (2.27) is also satisfied when $L(\operatorname{dom} f) - \operatorname{dom} g = \mathcal{H}$ and, a fortiori, when f is finite-valued and L is surjective, or when g is finite-valued. If $\mathcal{G}$ is finite-dimensional, then (2.27) reduces to [70, Section 6]

$$r \in \operatorname{ri} \big(L(\operatorname{dom} f) - \operatorname{dom} g \big) = (\operatorname{ri} L(\operatorname{dom} f)) - \operatorname{ri} \operatorname{dom} g, \tag{2.44}$$

i.e., $(\operatorname{ri} L(\operatorname{dom} f)) \cap (r + \operatorname{ri} \operatorname{dom} g) \neq \varnothing$.

Let $\varphi \in \Gamma_0(\mathcal{H})$. The conjugate of φ is the function $\varphi^* \in \Gamma_0(\mathcal{H})$ defined by

$$(\forall u \in \mathcal{H}) \quad \varphi^*(u) = \sup_{x \in \mathcal{H}} \langle x \mid u \rangle - \varphi(x). \tag{2.45}$$

The Fenchel-Moreau theorem states that $\varphi^{**} = \varphi$. The subdifferential of φ is the set-valued operator

$$\partial\varphi \colon \mathcal{H} \to 2^{\mathcal{H}} \colon x \mapsto \big\{u \in \mathcal{H} \;\big|\; (\forall y \in \mathcal{H}) \;\; \langle y - x \mid u \rangle + \varphi(x) \leqslant \varphi(y)\big\}. \tag{2.46}$$

We have

$$(\forall (x,u) \in \mathcal{H} \times \mathcal{H}) \quad u \in \partial\varphi(x) \quad \Leftrightarrow \quad x \in \partial\varphi^*(u). \tag{2.47}$$

Moreover, if φ is Gâteaux differentiable at x, then

$$\partial\varphi(x) = \{\nabla\varphi(x)\}. \tag{2.48}$$

Fermat's rule states that

$$(\forall x \in \mathcal{H}) \quad x \in \operatorname{Argmin} \varphi = \big\{x \in \operatorname{dom} \varphi \;\big|\; (\forall y \in \mathcal{H}) \;\; \varphi(x) \leqslant \varphi(y)\big\} \quad \Leftrightarrow \quad 0 \in \partial\varphi(x). \tag{2.49}$$

If $\operatorname{Argmin} \varphi$ is a singleton, we denote by $\operatorname{argmin}_{y \in \mathcal{H}} \varphi(y)$ the unique minimizer of φ.

Lemma 2.13 [83, Theorem 2.8.3] *Let $\varphi \in \Gamma_0(\mathcal{H})$, let $\psi \in \Gamma_0(\mathcal{G})$, and let $M \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ be such that $0 \in \operatorname{sri}(M(\operatorname{dom} \varphi) - \operatorname{dom} \psi)$. Then $\partial(\varphi + \psi \circ M) = \partial\varphi + M^* \circ (\partial\psi) \circ M$.*

2.2.2.3 Moreau envelopes and proximity operators

Essential to this paper is the notion of a proximity operator, which is due to Moreau [60] (see [39, 61] for detailed accounts and Section 2.2.2.4 for closed-form examples). The Moreau envelope of φ is the continuous convex function

$$\widetilde{\varphi} \colon \mathcal{H} \to \mathbb{R} \colon x \mapsto \min_{y \in \mathcal{H}} \varphi(y) + \frac{1}{2}\|x - y\|^2. \tag{2.50}$$

For every $x \in \mathcal{H}$, the function $y \mapsto \varphi(y) + \|x - y\|^2/2$ admits a unique minimizer, which is denoted by $\operatorname{prox}_\varphi x$. The proximity operator of φ is defined by

$$\operatorname{prox}_\varphi \colon \mathcal{H} \to \mathcal{H} \colon x \mapsto \operatorname*{argmin}_{y \in \mathcal{H}} \varphi(y) + \frac{1}{2}\|x - y\|^2 \tag{2.51}$$

and characterized by

$$(\forall (x,p) \in \mathcal{H} \times \mathcal{H}) \quad p = \operatorname{prox}_\varphi x \quad \Leftrightarrow \quad x - p \in \partial\varphi(p). \tag{2.52}$$

Lemma 2.14 [61] *Let $\varphi \in \Gamma_0(\mathcal{H})$. Then the following hold.*

(i) $(\forall x \in \mathcal{H})(\forall y \in \mathcal{H})\ \|\mathrm{prox}_\varphi x - \mathrm{prox}_\varphi y\|^2 \leqslant \langle x - y \mid \mathrm{prox}_\varphi x - \mathrm{prox}_\varphi y\rangle$.

(ii) $(\forall x \in \mathcal{H})(\forall y \in \mathcal{H})\ \|\mathrm{prox}_\varphi x - \mathrm{prox}_\varphi y\| \leqslant \|x - y\|$.

(iii) $\widetilde{\varphi} + \widetilde{\varphi^*} = \|\cdot\|^2/2$.

(iv) $\widetilde{\varphi^*}$ *is Fréchet differentiable and* $\nabla\widetilde{\varphi^*} = \mathrm{prox}_\varphi = \mathrm{Id} - \mathrm{prox}_{\varphi^*}$.

The identity $\mathrm{prox}_\varphi = \mathrm{Id} - \mathrm{prox}_{\varphi^*}$ can be stated in a slightly extended context.

Lemma 2.15 [39, Lemma 2.10] *Let $\varphi \in \Gamma_0(\mathcal{H})$, let $x \in \mathcal{H}$, and let $\gamma \in \left]0, +\infty\right[$. Then* $x = \mathrm{prox}_{\gamma\varphi} x + \gamma \mathrm{prox}_{\gamma^{-1}\varphi^*}(\gamma^{-1}x)$.

The following fact will also be required.

Lemma 2.16 *Let $\psi \in \Gamma_0(\mathcal{H})$, let $w \in \mathcal{H}$, and set $\varphi\colon x \mapsto \psi(x) + \|x - w\|^2/2$. Then* $\varphi^*\colon u \mapsto \widetilde{\psi^*}(u + w) - \|w\|^2/2$.

Proof. Let $u \in \mathcal{H}$. It follows from (2.45) and Lemma 2.14(iii) that

$$\begin{aligned}
\varphi^*(u) &= -\inf_{x\in\mathcal{H}} \psi(x) + \frac{1}{2}\|x - w\|^2 - \langle x \mid u\rangle \\
&= \frac{1}{2}\|u\|^2 + \langle w \mid u\rangle - \inf_{x\in\mathcal{H}} \psi(x) + \frac{1}{2}\|x - (w + u)\|^2 \\
&= \frac{1}{2}\|u + w\|^2 - \frac{1}{2}\|w\|^2 - \widetilde{\psi}(u + w) \\
&= \widetilde{\psi^*}(u + w) - \frac{1}{2}\|w\|^2, \qquad (2.53)
\end{aligned}$$

which yields the desired identity. □

2.2.2.4 Examples of proximity operators

To solve Problem 2.11, our algorithm will use (approximate) evaluations of the proximity operators of the functions f and g^* (or, equivalently, of g by Lemma 2.14(iv)). In this section, we supply examples of proximity operators which admit closed-form expressions.

Example 2.17 Let C be a nonempty closed convex subset of $\mathcal{H}$. Then the following hold.

(i) Set $\varphi = \iota_C$. Then $\mathrm{prox}_\varphi = P_C$ [61, Example 3.d].

(ii) Set $\varphi = \sigma_C$. Then $\mathrm{prox}_\varphi = \mathrm{Id} - P_C$ [39, Example 2.17].

(iii) Set $\varphi = d_C^2/(2\alpha)$. Then $(\forall x \in \mathcal{H})\ \mathrm{prox}_\varphi x = x + (1 + \alpha)^{-1}(P_C x - x)$ [39, Example 2.14].

(iv) Set $\varphi = (\|\cdot\|^2 - d_C^2)/(2\alpha)$. Then $(\forall x \in \mathcal{H})$ $\operatorname{prox}_\varphi x = x - \alpha^{-1}P_C(\alpha(\alpha+1)^{-1}x)$ [39, Lemma 2.7].

Example 2.18 [39, Lemma 2.7] Let $\psi \in \Gamma_0(\mathcal{H})$ and set $\varphi = \|\cdot\|^2/2 - \widetilde{\psi}$. Then $\varphi \in \Gamma_0(\mathcal{H})$ and $(\forall x \in \mathcal{H})$ $\operatorname{prox}_\varphi x = x - \operatorname{prox}_{\psi/2}(x/2)$.

Example 2.19 [36, Proposition 11] Let $\mathcal{G}$ be a real Hilbert space, let $\psi \in \Gamma_0(\mathcal{G})$, let $M \in \mathcal{B}(\mathcal{H},\mathcal{G})$, and set $\varphi = \psi \circ M$. Suppose that $M \circ M^* = \kappa \operatorname{Id}$, for some $\kappa \in \left]0,+\infty\right[$. Then $\varphi \in \Gamma_0(\mathcal{H})$ and

$$\operatorname{prox}_\varphi = \operatorname{Id} + \frac{1}{\kappa} M^* \circ (\operatorname{prox}_{\kappa\psi} - \operatorname{Id}) \circ M. \tag{2.54}$$

Example 2.20 [27, Proposition 2.10 and Remark 3.2(ii)] Set

$$\varphi \colon \mathcal{H} \to \left]-\infty,+\infty\right] : x \mapsto \sum_{k\in\mathbb{K}} \phi_k(\langle x \mid o_k\rangle), \tag{2.55}$$

where :

(i) $\varnothing \neq \mathbb{K} \subset \mathbb{N}$;

(ii) $(o_k)_{k\in\mathbb{K}}$ is an orthonormal basis of $\mathcal{H}$;

(iii) $(\phi_k)_{k\in\mathbb{K}}$ are functions in $\Gamma_0(\mathbb{R})$;

(iv) Either $\mathbb{K}$ is finite, or there exists a subset $\boldsymbol{L}$ of $\mathbb{K}$ such that :

  (a) $\mathbb{K} \smallsetminus \boldsymbol{L}$ is finite;

  (b) $(\forall k \in \boldsymbol{L})$ $\phi_k \geqslant \phi_k(0) = 0$.

Then $\varphi \in \Gamma_0(\mathcal{H})$ and

$$(\forall x \in \mathcal{H}) \quad \operatorname{prox}_\varphi x = \sum_{k\in\mathbb{K}} \big(\operatorname{prox}_{\phi_k}\langle x \mid o_k\rangle\big) o_k. \tag{2.56}$$

Example 2.21 [17, Proposition 2.1] Let C be a nonempty closed convex subset of $\mathcal{H}$, let $\phi \in \Gamma_0(\mathbb{R})$ be even, and set $\varphi = \phi \circ d_C$. Then $\varphi \in \Gamma_0(\mathcal{H})$. Moreover, $\operatorname{prox}_\varphi = P_C$ if $\phi = \iota_{\{0\}} + \eta$ for some $\eta \in \mathbb{R}$ and, otherwise,

$$(\forall x \in \mathcal{H}) \quad \operatorname{prox}_\varphi x = \begin{cases} x + \dfrac{\operatorname{prox}_{\phi^*} d_C(x)}{d_C(x)}(P_C x - x), & \text{if } d_C(x) > \max \partial\phi(0); \\ P_C x, & \text{if } x \notin C \text{ and } d_C(x) \leqslant \max\partial\phi(0); \\ x, & \text{if } x \in C. \end{cases} \tag{2.57}$$

Remark 2.22 Taking $C = \{0\}$ and $\phi \neq \iota_{\{0\}} + \eta$ ($\eta \in \mathbb{R}$) in Example 2.21 yields the proximity operator of $\phi \circ \|\cdot\|$, namely (using Lemma 2.14(iv))

$$(\forall x \in \mathcal{H}) \quad \operatorname{prox}_{\varphi} x = \begin{cases} \dfrac{\operatorname{prox}_{\phi}\|x\|}{\|x\|}x, & \text{if } \|x\| > \max \partial\phi(0); \\ 0, & \text{if } \|x\| \leqslant \max \partial\phi(0). \end{cases} \tag{2.58}$$

On the other hand, if ϕ is differentiable at 0 in Example 2.21, then $\partial\phi(0) = \{0\}$ and (2.57) yields

$$(\forall x \in \mathcal{H}) \quad \operatorname{prox}_{\varphi} x = \begin{cases} x + \dfrac{\operatorname{prox}_{\phi^*} d_C(x)}{d_C(x)}(P_C x - x), & \text{if } x \notin C; \\ x, & \text{if } x \in C. \end{cases} \tag{2.59}$$

Example 2.23 [17, Proposition 2.2] Let C be a nonempty closed convex subset of $\mathcal{H}$, let $\phi \in \Gamma_0(\mathbb{R})$ be even and nonconstant, and set $\varphi = \sigma_C + \phi \circ \|\cdot\|$. Then $\varphi \in \Gamma_0(\mathcal{H})$ and

$$(\forall x \in \mathcal{H}) \quad \operatorname{prox}_{\varphi} x = \begin{cases} \dfrac{\operatorname{prox}_{\phi} d_C(x)}{d_C(x)}(x - P_C x), & \text{if } d_C(x) > \max \operatorname{Argmin} \phi; \\ x - P_C x, & \text{if } x \notin C \text{ and } d_C(x) \leqslant \max \operatorname{Argmin} \phi; \\ 0, & \text{if } x \in C. \end{cases} \tag{2.60}$$

Example 2.24 Let $A \in \mathcal{B}(\mathcal{H})$ be positive and self-adjoint, let $b \in \mathcal{H}$, let $\alpha \in \mathbb{R}$, and set $\varphi\colon x \mapsto \langle Ax \mid x\rangle/2 + \langle x \mid b\rangle + \alpha$. Then $\varphi \in \Gamma_0(\mathcal{H})$ and $(\forall x \in \mathcal{H})$ $\operatorname{prox}_{\varphi} x = (\operatorname{Id} + A)^{-1}(x - b)$.

Proof. It is clear that φ is a finite-valued continuous convex function. Now fix $x \in \mathcal{H}$ and set $\psi\colon y \mapsto \|x - y\|^2/2 + \langle Ay \mid y\rangle/2 + \langle y \mid b\rangle + \alpha$. Then $\nabla\psi\colon y \mapsto y - x + Ay + b$. Hence, $(\forall y \in \mathcal{H})$ $\nabla\psi(y) = 0 \Leftrightarrow y = (\operatorname{Id} + A)^{-1}(x - b)$. ☐

Example 2.25 For every $i \in \{1, \ldots, m\}$, let $(\mathcal{G}_i, \|\cdot\|)$ be a real Hilbert space, let $r_i \in \mathcal{G}_i$, let $T_i \in \mathcal{B}(\mathcal{H}, \mathcal{G}_i)$, and let $\alpha_i \in \left]0, +\infty\right[$. Set $(\forall x \in \mathcal{H})$ $\varphi(x) = (1/2)\sum_{i=1}^m \alpha_i \|T_i x - r_i\|^2$. Then $\varphi \in \Gamma_0(\mathcal{H})$ and

$$(\forall x \in \mathcal{H}) \quad \operatorname{prox}_{\varphi} x = \left(\operatorname{Id} + \sum_{i=1}^m \alpha_i T_i^* T_i\right)^{-1}\left(x + \sum_{i=1}^m \alpha_i T_i^* r_i\right). \tag{2.61}$$

Proof. We have $\varphi\colon x \mapsto \sum_{i=1}^m \alpha_i \langle T_i x - r_i \mid T_i x - r_i\rangle/2 = \langle Ax \mid x\rangle/2 + \langle x \mid b\rangle + \alpha$, where $A = \sum_{i=1}^m \alpha_i T_i^* T_i$, $b = -\sum_{i=1}^m \alpha_i T_i^* r_i$, and $\alpha = \sum_{i=1}^m \alpha_i \|r_i\|^2/2$. Hence, (2.61) follows from Example 2.24. ☐

As seen in Example 2.20, Example 2.21, Remark 2.22, and Example 2.23, some important proximity operators can be decomposed in terms of those of functions in $\Gamma_0(\mathbb{R})$. Here are explicit expressions for the proximity operators of such functions.

Example 2.26 [27, Examples 4.2 and 4.4] Let $p \in [1,+\infty[$, let $\alpha \in]0,+\infty[$, let $\phi\colon \mathbb{R} \to \mathbb{R}\colon \eta \mapsto \alpha|\eta|^p$, let $\xi \in \mathbb{R}$, and set $\pi = \operatorname{prox}_\phi \xi$. Then the following hold.

(i) $\pi = \operatorname{sign}(\xi)\max\{|\xi|-\alpha, 0\}$, if $p = 1$;

(ii) $\pi = \xi + \dfrac{4\alpha}{3\cdot 2^{1/3}}\Big(|\rho-\xi|^{1/3} - |\rho+\xi|^{1/3}\Big)$, where $\rho = \sqrt{\xi^2 + 256\alpha^3/729}$, if $p = 4/3$;

(iii) $\pi = \xi + 9\alpha^2 \operatorname{sign}(\xi)\big(1 - \sqrt{1+16|\xi|/(9\alpha^2)}\,\big)/8$, if $p = 3/2$;

(iv) $\pi = \xi/(1+2\alpha)$, if $p = 2$;

(v) $\pi = \operatorname{sign}(\xi)\big(\sqrt{1+12\alpha|\xi|} - 1\big)/(6\alpha)$, if $p = 3$;

(vi) $\pi = \left|\dfrac{\rho+\xi}{8\alpha}\right|^{1/3} - \left|\dfrac{\rho-\xi}{8\alpha}\right|^{1/3}$, where $\rho = \sqrt{\xi^2 + 1/(27\alpha)}$, if $p = 4$.

Example 2.27 [39, Example 2.18] Let $\alpha \in]0,+\infty[$ and set

$$\phi\colon \xi \mapsto \begin{cases} -\alpha\ln(\xi), & \text{if } \xi > 0;\\ +\infty, & \text{if } \xi \leqslant 0. \end{cases} \tag{2.62}$$

Then $(\forall \xi \in \mathbb{R})$ $\operatorname{prox}_\phi \xi = (\xi + \sqrt{\xi^2+4\alpha})/2$.

Example 2.28 [35, Example 3.5] Let $\omega \in]0,+\infty[$ and set

$$\phi\colon \mathbb{R} \to]-\infty,+\infty] : \xi \mapsto \begin{cases} \ln(\omega) - \ln(\omega - |\xi|), & \text{if } |\xi| < \omega;\\ +\infty, & \text{otherwise.} \end{cases} \tag{2.63}$$

Then

$$(\forall \xi \in \mathbb{R}) \quad \operatorname{prox}_\phi \xi = \begin{cases} \operatorname{sign}(\xi)\,\dfrac{|\xi| + \omega - \sqrt{\big||\xi|-\omega\big|^2 + 4}}{2}, & \text{if } |\xi| > 1/\omega;\\ 0 & \text{otherwise.} \end{cases} \tag{2.64}$$

Example 2.29 [27, Example 4.5] Let $\omega \in]0,+\infty[$, $\tau \in]0,+\infty[$, and set

$$\phi\colon \mathbb{R} \to]-\infty,+\infty] : \xi \mapsto \begin{cases} \tau\xi^2, & \text{if } |\xi| \leqslant \omega/\sqrt{2\tau};\\ \omega\sqrt{2\tau}|\xi| - \omega^2/2, & \text{otherwise.} \end{cases} \tag{2.65}$$

Then

$$(\forall \xi \in \mathbb{R}) \quad \operatorname{prox}_\phi \xi = \begin{cases} \dfrac{\xi}{2\tau+1}, & \text{if } |\xi| \leqslant \omega(2\tau+1)/\sqrt{2\tau};\\ \xi - \omega\sqrt{2\tau}\operatorname{sign}(\xi), & \text{if } |\xi| > \omega(2\tau+1)/\sqrt{2\tau}. \end{cases} \tag{2.66}$$

Further examples can be constructed via the following rules.

Lemma 2.30 [35, Proposition 3.6] *Let $\phi = \psi + \sigma_\Omega$, where $\psi \in \Gamma_0(\mathbb{R})$ and $\Omega \subset \mathbb{R}$ is a nonempty closed interval. Suppose that ψ is differentiable at 0 with $\psi'(0) = 0$. Then* $\operatorname{prox}_\phi = \operatorname{prox}_\psi \circ \operatorname{soft}_\Omega$*, where*

$$\operatorname{soft}_\Omega : \mathbb{R} \to \mathbb{R} \colon \xi \mapsto \begin{cases} \xi - \underline{\omega}, & \text{if } \xi < \underline{\omega}; \\ 0, & \text{if } \xi \in \Omega; \\ \xi - \overline{\omega}, & \text{if } \xi > \overline{\omega}, \end{cases} \qquad \text{with} \quad \begin{cases} \underline{\omega} = \inf \Omega, \\ \overline{\omega} = \sup \Omega. \end{cases} \tag{2.67}$$

Lemma 2.31 [36, Proposition 12(ii)] *Let $\phi = \iota_C + \psi$, where $\psi \in \Gamma_0(\mathbb{R})$ and where C is a closed interval in $\mathbb{R}$ such that $C \cap \operatorname{dom} \psi \neq \varnothing$. Then* $\operatorname{prox}_{\iota_C + \psi} = P_C \circ \operatorname{prox}_\psi$.

2.2.3 Dualization and algorithm

2.2.3.1 Fenchel-Moreau-Rockafellar duality

Our analysis will revolve around the following version of the Fenchel-Moreau-Rockafellar duality formula (see [48], [62], and [69] for historical work). It will also exploit various aspects of the Baillon-Haddad theorem [6].

Lemma 2.32 [83, Corollary 2.8.5] *Let $\varphi \in \Gamma_0(\mathcal{H})$, let $\psi \in \Gamma_0(\mathcal{G})$, and let $M \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ be such that $0 \in \operatorname{sri}(M(\operatorname{dom}\varphi) - \operatorname{dom}\psi)$. Then*

$$\inf_{x \in \mathcal{H}} \varphi(x) + \psi(Mx) = -\min_{v \in \mathcal{G}} \varphi^*(-M^*v) + \psi^*(v). \tag{2.68}$$

The problem of minimizing $\varphi + \psi \circ M$ on $\mathcal{H}$ in (2.68) is referred to as the primal problem, and that of minimizing $\varphi^* \circ (-M^*) + \psi^*$ on $\mathcal{G}$ as the dual problem. Lemma 2.32 gives conditions under which a dual solution exists and the value of the dual problem coincides with the opposite of the value of the primal problem. We can now introduce the dual of Problem 2.11.

Problem 2.33 (dual problem) Under the same assumptions as in Problem 2.11,

$$\underset{v \in \mathcal{G}}{\text{minimize}} \;\; \widetilde{f}^*(z - L^*v) + g^*(v) + \langle v \mid r \rangle. \tag{2.69}$$

Proposition 2.34 *Problem 2.33 is the dual of Problem 2.11 and it admits at least one solution. Moreover, every solution v to Problem 2.33 is characterized by the inclusion*

$$L\big(\operatorname{prox}_f(z - L^*v)\big) - r \in \partial g^*(v). \tag{2.70}$$

Proof. Let us set $w = z$, $\varphi = f + \|\cdot - w\|^2/2$, $M = L$, and $\psi = g(\cdot - r)$. Then $(\forall x \in \mathcal{H})$ $\varphi(x) + \psi(Mx) = f(x) + g(Lx - r) + \|x - z\|^2/2$. Hence, it results from (2.68) and

Lemma 2.16 that the dual of Problem 2.11 is to minimize the function

$$\begin{aligned}\varphi^* \circ (-M^*) + \psi^* \colon v \mapsto \widetilde{f}^*(-M^*v + w) - \frac{1}{2}\|w\|^2 + \psi^*(v)\\ = \widetilde{f}^*(z - L^*v) - \frac{1}{2}\|z\|^2 + g^*(v) + \langle v \mid r\rangle \end{aligned} \tag{2.71}$$

or, equivalently, the function $v \mapsto \widetilde{f}^*(z - L^*v) + g^*(v) + \langle v \mid r\rangle$. In view of (2.27), the first two claims therefore follow from Lemma 2.32. To establish the last claim, note that (2.50) asserts that $\operatorname{dom} \widetilde{f}^* \circ (z - L^* \cdot) = \mathcal{G}$. Hence, using (2.49), Lemma 2.13, (2.48), and Lemma 2.14(iv), we get

$$\begin{aligned} v \text{ solves (2.69)} \quad &\Leftrightarrow \quad 0 \in \partial\Big(\widetilde{f}^* \circ (z - L^*\cdot) + g^* + \langle \cdot \mid r\rangle\Big)(v)\\ &\Leftrightarrow \quad 0 \in -L\big(\nabla \widetilde{f}^*(z - L^*v)\big) + \partial g^*(v) + r\\ &\Leftrightarrow \quad 0 \in -L\big(\operatorname{prox}_f(z - L^*v)\big) + \partial g^*(v) + r, \end{aligned} \tag{2.72}$$

which yields (2.70). □

A key property underlying our setting is that the primal solution can actually be recovered from any dual solution (this is property (c) in the Introduction).

Proposition 2.35 *Let v be a solution to Problem 2.33 and set*

$$x = \operatorname{prox}_f(z - L^*v). \tag{2.73}$$

Then x is the solution to Problem 2.11.

Proof. We derive from (2.73) and (2.52) that $z - L^*v - x \in \partial f(x)$. Therefore

$$-L^*v \in \partial f(x) + x - z. \tag{2.74}$$

On the other hand, it follows from (2.70), (2.73), and (2.47) that

$$\begin{aligned} v \text{ solves (2.69)} \quad &\Leftrightarrow \quad Lx - r \in \partial g^*(v)\\ &\Leftrightarrow \quad v \in \partial g(Lx - r)\\ &\Rightarrow \quad L^*v \in L^*\big(\partial g(Lx - r)\big). \end{aligned} \tag{2.75}$$

Upon adding (2.74) and (2.75), invoking Lemma 2.13, and then (2.49) we obtain

$$\begin{aligned} v \text{ solves (2.69)} \quad &\Rightarrow \quad 0 = L^*v - L^*v\\ &\qquad \in \partial f(x) + L^*\big(\partial g(Lx - r)\big) + x - z\\ &\qquad = \partial f(x) + L^*\big(\partial g(Lx - r)\big) + \nabla\Big(\frac{1}{2}\|\cdot - z\|^2\Big)(x)\\ &\qquad = \partial\Big(f + g(L\cdot - r) + \frac{1}{2}\|\cdot - z\|^2\Big)(x)\\ &\Leftrightarrow \quad x \text{ solves (2.28)}, \end{aligned} \tag{2.76}$$

which completes the proof. □

2.2.3.2 Algorithm

As seen in (2.31), the unique solution to Problem 2.11 is $\operatorname{prox}_h z$, where h is defined in (2.30). Since $\operatorname{prox}_h z$ cannot be computed directly, it will be constructed iteratively by the following algorithm, which produces a primal sequence $(x_n)_{n\in\mathbb{N}}$ as well as a dual sequence $(v_n)_{n\in\mathbb{N}}$.

Algorithm 2.36 Let $(a_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{G}$ such that $\sum_{n\in\mathbb{N}}\|a_n\| < +\infty$ and let $(b_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that $\sum_{n\in\mathbb{N}}\|b_n\| < +\infty$. Sequences $(x_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}}$ are generated by the following routine.

$$
\begin{array}{l}
\text{Initialization}\\
\left\lfloor
\begin{array}{l}
\varepsilon \in \left]0, \min\{1, \|L\|^{-2}\}\right[\\
v_0 \in \mathcal{G}
\end{array}
\right.\\
\text{For } n = 0, 1, \ldots\\
\left\lfloor
\begin{array}{l}
x_n = \operatorname{prox}_f(z - L^* v_n) + b_n\\
\gamma_n \in [\varepsilon, 2\|L\|^{-2} - \varepsilon]\\
\lambda_n \in [\varepsilon, 1]\\
v_{n+1} = v_n + \lambda_n\big(\operatorname{prox}_{\gamma_n g^*}(v_n + \gamma_n(Lx_n - r)) + a_n - v_n\big).
\end{array}
\right.
\end{array}
\tag{2.77}
$$

It is noteworthy that each iteration of Algorithm 2.36 achieves full splitting with respect to the operators L, prox_f, and $\operatorname{prox}_{g^*}$, which are used at separate steps. In addition, (2.77) incorporates tolerances a_n and b_n in the computation of the proximity operators at iteration n.

2.2.3.3 Convergence

Our main convergence result will be a consequence of Proposition 2.35 and the following results on the convergence of the forward-backward splitting method.

Theorem 2.37 [39, Theorem 3.4] *Let f_1 and f_2 be functions in $\Gamma_0(\mathcal{G})$ such that the set G of minimizers of $f_1 + f_2$ is nonempty and such that f_2 is differentiable on $\mathcal{G}$ with a $1/\beta$-Lipschitz continuous gradient for some $\beta \in \left]0, +\infty\right[$. Let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $\left]0, 2\beta\right[$ such that $\inf_{n\in\mathbb{N}}\gamma_n > 0$ and $\sup_{n\in\mathbb{N}}\gamma_n < 2\beta$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $\left]0, 1\right]$ such that $\inf_{n\in\mathbb{N}}\lambda_n > 0$, and let $(a_{1,n})_{n\in\mathbb{N}}$ and $(a_{2,n})_{n\in\mathbb{N}}$ be sequences in $\mathcal{G}$ such that $\sum_{n\in\mathbb{N}}\|a_{1,n}\| < +\infty$ and $\sum_{n\in\mathbb{N}}\|a_{2,n}\| < +\infty$. Fix $v_0 \in \mathcal{G}$ and, for every $n \in \mathbb{N}$, set*

$$
v_{n+1} = v_n + \lambda_n\Big(\operatorname{prox}_{\gamma_n f_1}\big(v_n - \gamma_n(\nabla f_2(v_n) + a_{2,n})\big) + a_{1,n} - v_n\Big). \tag{2.78}
$$

Then $(v_n)_{n\in\mathbb{N}}$ converges weakly to a point $v \in G$ and $\sum_{n\in\mathbb{N}}\|\nabla f_2(v_n) - \nabla f_2(v)\|^2 < +\infty$.

The following theorem describes the asymptotic behavior of Algorithm 2.36.

Theorem 2.38 *Let $(x_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}}$ be sequences generated by Algorithm 2.36, and let x be the solution to Problem 2.11. Then the following hold.*

(i) *$(v_n)_{n\in\mathbb{N}}$ converges weakly to a solution v to Problem 2.33 and $x = \operatorname{prox}_f(z - L^*v)$.*

(ii) *$(x_n)_{n\in\mathbb{N}}$ converges strongly to x.*

Proof. Let us define two functions f_1 and f_2 on $\mathcal{G}$ by $f_1 \colon v \mapsto g^*(v) + \langle v \mid r\rangle$ and $f_2 \colon v \mapsto \widetilde{f}^*(z - L^*v)$. Then (2.69) amounts to minimizing $f_1 + f_2$ on $\mathcal{G}$. Let us first check that all the assumptions specified in Theorem 2.37 are satisfied. First, f_1 and f_2 are in $\Gamma_0(\mathcal{G})$ and, by Proposition 2.34, $\operatorname{Argmin} f_1 + f_2 \neq \varnothing$. Moreover, it follows from Lemma 2.14(iv) that f_2 is differentiable on $\mathcal{G}$ with gradient

$$\nabla f_2 \colon v \mapsto -L\big(\operatorname{prox}_f(z - L^*v)\big). \tag{2.79}$$

Hence, we derive from Lemma 2.14(ii) that

$$\begin{aligned}(\forall v \in \mathcal{G})(\forall w \in \mathcal{G}) \quad \|\nabla f_2(v) - \nabla f_2(w)\| &\leqslant \|L\|\, \|\operatorname{prox}_f(z - L^*v) - \operatorname{prox}_f(z - L^*w)\| \\ &\leqslant \|L\|\, \|L^*v - L^*w\| \\ &\leqslant \|L\|^2\, \|v - w\|. \end{aligned} \tag{2.80}$$

The reciprocal of the Lipschitz constant of ∇f_2 is therefore $\beta = \|L\|^{-2}$. Now set

$$(\forall n \in \mathbb{N}) \quad a_{1,n} = a_n \quad \text{and} \quad a_{2,n} = -Lb_n. \tag{2.81}$$

Then $\sum_{n\in\mathbb{N}} \|a_{1,n}\| = \sum_{n\in\mathbb{N}} \|a_n\| < +\infty$ and $\sum_{n\in\mathbb{N}} \|a_{2,n}\| \leqslant \|L\| \sum_{n\in\mathbb{N}} \|b_n\| < +\infty$. Moreover, for every $n \in \mathbb{N}$, (2.77) yields

$$x_n = \operatorname{prox}_f(z - L^*v_n) + b_n \tag{2.82}$$

and, together with [39, Lemma 2.6(i)],

$$\begin{aligned} v_{n+1} &= v_n + \lambda_n\Big(\operatorname{prox}_{\gamma_n g^*}\big(v_n + \gamma_n(Lx_n - r)\big) + a_n - v_n\Big) \\ &= v_n + \lambda_n\Big(\operatorname{prox}_{\gamma_n g^* + \langle \cdot \mid \gamma_n r\rangle}\big(v_n + \gamma_n Lx_n\big) + a_n - v_n\Big) \\ &= v_n + \lambda_n\Big(\operatorname{prox}_{\gamma_n (g^* + \langle \cdot \mid r\rangle)}\big(v_n + \gamma_n L(\operatorname{prox}_f(z - L^*v_n) + b_n)\big) + a_n - v_n\Big) \\ &= v_n + \lambda_n\Big(\operatorname{prox}_{\gamma_n f_1}\big(v_n - \gamma_n(\nabla f_2(v_n) + a_{2,n})\big) + a_{1,n} - v_n\Big). \end{aligned} \tag{2.83}$$

This provides precisely the update rule (2.78), which allows us to apply Theorem 2.37.

(i) : In view of the above, we derive from Theorem 2.37 that $(v_n)_{n\in\mathbb{N}}$ converges weakly to a solution v to (2.69). The second assertion follows from Proposition 2.35.

(ii) : Let us set

$$(\forall n \in \mathbb{N}) \quad y_n = x_n - b_n = \operatorname{prox}_f(z - L^*v_n). \tag{2.84}$$

As seen in (i), $v_n \rightharpoonup v$, where v is a solution to (2.69), and $x = \operatorname{prox}_f(z - L^*v)$. Now set $\rho = \sup_{n\in\mathbb{N}} \|v_n - v\|$. Then $\rho < +\infty$ and, using Lemma 2.14(i) and (2.79), we obtain

$$\begin{aligned}
\|y_n - x\|^2 &= \|\operatorname{prox}_f(z - L^*v_n) - \operatorname{prox}_f(z - L^*v)\|^2 \\
&\leqslant \big\langle L^*v - L^*v_n \mid \operatorname{prox}_f(z - L^*v_n) - \operatorname{prox}_f(z - L^*v)\big\rangle \\
&= \big\langle v_n - v \mid -L\big(\operatorname{prox}_f(z - L^*v_n)\big) + L\big(\operatorname{prox}_f(z - L^*v)\big)\big\rangle \\
&= \langle v_n - v \mid \nabla f_2(v_n) - \nabla f_2(v)\rangle \\
&\leqslant \rho\|\nabla f_2(v_n) - \nabla f_2(v)\|.
\end{aligned} \tag{2.85}$$

However, as seen in Theorem 2.37, $\|\nabla f_2(v_n) - \nabla f_2(v)\| \to 0$. Hence, we derive from (2.85) that $y_n \to x$. In turn, since $b_n \to 0$, (2.84) yields $x_n \to x$. □

Remark 2.39 (Dykstra-like algorithm) Suppose that, in Problem 2.11, $\mathcal{G} = \mathcal{H}$, $L = \operatorname{Id}$, and $r = 0$. Then it follows from Theorem 2.38(ii) that the sequence $(x_n)_{n\in\mathbb{N}}$ produced by Algorithm 2.36 converges strongly to $x = \operatorname{prox}_{f+g} z$. Now let us consider the special case when Algorithm 2.36 is implemented with $v_0 = 0$, $\gamma_n \equiv 1$, $\lambda_n \equiv 1$, and no errors, i.e., $a_n \equiv 0$ and $b_n \equiv 0$. Then it follows from Lemma 2.14(iv) that (2.77) simplifies to

$$\begin{array}{l}
\text{Initialization} \\
\left\lfloor \; v_0 = 0 \right. \\
\text{For } n = 0, 1, \ldots \\
\left\lfloor \begin{array}{l} x_n = \operatorname{prox}_f(z - v_n) \\ v_{n+1} = x_n + v_n - \operatorname{prox}_g(x_n + v_n). \end{array} \right.
\end{array} \tag{2.86}$$

Using [5, Eq. (2.10)] it can then easily be shown by induction that the resulting sequence $(x_n)_{n\in\mathbb{N}}$ coincides with that produced by the Dykstra-like algorithm (2.37) (with $h_1 = g$ and $h_2 = f$) and that the sequence $(v_n)_{n\in\mathbb{N}}$ coincides with the sequence $(p_n)_{n\in\mathbb{N}}$ of (2.37). The fact that $x_n \to \operatorname{prox}_{f+g} z$ was established in [5, Theorem 3.3(i)] using different tools. Thus, Algorithm 2.36 can be regarded as a generalization of the Dykstra-like algorithm (2.37).

Remark 2.40 Theorem 2.38 remains valid if we introduce explicitly errors in the implementation of the operators L and L^* in Algorithm 2.36. More precisely, we can replace the steps defining x_n and v_n in (2.77) by

$$\left\lfloor \begin{array}{l} x_n = \operatorname{prox}_f(z - L^*v_n - d_{2,n}) + d_{1,n} \\ v_{n+1} = v_n + \lambda_n\big(\operatorname{prox}_{\gamma_n g^*}(v_n + \gamma_n(Lx_n + c_{2,n} - r)) + c_{1,n} - v_n\big), \end{array} \right. \tag{2.87}$$

where $(d_{1,n})_{n\in\mathbb{N}}$ and $(d_{2,n})_{n\in\mathbb{N}}$ are sequences in $\mathcal{H}$ such that $\sum_{n\in\mathbb{N}} \|d_{1,n}\| < +\infty$ and $\sum_{n\in\mathbb{N}} \|d_{2,n}\| < +\infty$, and where $(c_{1,n})_{n\in\mathbb{N}}$ and $(c_{2,n})_{n\in\mathbb{N}}$ are sequences in $\mathcal{G}$ such that $\sum_{n\in\mathbb{N}} \|c_{1,n}\| < +\infty$ and $\sum_{n\in\mathbb{N}} \|c_{2,n}\| < +\infty$. Indeed set, for every $n \in \mathbb{N}$,

$$\begin{cases} a_n = c_{1,n} + \operatorname{prox}_{\gamma_n g^*}(v_n + \gamma_n(Lx_n + c_{2,n} - r)) - \operatorname{prox}_{\gamma_n g^*}(v_n + \gamma_n(Lx_n - r)) \\ b_n = d_{1,n} + \operatorname{prox}_f(z - L^*v_n - d_{2,n}) - \operatorname{prox}_f(z - L^*v_n). \end{cases} \tag{2.88}$$

Then (2.87) reverts to

$$\left\lfloor \begin{array}{l} x_n = \operatorname{prox}_f(z - L^* v_n) + b_n \\ v_{n+1} = v_n + \lambda_n\big(\operatorname{prox}_{\gamma_n g^*}(v_n + \gamma_n(Lx_n - r)) + a_n - v_n\big), \end{array} \right. \tag{2.89}$$

as in (2.77). Moreover, by Lemma 2.14(ii),

$$\begin{aligned} (\forall n \in \mathbb{N}) \quad \|a_n\| &\leqslant \|c_{1,n}\| + \|\operatorname{prox}_{\gamma_n g^*}(v_n + \gamma_n(Lx_n + c_{2,n} - r)) \\ &\qquad\qquad\qquad\qquad - \operatorname{prox}_{\gamma_n g^*}(v_n + \gamma_n(Lx_n - r))\| \\ &\leqslant \|c_{1,n}\| + \gamma_n \|c_{2,n}\| \\ &\leqslant \|c_{1,n}\| + 2\|L\|^{-2}\|c_{2,n}\|. \end{aligned} \tag{2.90}$$

Thus, $\sum_{n\in\mathbb{N}} \|a_n\| < +\infty$. Likewise, we have $\sum_{n\in\mathbb{N}} \|b_n\| < +\infty$.

2.2.4 Application to specific signal recovery problems

In this section, we present a few applications of the duality framework presented in Section 2.2.3, which correspond to specific choices of $\mathcal{H}$, $\mathcal{G}$, L, f, g, r, and z in Problem 2.11.

2.2.4.1 Best feasible approximation

A standard feasibility problem in signal recovery is to find a signal in the intersection of two closed convex sets modeling constraints on the ideal solution [32, 73, 76, 82]. A more structured variant of this problem, is the so-called split feasibility problem [18, 21, 22], which requires to find a signal in a closed convex set $C \subset \mathcal{H}$ and such that some affine transformation of it lies in a closed convex set $D \subset \mathcal{G}$. Such problems typically admit infinitely many solutions and one often seeks to find the solution that lies closest to a nominal signal $z \in \mathcal{H}$ [30, 68]. This leads to the formulation (2.29), which consists in finding the best approximation to a reference signal $z \in \mathcal{H}$ from the feasibility set $C \cap L^{-1}(r + D)$.

Problem 2.41 Let $z \in \mathcal{H}$, let $r \in \mathcal{G}$, let $C \subset \mathcal{H}$ and $D \subset \mathcal{G}$ be closed convex sets, and let L be a nonzero operator in $\mathcal{B}(\mathcal{H}, \mathcal{G})$ such that

$$r \in \operatorname{sri}\big(L(C) - D\big). \tag{2.91}$$

The problem is to

$$\underset{\substack{x \in C \\ Lx - r \in D}}{\text{minimize}} \; \frac{1}{2}\|x - z\|^2, \tag{2.92}$$

and its dual is to

$$\underset{v\in\mathcal{G}}{\text{minimize}}\ \frac{1}{2}\|z-L^*v\|^2-\frac{1}{2}d_C^2(z-L^*v)+\sigma_D(v)+\langle v\mid r\rangle. \tag{2.93}$$

Proposition 2.42 *Let $(b_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that $\sum_{n\in\mathbb{N}}\|b_n\|<+\infty$, let $(c_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{G}$ such that $\sum_{n\in\mathbb{N}}\|c_n\|<+\infty$, and let $(x_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}}$ be sequences generated by the following routine.*

$$\begin{array}{l}
\text{Initialization}\\
\left\lfloor\begin{array}{l}
\varepsilon\in\left]0,\min\{1,\|L\|^{-2}\}\right[\\
v_0\in\mathcal{G}
\end{array}\right.\\
\text{For } n=0,1,\ldots\\
\left\lfloor\begin{array}{l}
x_n=P_C(z-L^*v_n)+b_n\\
\gamma_n\in[\varepsilon,2\|L\|^{-2}-\varepsilon]\\
\lambda_n\in[\varepsilon,1]\\
v_{n+1}=v_n+\lambda_n\gamma_n\big(Lx_n-r-P_D(\gamma_n^{-1}v_n+Lx_n-r)+c_n\big).
\end{array}\right.
\end{array} \tag{2.94}$$

Then the following hold, where x designates the primal solution to Problem 2.41.

(i) $(v_n)_{n\in\mathbb{N}}$ *converges weakly to a solution v to (2.93) and $x=P_C(z-L^*v)$.*

(ii) $(x_n)_{n\in\mathbb{N}}$ *converges strongly to x.*

Proof. Set $f=\iota_C$ and $g=\iota_D$. Then (2.28) reduces to (2.92) and (2.27) reduces to (2.91). In addition, we derive from Lemma 2.14(iii) that $\widetilde{f^*}=\|\cdot\|^2/2-\widetilde{\iota_C}=(\|\cdot\|^2-d_C^2)/2$. Hence, in view of (2.69), (2.93) in indeed the dual of (2.92). Furthermore, items (i) and (ii) in Example 2.17 yield $\text{prox}_f=P_C$ and

$$(\forall n\in\mathbb{N})\quad \text{prox}_{\gamma_n g^*}=\text{prox}_{\gamma_n\sigma_D}=\text{prox}_{\sigma_{\gamma_n D}}=\text{Id}-P_{\gamma_n D}=\text{Id}-\gamma_n P_D(\cdot/\gamma_n). \tag{2.95}$$

Finally, set $(\forall n\in\mathbb{N})$ $a_n=\gamma_n c_n$. Then $\sum_{n\in\mathbb{N}}\|a_n\|\leqslant 2\|L\|^{-2}\sum_{n\in\mathbb{N}}\|c_n\|<+\infty$ and, altogether, (2.77) reduces to (2.94). Hence, the results follow from Theorem 2.38. □

Our investigation was motivated in the Introduction by the duality framework of [68]. In the next example we recover and sharpen Proposition 2.10.

Example 2.43 Consider the special case of Problem 2.41 in which $z=0$, $\mathcal{G}=\mathbb{R}^N$, $D=\{0\}$, $r=(\rho_i)_{1\leqslant i\leqslant N}$, and $L\colon x\mapsto(\langle x\mid s_i\rangle)_{1\leqslant i\leqslant N}$, where $(s_i)_{1\leqslant i\leqslant N}\in\mathcal{H}^N$ satisfies $\sum_{i=1}^N\|s_i\|^2\leqslant 1$. Then, by (2.44), (2.91) reduces to $r\in\text{ri}\,L(C)$ and (2.92) to (2.24). Since $\|L\|\leqslant 1$, specializing (2.94) to the case when $c_n\equiv 0$ and $\lambda_n\equiv 1$, and introducing the sequence $(w_n)_{n\in\mathbb{N}} = (-v_n)_{n\in\mathbb{N}}$ for convenience yields the following routine.

$$
\begin{array}{l}
\text{Initialization} \\
\left\lfloor
\begin{array}{l}
\varepsilon \in \left]0,1\right[\\
w_0 \in \mathbb{R}^N
\end{array}
\right. \\
\text{For } n = 0,1,\ldots \\
\left\lfloor
\begin{array}{l}
x_n = P_C(L^* w_n) + b_n \\
\gamma_n \in [\varepsilon, 2\|L\|^{-2} - \varepsilon] \\
w_{n+1} = w_n + \gamma_n\big(r - Lx_n\big).
\end{array}
\right.
\end{array}
\tag{2.96}
$$

Thus, if $\sum_{n\in\mathbb{N}} \|b_n\| < +\infty$, we deduce from Proposition 2.42(i) and Proposition 2.34 the weak convergence of $(w_n)_{n\in\mathbb{N}}$ to a point w such that $v = -w$ satisfies (2.70), i.e., $L(P_C(-L^*v)) - r \in \partial\iota^*_{\{0\}}(v) = \{0\}$ or, equivalently, $L(P_C(L^*w)) = r$, and such that $P_C(-L^*v) = P_C(L^*w)$ is the solution to (2.24). In addition, we derive from Proposition 2.42(ii), the strong convergence of $(x_n)_{n\in\mathbb{N}}$ to the solution to (2.24). These results sharpen the conclusion of Proposition 2.10 (note that (2.25) corresponds to setting $b_n \equiv 0$ and $\gamma_n \equiv \gamma \in \left]0,2\right[$ in (2.96)).

Example 2.44 We consider the standard linear inverse problem of recovering an ideal signal $\overline{x} \in \mathcal{H}$ from an observation

$$
r = L\overline{x} + s \tag{2.97}
$$

in $\mathcal{G}$, where $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ and where $s \in \mathcal{G}$ models noise. Given an estimate x of $\overline{x}$, the residual $r - Lx$ should ideally behave like the noise process. Thus, any known probabilistic attribute of the noise process can give rise to a constraint. This observation was used in [38, 76] to construct various constraints of the type $Lx - r \in D$, where D is closed and convex. In this context, (2.92) amounts to finding the signal which is closest to some nominal signal z and which satisfies a noise-based constraint and some convex constraint on $\overline{x}$ represented by C. Such problems were considered for instance in [30], where they were solved by methods that require the projection onto the set $\{x \in \mathcal{H} \mid Lx - r \in D\}$, which is typically hard to compute, even in the simple case when D is a closed Euclidean ball [76]. By contrast, the iterative method (2.94) requires only the projection onto D to enforce such constraints.

2.2.4.2 Soft best feasible approximation

It follows from (2.91) that the underlying feasibility set $C \cap L^{-1}(r + D)$ in Problem 2.41 is nonempty. In many situations, feasibility may not guaranteed due to, for instance, imprecise prior information or unmodeled dynamics in the data formation process [31, 81]. In such instances, one can relax the hard constraints $x \in C$ and $Lx - r \in D$ in (2.92) by merely forcing that x be close to C and $Lx - r$ be close to D. Let us formulate this problem within the framework of Problem 2.11.

Problem 2.45 Let $z \in \mathcal{H}$, let $r \in \mathcal{G}$, let $C \subset \mathcal{H}$ and $D \subset \mathcal{G}$ be nonempty closed convex sets, let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ be a nonzero operator, and let ϕ and ψ be even functions in $\Gamma_0(\mathbb{R}) \smallsetminus \{\iota_{\{0\}}\}$ such that

$$r \in \operatorname{sri}\big(L(\{x \in \mathcal{H} \mid d_C(x) \in \operatorname{dom}\phi\}) - \{y \in \mathcal{G} \mid d_D(y) \in \operatorname{dom}\psi\}\big). \tag{2.98}$$

The problem is to

$$\underset{x \in \mathcal{H}}{\text{minimize}} \;\; \phi\big(d_C(x)\big) + \psi\big(d_D(Lx - r)\big) + \frac{1}{2}\|x - z\|^2, \tag{2.99}$$

and its dual is to

$$\underset{v \in \mathcal{G}}{\text{minimize}} \;\; \frac{1}{2}\|z - L^*v\|^2 - (\phi \circ d_C)^{\sim}(z - L^*v) + \sigma_D(v) + \psi^*(\|v\|) + \langle v \mid r\rangle. \tag{2.100}$$

Since ϕ and ψ are even functions in $\Gamma_0(\mathbb{R}) \smallsetminus \{\iota_{\{0\}}\}$, we can use Example 2.21 to get an explicitly expression of the proximity operators involved and solve the minimization problems (2.99) and (2.100) as follows.

Proposition 2.46 *Let $(b_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that $\sum_{n\in\mathbb{N}} \|b_n\| < +\infty$, let $(c_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{G}$ such that $\sum_{n\in\mathbb{N}} \|c_n\| < +\infty$, and let $(x_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}}$ be sequences generated by the following routine.*

$$
\begin{array}{l}
\text{Initialization}\\
\left\lfloor
\begin{array}{l}
\varepsilon \in \left]0, \min\{1, \|L\|^{-2}\}\right[\\
v_0 \in \mathcal{G}
\end{array}
\right.\\
\text{For } n = 0, 1, \ldots\\
\left\lfloor
\begin{array}{l}
y_n = z - L^*v_n\\
\text{if } \; d_C(y_n) > \max \partial\phi(0)\\
\left\lfloor \; x_n = y_n + \dfrac{\operatorname{prox}_{\phi^*} d_C(y_n)}{d_C(y_n)}(P_C y_n - y_n) + b_n \right.\\
\text{if } \; d_C(y_n) \leqslant \max \partial\phi(0)\\
\left\lfloor \; x_n = P_C y_n + b_n \right.\\
\gamma_n \in [\varepsilon, 2\|L\|^{-2} - \varepsilon]\\
w_n = \gamma_n^{-1} v_n + Lx_n - r\\
\text{if } \; d_D(w_n) > \gamma_n^{-1} \max \partial\psi(0)\\
\left\lfloor \; p_n = \dfrac{\operatorname{prox}_{(\gamma_n^{-1}\psi)^*} d_D(w_n)}{d_D(w_n)}(w_n - P_D w_n) + c_n \right.\\
\text{if } \; d_D(w_n) \leqslant \gamma_n^{-1} \max \partial\psi(0)\\
\left\lfloor \; p_n = w_n - P_D w_n + c_n \right.\\
\lambda_n \in [\varepsilon, 1]\\
v_{n+1} = v_n + \lambda_n\big(\gamma_n p_n - v_n\big).
\end{array}
\right.
\end{array}
\tag{2.101}
$$

Then the following hold, where x designates the primal solution to Problem 2.45.

(i) $(v_n)_{n\in\mathbb{N}}$ *converges weakly to a solution* v *to* (2.100) *and, if we set* $y = z - L^*v$,

$$x = \begin{cases} y + \dfrac{\mathrm{prox}_{\phi^*} d_C(y)}{d_C(y)}(P_C y - y), & \text{if } d_C(y) > \max \partial\phi(0); \\ P_C y, & \text{if } d_C(y) \leqslant \max \partial\phi(0). \end{cases} \tag{2.102}$$

(ii) $(x_n)_{n\in\mathbb{N}}$ *converges strongly to* x.

Proof. Set $f = \phi\circ d_C$ and $g = \psi\circ d_D$. Since d_C and d_D are continuous convex functions, $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$. Moreover, (2.98) implies that (2.27) holds. Thus, Problem 2.45 is a special case of Problem 2.11. On the other hand, it follows from Lemma 2.14(iii) that $\widetilde{f^*} = \|\cdot\|^2/2 - (\phi\circ d_C)^\sim$ and from [17, Lemma 2.2] that $g^* = \sigma_D + \psi^*\circ\|\cdot\|$. This shows that (2.100) is the dual of (2.99). Let us now examine iteration n of the algorithm. In view of Example 2.21, the vector x_n in (2.101) is precisely the vector $x_n = \mathrm{prox}_f(z - L^*v_n) + b_n$ of (2.77). Moreover, using successively the definition of w_n in (2.101), Lemma 2.15, Example 2.21, and the definition of p_n in (2.101), we obtain

$$\begin{aligned} &\gamma_n^{-1}\mathrm{prox}_{\gamma_n g^*}(v_n + \gamma_n(Lx_n - r)) \\ &\quad= \gamma_n^{-1}\mathrm{prox}_{\gamma_n g^*}(\gamma_n w_n) \\ &\quad= w_n - \mathrm{prox}_{\gamma_n^{-1} g} w_n \\ &\quad= w_n - \mathrm{prox}_{(\gamma_n^{-1}\psi)\circ d_D} w_n \\ &\quad= \begin{cases} \dfrac{\mathrm{prox}_{(\gamma_n^{-1}\psi)^*} d_D(w_n)}{d_D(w_n)}(w_n - P_D w_n) & \text{if } d_D(w_n) > \gamma_n^{-1}\max\partial\psi(0) \\ w_n - P_D w_n & \text{if } d_D(w_n) \leqslant \gamma_n^{-1}\max\partial\psi(0) \end{cases} \\ &\quad= p_n - c_n. \end{aligned} \tag{2.103}$$

Altogether, (2.101) is a special instance of (2.77) in which $(\forall n \in \mathbb{N})\ a_n = \gamma_n c_n$. Therefore, since $\sum_{n\in\mathbb{N}} \|a_n\| \leqslant 2\|L\|^{-2} \sum_{n\in\mathbb{N}} \|c_n\| < +\infty$, the assertions follow from Theorem 2.38, where we have used (2.57) to get (2.102). □

Example 2.47 We can obtain a soft-constrained version of the Potter-Arun problem (2.24) revisited in Example 2.43 by specializing Problem 2.45 as follows : $z = 0$, $\mathcal{G} = \mathbb{R}^N$, $D = \{0\}$, $r = (\rho_i)_{1\leqslant i\leqslant N}$, and $L\colon x \mapsto (\langle x \mid s_i\rangle)_{1\leqslant i\leqslant N}$, where $(s_i)_{1\leqslant i\leqslant N} \in \mathcal{H}^N$ satisfies $\sum_{i=1}^N \|s_i\|^2 \leqslant 1$. We thus arrive at the relaxed version of (2.24)

$$\underset{x\in\mathcal{H}}{\text{minimize}}\ \ \phi(d_C(x)) + \psi\Big(\sqrt{\textstyle\sum_{i=1}^N |\langle x \mid s_i\rangle - \rho_i|^2}\Big) + \frac{1}{2}\|x\|^2. \tag{2.104}$$

Since $D = \{0\}$, we can replace each occurrence of $d_D(w_n)$ by $\|w_n\|$ and each occurrence of $w_n - P_D w_n$ by w_n in (2.101). Proposition 2.46(ii) asserts that any sequence $(x_n)_{n\in\mathbb{N}}$ produced by the resulting algorithm converges strongly to the solution to (2.104). For the sake of illustration, let us consider the case when $\phi = \alpha|\cdot|^{4/3}$ and $\psi = \beta|\cdot|$, for some

α and β in $]0,+\infty[$. Then $\operatorname{dom}\psi=\mathbb{R}$ and (2.98) is trivially satisfied. In addition, (2.104) becomes

$$\underset{x\in\mathcal{H}}{\text{minimize}}\;\; \alpha d_C^{4/3}(x)+\beta\sqrt{\textstyle\sum_{i=1}^N|\langle x\mid s_i\rangle-\rho_i|^2}+\frac{1}{2}\|x\|^2. \tag{2.105}$$

Since $\phi^*\colon\mu\mapsto 27|\mu|^4/(256\alpha^3)$, $\operatorname{prox}_{\phi^*}$ in (2.101) can be derived from Example 2.26(vi). On the other hand, since $\psi^*=\iota_{[-\beta,\beta]}$, Example 2.17(i) yields $\operatorname{prox}_{\psi^*}=P_{[-\beta,\beta]}$. Thus, upon setting, for simplicity, $b_n\equiv 0$, $c_n\equiv 0$, $\lambda_n\equiv 1$, and $\gamma_n\equiv 1$ (note that $\|L\|\leqslant 1$) in (2.101) and observing that $\partial\phi(0)=\{0\}$ and $\partial\psi(0)=[-\beta,\beta]$, we obtain the following algorithm, where $L^*\colon(\nu_i)_{1\leqslant i\leqslant N}\mapsto\sum_{i=1}^N\nu_i s_i$.

Initialization
$$\left\lfloor\begin{array}{l}\tau=3/(2\alpha 4^{1/3}),\;\;\sigma=256\alpha^3/729\\ v_0\in\mathbb{R}^N\end{array}\right.$$
For $n=0,1,\ldots$
$$\left\lfloor\begin{array}{l}y_n=z-L^*v_n\\ \text{if}\;\; y_n\notin C\\ \left\lfloor\; x_n=y_n+\dfrac{\left|\sqrt{d_C^2(y_n)+\sigma}+d_C(y_n)\right|^{1/3}-\left|\sqrt{d_C^2(y_n)+\sigma}-d_C(y_n)\right|^{1/3}}{\tau d_C(y_n)}(P_Cy_n-y_n)\right.\\ \text{if}\;\; y_n\in C\\ \lfloor\; x_n=y_n\\ w_n=v_n+Lx_n-r\\ \text{if}\;\;\|w_n\|>\beta\\ \left\lfloor\; v_{n+1}=\dfrac{\beta}{\|w_n\|}w_n\right.\\ \text{if}\;\;\|w_n\|\leqslant\beta\\ \lfloor\; v_{n+1}=w_n.\end{array}\right.$$

As shown above, the sequence $(x_n)_{n\in\mathbb{N}}$ converges strongly to the solution to (2.105).

Remark 2.48 Alternative relaxations of (2.24) can be derived from Problem 2.11. For instance, given an even function $\phi\in\Gamma_0(\mathbb{R})\smallsetminus\{\iota_{\{0\}}\}$ and $\alpha\in\,]0,+\infty[$, an alternative to (2.104) is

$$\underset{x\in\mathcal{H}}{\text{minimize}}\;\; \phi(d_C(x))+\alpha\max_{1\leqslant i\leqslant N}|\langle x\mid s_i\rangle-\rho_i|+\frac{1}{2}\|x\|^2. \tag{2.106}$$

This formulation results from (2.28) with $z=0$, $f=\phi\circ d_C$, $\mathcal{G}=\mathbb{R}^N$, $r=(\rho_i)_{1\leqslant i\leqslant N}$, $L\colon x\mapsto(\langle x\mid s_i\rangle)_{1\leqslant i\leqslant N}$, and $g=\alpha\|\cdot\|_\infty$ (note that (2.27) holds since $\operatorname{dom}g=\mathcal{G}$). Since

$g^* = \iota_D$, where $D = \big\{(\nu_i)_{1\leqslant i\leqslant N} \in \mathbb{R}^N \;\big|\; \sum_{i=1}^N |\nu_i| \leqslant \alpha\big\}$, the dual problem (2.69) therefore assumes the form

$$\underset{(\nu_i)_{1\leqslant i\leqslant N}\in D}{\text{minimize}} \;\; \frac{1}{2}\Big\|\sum_{i=1}^N \nu_i s_i\Big\|^2 - (\phi\circ d_C)^{\sim}\Big(-\sum_{i=1}^N \nu_i s_i\Big) + \sum_{i=1}^N \rho_i\nu_i. \tag{2.107}$$

The proximity operators of $f = \phi\circ d_C$ and $\gamma_n g^* = \iota_D$ required by Algorithm 2.36 are supplied by Example 2.21 and Example 2.17(i), respectively. Strong convergence of the resulting sequence $(x_n)_{n\in\mathbb{N}}$ to the solution to (2.106) is guaranteed by Theorem 2.38(ii).

2.2.4.3 Denoising over dictionaries

In denoising problems, the goal is to recover the original form of an ideal signal $\overline{x} \in \mathcal{H}$ from a corrupted observation

$$z = \overline{x} + s, \tag{2.108}$$

where $s \in \mathcal{H}$ is the realization of a noise process which may for instance model imperfections in the data recording instruments, uncontrolled dynamics, or physical interferences. A common approach to solve this problem is to minimize the least-squares data fitting functional $x \mapsto \|x - z\|^2/2$ subject to some constraints on x that represent a priori knowledge on the ideal solution $\overline{x}$ and some affine transformation $L\overline{x} - r$ thereof, where $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ and $r \in \mathcal{G}$. By measuring the degree of violation of these constraints via potentials $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$, we arrive at (2.28). In this context, L can be a gradient [23, 45, 54, 72], a low-pass filter [2, 77], a wavelet or a frame decomposition operator [36, 44, 78]. Alternatively, the vector $r \in \mathcal{G}$ may arise from the availability of a second observation in the form of a noise-corrupted linear measurement of $\overline{x}$, as in (2.97) [27].

In this section, the focus is placed on models in which information on the scalar products $(\langle \overline{x} \mid e_k\rangle)_{k\in\mathbb{K}}$ of the original signal $\overline{x}$ against a finite or infinite a sequence of reference unit norm vectors $(e_k)_{k\in\mathbb{K}}$ of $\mathcal{H}$, called a dictionary, is available. In practice, such information can take various forms, e.g., sparsity, distribution type, statistical properties [27, 35, 41, 49, 57, 75], and they can often be modeled in a variational framework by introducing a sequence of convex potentials $(\phi_k)_{k\in\mathbb{K}}$. If we model the rest of the information available about $\overline{x}$ via a potential f, we obtain the following formulation.

Problem 2.49 Let $z \in \mathcal{H}$, let $f \in \Gamma_0(\mathcal{H})$, let $(e_k)_{k\in\mathbb{K}}$ be a sequence of unit norm vectors in $\mathcal{H}$ such that

$$(\exists\, \delta \in \left]0,+\infty\right[)(\forall x \in \mathcal{H}) \quad \sum_{k\in\mathbb{K}} |\langle x \mid e_k\rangle|^2 \leqslant \delta\|x\|^2, \tag{2.109}$$

and let $(\phi_k)_{k\in\mathbb{K}}$ be functions in $\Gamma_0(\mathbb{R})$ such that

$$(\forall k \in \mathbb{K}) \quad \phi_k \geqslant \phi_k(0) = 0 \tag{2.110}$$

and

$$0 \in \operatorname{sri}\left\{ \big(\langle x \mid e_k\rangle - \xi_k\big)_{k\in\mathbb{K}} \;\middle|\; (\xi_k)_{k\in\mathbb{K}} \in \ell^2(\mathbb{K}),\;\; \sum_{k\in\mathbb{K}} \phi_k(\xi_k) < +\infty,\;\; \text{and}\;\; x \in \operatorname{dom} f \right\}. \tag{2.111}$$

The problem is to

$$\underset{x\in\mathcal{H}}{\text{minimize}}\;\; f(x) + \sum_{k\in\mathbb{K}} \phi_k(\langle x \mid e_k\rangle) + \frac{1}{2}\|x-z\|^2, \tag{2.112}$$

and its dual is to

$$\underset{(\nu_k)_{k\in\mathbb{K}}\in\ell^2(\mathbb{K})}{\text{minimize}}\;\; \widetilde{f}^*\Big(z - \sum_{k\in\mathbb{K}} \nu_{n,k} e_k\Big) + \sum_{k\in\mathbb{K}} \phi_k^*(\nu_k). \tag{2.113}$$

Problems (2.112) and (2.113) can be solved by the following algorithm, where $\alpha_{n,k}$ stands for a numerical tolerance in the implementation of the operator $\operatorname{prox}_{\gamma_n\phi_k^*}$. Let us note that closed-form expressions for the proximity operators of a wide range of functions in $\Gamma_0(\mathbb{R})$ are available [27, 35, 39], in particular in connection with Bayesian formulations involving log-concave densities, and with problems involving sparse representations (see also Examples 2.26–2.29 and Lemmas 2.30–2.31).

Proposition 2.50 *Let $((\alpha_{n,k})_{n\in\mathbb{N}})_{k\in\mathbb{K}}$ be sequences in $\mathbb{R}$ such that $\sum_{n\in\mathbb{N}} \sqrt{\sum_{k\in\mathbb{K}} |\alpha_{n,k}|^2} < +\infty$, let $(b_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that $\sum_{n\in\mathbb{N}} \|b_n\| < +\infty$, and let $(x_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}} = ((\nu_{n,k})_{k\in\mathbb{K}})_{n\in\mathbb{N}}$ be sequences generated by the following routine.*

$$\begin{array}{l}
\text{Initialization} \\
\left\lfloor \begin{array}{l} \varepsilon \in \left]0, \min\{1, \delta^{-1}\}\right[\\ (\nu_{0,k})_{k\in\mathbb{K}} \in \ell^2(\mathbb{K}) \end{array} \right. \\
\text{For } n = 0, 1, \ldots \\
\left\lfloor \begin{array}{l} x_n = \operatorname{prox}_f\big(z - \sum_{k\in\mathbb{K}} \nu_{n,k} e_k\big) + b_n \\ \gamma_n \in [\varepsilon, 2\delta^{-1} - \varepsilon] \\ \lambda_n \in [\varepsilon, 1] \\ \text{For every } k \in \mathbb{K} \\ \left\lfloor \nu_{n+1,k} = \nu_{n,k} + \lambda_n\big(\operatorname{prox}_{\gamma_n\phi_k^*}(\nu_{n,k} + \gamma_n\langle x_n \mid e_k\rangle) + \alpha_{n,k} - \nu_{n,k}\big). \right. \end{array} \right.
\end{array} \tag{2.114}$$

Then the following hold, where x designates the primal solution to Problem 2.49.

(i) *$(v_n)_{n\in\mathbb{N}}$ converges weakly to a solution $(\nu_k)_{k\in\mathbb{K}}$ to (2.113) and $x = \operatorname{prox}_f(z - \sum_{k\in\mathbb{K}} \nu_k e_k)$.*

(ii) *$(x_n)_{n\in\mathbb{N}}$ converges strongly to x.*

Proof. Set $\mathcal{G} = \ell^2(\mathbb{K})$ and $r = 0$. Define

$$L\colon \mathcal{H} \to \mathcal{G}\colon x \mapsto (\langle x \mid e_k\rangle)_{k\in\mathbb{K}} \quad \text{and} \quad g\colon \mathcal{G} \to \left]-\infty,+\infty\right] : (\xi_k)_{k\in\mathbb{K}} \mapsto \sum_{k\in\mathbb{K}} \phi_k(\xi_k). \tag{2.115}$$

Then $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ and its adjoint is the operator $L^* \in \mathcal{B}(\mathcal{G},\mathcal{H})$ defined by

$$L^*\colon (\xi_k)_{k\in\mathbb{K}} \mapsto \sum_{k\in\mathbb{K}} \xi_k e_k. \tag{2.116}$$

On the other hand, it follows from our assumptions that $g \in \Gamma_0(\mathcal{G})$ (Example 2.20) and that

$$g^*\colon \mathcal{G} \to \left]-\infty,+\infty\right] : (\nu_k)_{k\in\mathbb{K}} \mapsto \sum_{k\in\mathbb{K}} \phi_k^*(\nu_k). \tag{2.117}$$

In addition, (2.111) implies that (2.27) holds. This shows that (2.112) is a special case of (2.28) and that (2.113) is a special case of (2.69). We also observe that (2.109) and (2.115) yield

$$\|L\|^2 = \sup_{\|x\|=1} \|Lx\|^2 = \sup_{\|x\|=1} \sum_{k\in\mathbb{K}} |\langle x \mid e_k\rangle|^2 \leqslant \delta. \tag{2.118}$$

Hence, $[\varepsilon, 2\delta^{-1}-\varepsilon] \subset [\varepsilon, 2\|L\|^{-2}-\varepsilon]$. Next, we derive from (2.45) and (2.110) that, for every $k \in \mathbb{K}$, $\phi_k^*(0) = \sup_{\xi\in\mathbb{R}} -\phi_k(\xi) = -\inf_{\xi\in\mathbb{R}} \phi_k(\xi) = \phi_k(0) = 0$ and that $(\forall \nu \in \mathbb{R})$ $\phi_k^*(\nu) = \sup_{\xi\in\mathbb{R}} \xi\nu - \phi_k(\xi) \geqslant -\phi_k(0) = 0$. In turn, we derive from (2.117) and Example 2.20 (applied to the canonical orthonormal basis of $\ell^2(\mathbb{K})$) that

$$(\forall \gamma \in \left]0,+\infty\right[)(\forall v = (\nu_k)_{k\in\mathbb{K}} \in \mathcal{G}) \quad \operatorname{prox}_{\gamma g^*} v = \big(\operatorname{prox}_{\gamma\phi_k^*} \nu_k\big)_{k\in\mathbb{K}}. \tag{2.119}$$

Altogether, (2.114) is a special case of Algorithm 2.36 with $(\forall n \in \mathbb{N})$ $a_n = (\alpha_{n,k})_{k\in\mathbb{K}}$. Hence, the assertions follow from Theorem 2.38. ◻

Remark 2.51 Using (2.115), we can write the potential on the dictionary coefficients in Problem 2.49 as

$$g \circ L\colon x \mapsto \sum_{k\in\mathbb{K}} \phi_k(\langle x \mid e_k\rangle). \tag{2.120}$$

(i) If $(e_k)_{k\in\mathbb{K}}$ were an orthonormal basis in Problem 2.49, we would have $L^{-1} = L^*$ and $\operatorname{prox}_{g\circ L}$ would be decomposable as $L^* \circ \operatorname{prox}_g \circ L$ [39, Lemma 2.8]. As seen in the Introduction, we could then approach (2.112) directly via forward-backward, Douglas-Rachford, or Dykstra-like splitting, depending on the properties of f. Our duality framework allows us to solve (2.112) for the much broader class of dictionaries satisfying (2.109) and, in particular, for frames [40].

(ii) Suppose that each ϕ_k in Problem 2.49 is of the form $\phi_k = \psi_k + \sigma_{\Omega_k}$, where $\psi_k \in \Gamma_0(\mathbb{R})$ satisfies $\psi_k \geqslant \psi_k(0) = 0$ and is differentiable at 0 with $\psi_k'(0) = 0$, and where Ω_k is a nonempty closed interval. In this case, (2.120) aims at promoting the sparsity of the solution in the dictionary $(e_k)_{k\in\mathbb{K}}$ [35] (a standard case is when, for every $k \in \mathbb{K}$, $\psi_k = 0$ and $\Omega_k = [-\omega_k, \omega_k]$, which gives rise to the standard weighted ℓ^1 potential $x \mapsto \sum_{k\in\mathbb{K}} \omega_k |\langle x \mid e_k\rangle|$). Moreover, the proximity operator $\text{prox}_{\gamma_n \phi_k^*}$ in (2.114) can be evaluated via Lemma 2.15 and Lemma 2.30.

2.2.4.4 Denoising with support functions

Suppose that g in Problem 2.11 is positively homogeneous, i.e.,

$$(\forall \lambda \in \,]0, +\infty[)(\forall y \in \mathcal{G}) \quad g(\lambda y) = \lambda g(y). \tag{2.121}$$

Instances of such functions arising in denoising problems can be found in [1, 8, 9, 24, 35, 39, 42, 65, 72, 79] and in the examples below. It follows from (2.121) and [4, Theorem 2.4.2] that g is the support function of a nonempty closed convex set $D \subset \mathcal{G}$, namely

$$g = \sigma_D = \sup_{v\in D} \langle \cdot \mid v\rangle, \quad \text{where} \quad D = \partial g(0) = \big\{ v \in \mathcal{G} \;\big|\; (\forall y \in \mathcal{G}) \;\; \langle y \mid v\rangle \leqslant g(y) \big\}. \tag{2.122}$$

If we denote by $\text{bar}\, D = \big\{ y \in \mathcal{G} \;\big|\; \sup_{v\in D} \langle y \mid v\rangle < +\infty \big\}$ the barrier cone of D, we thus obtain the following instance of Problem 2.11.

Problem 2.52 Let $z \in \mathcal{H}$, $r \in \mathcal{G}$, let $f \in \Gamma_0(\mathcal{H})$, let D be a nonempty closed convex subset of $\mathcal{G}$, and let L be a nonzero operator in $\mathcal{B}(\mathcal{H}, \mathcal{G})$ such that

$$r \in \text{sri}\, \big(L(\text{dom}\, f) - \text{bar}\, D \big). \tag{2.123}$$

The problem is to

$$\underset{x\in\mathcal{H}}{\text{minimize}} \;\; f(x) + \sigma_D(Lx - r) + \frac{1}{2}\|x - z\|^2, \tag{2.124}$$

and its dual is to

$$\underset{v\in D}{\text{minimize}} \;\; \widetilde{f^*}(z - L^* v) + \langle v \mid r\rangle. \tag{2.125}$$

Proposition 2.53 *Let $(a_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{G}$ such that $\sum_{n\in\mathbb{N}} \|a_n\| < +\infty$, let $(b_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that $\sum_{n\in\mathbb{N}} \|b_n\| < +\infty$, and let $(x_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}}$ be sequences* *generated by the following routine.*

$$
\begin{array}{l}
\text{Initialization}\\
\left\lfloor
\begin{array}{l}
\varepsilon \in \left]0, \min\{1, \|L\|^{-2}\}\right[\\
v_0 \in \mathcal{G}
\end{array}
\right.\\
\text{For } n = 0, 1, \ldots\\
\left\lfloor
\begin{array}{l}
x_n = \operatorname{prox}_f(z - L^* v_n) + b_n\\
\gamma_n \in [\varepsilon, 2\|L\|^{-2} - \varepsilon]\\
\lambda_n \in [\varepsilon, 1]\\
v_{n+1} = v_n + \lambda_n\big(P_D(v_n + \gamma_n(Lx_n - r)) + a_n - v_n\big).
\end{array}
\right.
\end{array}
\tag{2.126}
$$

Then the following hold, where x designates the primal solution to Problem 2.52.

(i) $(v_n)_{n\in\mathbb{N}}$ *converges weakly to a solution v to (2.125) and* $x = \operatorname{prox}_f(z - L^* v)$.

(ii) $(x_n)_{n\in\mathbb{N}}$ *converges strongly to x.*

Proof. The assertions follow from Theorem 2.38 with $g = \sigma_D$. Indeed, $g^* = \iota_D$ and, therefore, $(\forall \gamma \in \left]0, +\infty\right[)\ \operatorname{prox}_{\gamma g^*} = P_D$. □

Remark 2.54 Condition (2.123) is trivially satisfied when D is bounded, in which case $\operatorname{bar} D = \mathcal{G}$.

In the remainder of this section, we focus on examples that feature a bounded set D onto which projections are easily computed.

Example 2.55 In Problem 2.52, let D be the closed unit ball of $\mathcal{G}$. Then $P_D \colon y \mapsto y/\max\{\|y\|, 1\}$ and $\sigma_D = \|\cdot\|$. Hence, (2.124) becomes

$$
\underset{x\in\mathcal{H}}{\text{minimize}}\ \ f(x) + \|Lx - r\| + \frac{1}{2}\|x - z\|^2, \tag{2.127}
$$

and the dual problem (2.125) becomes

$$
\underset{v\in\mathcal{G},\,\|v\|\leqslant 1}{\text{minimize}}\ \ \widetilde{f}^*(z - L^* v) + \langle v \mid r\rangle. \tag{2.128}
$$

In signal recovery, variational formulations involving positively homogeneous functionals to control the behavior of the gradient of the solutions play a prominent role, e.g., [3, 14, 52, 65, 72]. In the context of image recovery, such a formulation can be obtained by revisiting Problem 2.52 with $\mathcal{H} = H_0^1(\Omega)$, where Ω is a bounded open domain in $\mathbb{R}^2$, $\mathcal{G} = L^2(\Omega) \oplus L^2(\Omega)$, $L = \nabla$, $D = \big\{y \in \mathcal{G} \bigm| |y|_2 \leqslant \mu \text{ a.e.}\big\}$ where $\mu \in \left]0, +\infty\right[$, and $r = 0$. With this scenario, (2.124) is equivalent to

$$
\underset{x\in H_0^1(\Omega)}{\text{minimize}}\ \ f(x) + \mu\,\mathrm{tv}(x) + \frac{1}{2}\|x - z\|^2, \tag{2.129}
$$

where $\mathrm{tv}(x) = \int_\Omega |\nabla x(\omega)|_2 d\omega$. In mechanics, such minimization problems have been studied extensively for certain potentials f [46]. For instance, $f = 0$ yields Mossolov's problem and its dual analysis is carried out in [46, Section IV.3.1]. In image processing, Mossolov's problem corresponds to the total variation denoising problem. Interestingly, in 1980, Mercier [59] proposed a dual projection algorithm to solve Mossolov's problem. This approach was independently rediscovered by Chambolle in a discrete setting [23, 24]. Next, we apply our framework to a discrete version of (2.129) for $N \times N$ images. This will extend the method of [24], which is restricted to $f = 0$, and provide a formal proof for its convergence (see also [79] for an alternative scheme based on Nesterov's algorithm [64]).

By way of preamble, let us introduce some notation. We denote by $y = \big(\eta_{k,l}^{(1)}, \eta_{k,l}^{(2)}\big)_{1\leqslant k,l\leqslant N}$ a generic element in $\mathbb{R}^{N\times N} \oplus \mathbb{R}^{N\times N}$ and by

$$\nabla\colon \mathbb{R}^{N\times N} \to \mathbb{R}^{N\times N} \oplus \mathbb{R}^{N\times N}\colon \big(\xi_{k,l}\big)_{1\leqslant k,l\leqslant N} \mapsto \big(\eta_{k,l}^{(1)}, \eta_{k,l}^{(2)}\big)_{1\leqslant k,l\leqslant N} \tag{2.130}$$

the discrete gradient operator, where

$$(\forall (k,l) \in \{1,\ldots,N\}^2) \quad \begin{cases} \eta_{k,l}^{(1)} = \xi_{k+1,l} - \xi_{k,l}, & \text{if } \ k < N; \\ \eta_{N,l}^{(1)} = 0; \\ \eta_{k,l}^{(2)} = \xi_{k,l+1} - \xi_{k,l}, & \text{if } \ l < N; \\ \eta_{k,N}^{(2)} = 0. \end{cases} \tag{2.131}$$

Now let $p \in [1, +\infty]$. Then p^* is the conjugate index of p, i.e., $p^* = +\infty$ if $p = 1$, $p^* = 1$ if $p = +\infty$, and $p^* = p/(p-1)$ otherwise. We define the p-th order discrete total variation function as

$$\mathrm{tv}_p\colon \mathbb{R}^{N\times N} \to \mathbb{R}\colon x \mapsto ||\nabla x||_{p,1}\,, \tag{2.132}$$

where

$$(\forall y \in \mathbb{R}^{N\times N} \oplus \mathbb{R}^{N\times N}) \quad \|y\|_{p,1} = \sum_{1\leqslant k,l\leqslant N} \big|(\eta_{k,l}^{(1)}, \eta_{k,l}^{(2)})\big|_p, \tag{2.133}$$

with

$$\big(\forall (\eta^{(1)}, \eta^{(2)}) \in \mathbb{R}^2\big) \quad \big|(\eta^{(1)}, \eta^{(2)})\big|_p = \begin{cases} \sqrt[p]{|\eta^{(1)}|^p + |\eta^{(2)}|^p}, & \text{if } \ p < +\infty; \\ \max\big\{|\eta^{(1)}|, |\eta^{(2)}|\big\}, & \text{if } \ p = +\infty. \end{cases} \tag{2.134}$$

In addition, the discrete divergence operator is defined as [23]

$$\mathrm{div}\colon \mathbb{R}^{N\times N} \oplus \mathbb{R}^{N\times N} \to \mathbb{R}^{N\times N}\colon \big(\eta_{k,l}^{(1)}, \eta_{k,l}^{(2)}\big)_{1\leqslant k,l\leqslant N} \mapsto \big(\xi_{k,l}^{(1)} + \xi_{k,l}^{(2)}\big)_{1\leqslant k,l\leqslant N}, \tag{2.135}$$

where

$$\xi_{k,l}^{(1)} = \begin{cases} \eta_{1,l}^{(1)} & \text{if } \ k = 1; \\ \eta_{k,l}^{(1)} - \eta_{k-1,l}^{(1)} & \text{if } \ 1 < k < N; \\ -\eta_{N-1,l}^{(1)} & \text{if } \ k = N; \end{cases} \quad \text{and} \quad \xi_{k,l}^{(2)} = \begin{cases} \eta_{k,1}^{(2)} & \text{if } \ l = 1; \\ \eta_{k,l}^{(2)} - \eta_{k,l-1}^{(2)} & \text{if } \ 1 < l < N; \\ -\eta_{k,N-1}^{(2)} & \text{if } \ l = N. \end{cases} \tag{2.136}$$

Problem 2.56 Let $z \in \mathbb{R}^{N\times N}$, let $f \in \Gamma_0(\mathbb{R}^{N\times N})$, let $\mu \in]0,+\infty[$, let $p \in [1,+\infty]$, and set

$$D_p = \left\{ \big(\nu^{(1)}_{k,l}, \nu^{(2)}_{k,l}\big)_{1\leqslant k,l\leqslant N} \in \mathbb{R}^{N\times N} \oplus \mathbb{R}^{N\times N} \;\middle|\; \max_{1\leqslant k,l\leqslant N} \big|(\nu^{(1)}_{k,l}, \nu^{(2)}_{k,l})\big|_{p^*} \leqslant 1 \right\}. \tag{2.137}$$

The problem is to

$$\underset{x\in\mathbb{R}^{N\times N}}{\text{minimize}} \;\; f(x) + \mu\, \mathrm{tv}_p(x) + \frac{1}{2}\|x-z\|^2, \tag{2.138}$$

and its dual is to

$$\underset{v\in D_p}{\text{minimize}} \;\; \widetilde{f}^*(z+\mu\,\mathrm{div}\, v). \tag{2.139}$$

Proposition 2.57 *Let $\big(\alpha^{(1)}_{n,k,l}\big)_{n\in\mathbb{N}}$ and $\big(\alpha^{(2)}_{n,k,l}\big)_{n\in\mathbb{N}}$ be sequences in $\mathbb{R}^{N\times N}$ such that*

$$\sum_{n\in\mathbb{N}} \sqrt{\sum_{1\leqslant k,l\leqslant N} \big|\alpha^{(1)}_{n,k,l}\big|^2 + \big|\alpha^{(2)}_{n,k,l}\big|^2} < +\infty, \tag{2.140}$$

let $(b_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}^{N\times N}$ such that $\sum_{n\in\mathbb{N}}\|b_n\| < +\infty$, and let $(x_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}}$ be sequences generated by the following routine, where $(\pi^{(1)}_p \mathsf{y}, \pi^{(2)}_p \mathsf{y})$ denotes the projection of a point $\mathsf{y}\in\mathbb{R}^2$ onto the closed unit ℓ^{p^*} ball in the Euclidean plane.

$$\begin{array}{l}
\text{Initialization}\\
\left\lfloor\begin{array}{l}
\varepsilon \in \left]0, \min\{1, \mu^{-1}/8\}\right[\\
v_0 = \big(\nu^{(1)}_{0,k,l}, \nu^{(2)}_{0,k,l}\big)_{1\leqslant k,l\leqslant N} \in \mathbb{R}^{N\times N}\oplus\mathbb{R}^{N\times N}
\end{array}\right.\\
\text{For } n = 0,1,\ldots\\
\left\lfloor\begin{array}{l}
x_n = \mathrm{prox}_f(z+\mu\,\mathrm{div}\, v_n) + b_n\\
\tau_n \in [\varepsilon, \mu^{-1}/4 - \varepsilon]\\
\big(\zeta^{(1)}_{n,k,l}, \zeta^{(2)}_{n,k,l}\big)_{1\leqslant k,l\leqslant N} = v_n + \tau_n \nabla x_n\\
\lambda_n \in [\varepsilon, 1]\\
\text{For every } (k,l)\in\{1,\ldots,N\}^2\\
\left\lfloor\begin{array}{l}
\nu^{(1)}_{n+1,k,l} = \nu^{(1)}_{n,k,l} + \lambda_n\Big(\pi^{(1)}_p\big(\zeta^{(1)}_{n,k,l}, \zeta^{(2)}_{n,k,l}\big) + \alpha^{(1)}_{n,k,l} - \nu^{(1)}_{n,k,l}\Big)\\
\nu^{(2)}_{n+1,k,l} = \nu^{(2)}_{n,k,l} + \lambda_n\Big(\pi^{(2)}_p\big(\zeta^{(1)}_{n,k,l}, \zeta^{(2)}_{n,k,l}\big) + \alpha^{(2)}_{n,k,l} - \nu^{(2)}_{n,k,l}\Big)
\end{array}\right.\\
v_{n+1} = \big(\nu^{(1)}_{n+1,k,l}, \nu^{(2)}_{n+1,k,l}\big)_{1\leqslant k,l\leqslant N}
\end{array}\right.
\end{array} \tag{2.141}$$

Then $(v_n)_{n\in\mathbb{N}}$ converges to a solution v to (2.139), $x = \mathrm{prox}_f(z+\mu\,\mathrm{div}\, v)$ is the primal solution to Problem 2.56, and $x_n \to x$.

Proof. It follows from (2.133) and (2.137) that $\|\cdot\|_{p,1} = \sigma_{D_p}$. Hence, Problem 2.56 is a special case of Problem 2.52 with $\mathcal{H} = \mathbb{R}^{N\times N}$, $\mathcal{G} = \mathbb{R}^{N\times N} \oplus \mathbb{R}^{N\times N}$, $L = \mu\nabla$ (see (2.130)), $D = D_p$, and $r = 0$. Moreover, $L^* = -\mu\,\mathrm{div}$ (see (2.135)), $\|L\| = \mu\|\nabla\| \leqslant 2\sqrt{2}\mu$ [23], and the projection of y onto the set D_p of (2.137) can be decomposed coordinatewise as

$$P_{D_p} y = \Big(\pi_p^{(1)}\big(\eta_{k,l}^{(1)}, \eta_{k,l}^{(2)}\big), \pi_p^{(2)}\big(\eta_{k,l}^{(1)}, \eta_{k,l}^{(2)}\big)\Big)_{1\leqslant k,l\leqslant N}. \tag{2.142}$$

Altogether, upon setting, for every $n \in \mathbb{N}$, $\tau_n = \mu\gamma_n$ and $a_n = \big(\alpha_{n,k,l}^{(1)}, \alpha_{n,k,l}^{(2)}\big)_{1\leqslant k,l\leqslant N}$, (2.141) appears as a special case of (2.126). The results therefore follow from (2.140) and Proposition 2.53. ☐

Remark 2.58 The inner loop in (2.141) performs the projection step. For certain values of p, this projection can be computed explicitly and we can therefore dispense with errors. Thus, if $p = 1$, then $p^* = +\infty$ and the projection loop becomes

$$\begin{array}{l}\text{For every } (k,l) \in \{1,\ldots,N\}^2 \\ \left\lfloor \begin{array}{l} \nu_{n+1,k,l}^{(1)} = \nu_{n,k,l}^{(1)} + \lambda_n\left(\dfrac{\zeta_{n,k,l}^{(1)}}{\max\left\{1, \left|\zeta_{n,k,l}^{(1)}\right|\right\}} - \nu_{n,k,l}^{(1)}\right) \\ \nu_{n+1,k,l}^{(2)} = \nu_{n,k,l}^{(2)} + \lambda_n\left(\dfrac{\zeta_{n,k,l}^{(2)}}{\max\left\{1, \left|\zeta_{n,k,l}^{(2)}\right|\right\}} - \nu_{n,k,l}^{(2)}\right). \end{array}\right.\end{array} \tag{2.143}$$

Likewise, if $p = 2$, then $p^* = 2$ and the projection loop becomes

$$\begin{array}{l}\text{For every } (k,l) \in \{1,\ldots,N\}^2 \\ \left\lfloor \begin{array}{l} \nu_{n+1,k,l}^{(1)} = \nu_{n,k,l}^{(1)} + \lambda_n\left(\dfrac{\zeta_{n,k,l}^{(1)}}{\max\left\{1, \left|\big(\zeta_{n,k,l}^{(1)}, \zeta_{n,k,l}^{(2)}\big)\right|_2\right\}} - \nu_{n,k,l}^{(1)}\right) \\ \nu_{n+1,k,l}^{(2)} = \nu_{n,k,l}^{(2)} + \lambda_n\left(\dfrac{\zeta_{n,k,l}^{(2)}}{\max\left\{1, \left|\big(\zeta_{n,k,l}^{(1)}, \zeta_{n,k,l}^{(2)}\big)\right|_2\right\}} - \nu_{n,k,l}^{(2)}\right). \end{array}\right.\end{array} \tag{2.144}$$

In the special case when $f = 0$, $\lambda_n \equiv 1$, and $\tau_n \equiv \tau \in \left]0, \mu^{-1}/4\right[$ the two resulting algorithms reduce to the popular methods proposed in [24]. Finally, if $p = +\infty$, then $p^* = 1$ and the efficient scheme described in [11] to project onto the ℓ^1 ball can be used.

2.3 Débruitage par variation totale sous contrainte

Nous proposons quelques simulations pour comparer le comportement numérique de l'Algorithme 2.3 avec d'autres méthodes. Pour chaque simulation, l'algorithme est exécuté avec $n = 5000$ itérations afin de trouver une solution approchée fiable x_{5000}. Ensuite, chaque algorithme est relancé avec le test d'arrêt relatif $\|x_n - x_{5000}\|/\|x_0 - x_{5000}\| \leqslant 10^{-9}$ afin de calculer le temps d'exécution.

Nous nous intéressons à la restauration d'une image originale $\widetilde{x} \in \mathbb{R}^{N\times N}$ à partir d'une image bruitée z dans $\mathbb{R}^{N\times N}$ selon le modèle

$$z = \widetilde{x} + w, \tag{2.145}$$

où w est un bruit additif. La formulation variationnelle suivante a été proposée dans [72] pour résoudre ce problème :

$$\underset{x\in\mathbb{R}^{N\times N}}{\text{minimiser}}\ \mu\|\nabla x\|_{2,1} + \frac{1}{2}\|x - z\|^2, \tag{2.146}$$

où μ est un paramètre strictement positif et $\|\nabla x\|_{2,1}$ est la variation totale de x (voir (2.130), (2.134)). L'avantage du modèle (2.146) est sa capacité à préserver des contours de l'image. Pourtant, ce modèle n'incorpore pas d'informations a priori sur l'image originale. Donc, on utilisera le modèle suivant (voir (2.138)) au lieu de (2.146),

$$\underset{x\in C}{\text{minimiser}}\ \mu\|\nabla x\|_{2,1} + \frac{1}{2}\|x - z\|^2, \tag{2.147}$$

où $C \subset \mathbb{R}^{N\times N}$ est un sous-ensemble convexe fermé non vide qui représente les information a priori sur l'image originale. Le problème (2.147) est un cas particulier du Problème 2.56 avec

$$f = \iota_C \quad \text{et} \quad p = 2. \tag{2.148}$$

On suppose que la composante basse-fréquence de l'image est connue [32],

$$C = \big\{x \in \mathbb{R}^{N\times N} \;\big|\; (\forall (i,j) \in \{1,\cdots,N/8\}^2)\ \hat{x}(i,j) = \widehat{\widetilde{x}}(i,j)\big\}, \tag{2.149}$$

où $\hat{x}$ est la transformée de Fourier discrète de x. La projection de x sur C est donnée explicitement dans [32, Eq. (6.27)] par

$$P_C x = \mathcal{F}^{-1}(\widehat{\widetilde{x}} 1_{K_1} + \hat{x} 1_{K_2}), \tag{2.150}$$

où $K_1 = \big\{(i,j) \;\big|\; 1 \leqslant i,j \leqslant N/8\big\}$ et $K_2 = \big\{(i,j) \;\big|\; 1 \leqslant i,j \leqslant N\big\}\backslash K_1$, 1_{K_1} et 1_{K_2} sont les fonctions caractéristiques de K_1 et K_2, respectivement, et $\mathcal{F}^{-1}$ est la transformée de Fourier inverse. Pour évaluer la performance, nous utilisons les algorithmes avec les paramètres suivants :

(i) L'algorithme (2.141) avec

$$\begin{cases} N = 64, p = 2, \varepsilon = 10^{-3}, v_0 = 0, \mu = 0.1 \\ (\forall n \in \mathbb{N}) \quad \lambda_n = 1, \tau_n = \mu^{-1}/4 - \varepsilon, b_n = 0 \\ (\forall n \in \mathbb{N})(\forall (k,l) \in \{1,\ldots,N\}^2) \quad \alpha^{(1)}_{n,k,l} = 0, \alpha^{(2)}_{n,k,l} = 0. \end{cases} \tag{2.151}$$

(ii) L'algorithme de Chen-Teboulle [29, Algorithm I] pour le problème (2.147) :

$$
\begin{array}{l}
\text{Initialisation} \\
\left\lfloor \begin{array}{l}
\varepsilon \in \left]0, \min\{1/3, 1/(2\mu\|\nabla\|+1)\}\right[\\
(x_0, v_0, y_0) \in \mathbb{R}^{N\times N} \times \left(\mathbb{R}^{N\times N} \times \mathbb{R}^{N\times N}\right) \times \left(\mathbb{R}^{N\times N} \times \mathbb{R}^{N\times N}\right)
\end{array} \right. \\
\text{Pour } n = 0, 1, \ldots \\
\left\lfloor \begin{array}{l}
\gamma_n \in [\varepsilon, \min\{(1-\varepsilon)/2, (1-\varepsilon)/(2\mu\|\nabla\|)\}] \\
p_{n+1} = v_n + \gamma_n(\mu\nabla x_n - y_n) \\
x_{n+1} = P_C\big((1+\gamma_n)^{-1}(x_n + \mu\gamma_n \operatorname{div} p_{n+1} + \gamma_n z)\big) \\
y_{n+1} = y_n + \gamma_n p_{n+1} - \gamma_n P_{D_2}(\gamma_n^{-1} y_n + p_{n+1}) \\
v_{n+1} = v_n + \gamma_n(\mu\nabla x_{n+1} - y_{n+1}),
\end{array} \right.
\end{array}
\tag{2.152}
$$

où P_{D_2} est donnée explicitement dans (2.142). On utilisera

$$
\varepsilon = 10^{-3}, (x_0, v_0, y_0) = 0, \mu = 0.1, \quad \text{et} \quad (\forall n \in \mathbb{N})\ \gamma_n = (1-\varepsilon)/2. \tag{2.153}
$$

Nous obtenons les résultats dans le tableau suivant et le figure 2.1 pour l'image de Lena

Lena, $N = 64$	L'algorithme (2.141)	L'algorithme de Chen-Teboulle
Temps d'exécution (s)	57	726

et dans le tableau suivant et le figure 2.6 pour l'image de Cameraman,

Cameraman, $N = 256$	L'algorithme (2.141)	L'algorithme de Chen-Teboulle
Temps d'exécution (s)	243	7745

On voit que l'algorithme (2.141) est plus rapide dans cet exemple.

Remarque 2.59 Depuis la parution de notre article dans [34] en 2010, un autre algorithme a été proposé dans [25, Algorithm 2] en 2011. Nous comparons ces deux algorithmes dans le tableau ci-dessous. Remarquons que notre algorithme est plus simple à mettre en œuvre, et ne nécessité à chaque itération que le stokage de deux variables de grande taille, à savoir x_n et v_n. Deux comparatifs sont fournis dans les figures 2.2 et 2.7 pour les images de Lena et Cameraman, respectivement.

L'Algorithme 2.3 (2010)	L'algorithme de Chambolle-Pock (2011)
Notons : $a_n \approx b_n \Leftrightarrow \sum_{n\in\mathbb{N}} \|a_n - b_n\| < +\infty$	$r = 0, h = f + \frac{1}{2}\|\cdot - z\|^2$
$(\gamma_n)_{n\in\mathbb{N}} \in \left]\varepsilon, 2\|L\|^{-2} - \varepsilon\right[^{\mathbb{N}}$ $x_n \approx \operatorname{prox}_f(z - L^* v_n)$ $v_{n+1} \approx \operatorname{prox}_{\gamma_n g^*}(v_n + \gamma_n(Lx_n - r)).$ (2.154)	$v_{n+1} = \operatorname{prox}_{\sigma_n g^*}(v_n + \sigma_n L\overline{x}_n)$ $x_{n+1} = \operatorname{prox}_{\tau_n h}(x_n - \tau_n L^* v_{n+1})$ $\lambda_n = 1 + \sqrt{1 + 2\tau_n}^{\,-1}$ $\tau_{n+1} = (\lambda_n - 1)\tau_n$ $\sigma_{n+1} = \sigma_n(\lambda_n - 1)^{-1}$ $\overline{x}_{n+1} = x_n + \lambda_n(x_{n+1} - x_n).$ (2.155)
Implémentation approchée des opérateur proximaux	Implémentation exacte des opérateur proximaux
Convergence forte de la suite primale	Pour n grand, $\|x_n - \overline{x}\| \leqslant c(\overline{x})/n$
Convergence faible de la suite duale	Pas de résultat correspondant

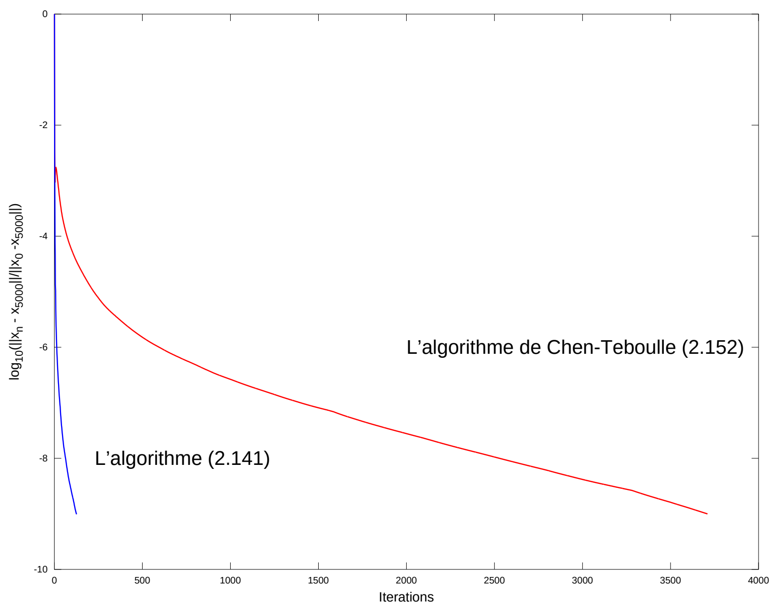

FIGURE 2.1 – Convergence de l’algorithme de Chen-Teboulle (2.152) et de l’algorithme (2.141) pour l’image de Lena.

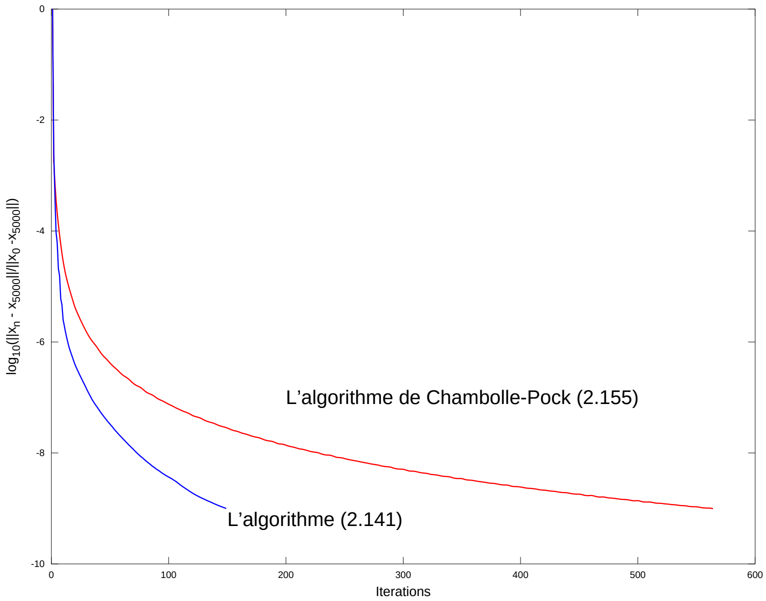

FIGURE 2.2 – Convergence de l'algorithme de Chambolle-Pock (2.155) et de l'algorithme (2.141) pour l'image de Lena.

FIGURE 2.3 – L'image originale.

FIGURE 2.4 – L'image bruitée (le bruit blanc de moyenne 0 et SNR $= 20$ dB).

FIGURE 2.5 – L'image débruitée par l'algorithme (2.141).

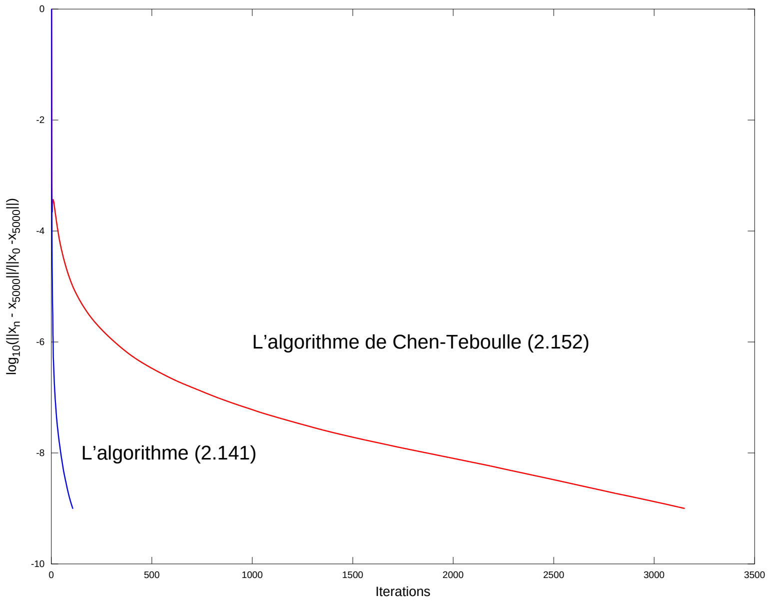

FIGURE 2.6 – Convergence de l'algorithme de Chen-Teboulle (2.152) et de l'algorithme (2.141) pour l'image de Cameraman.

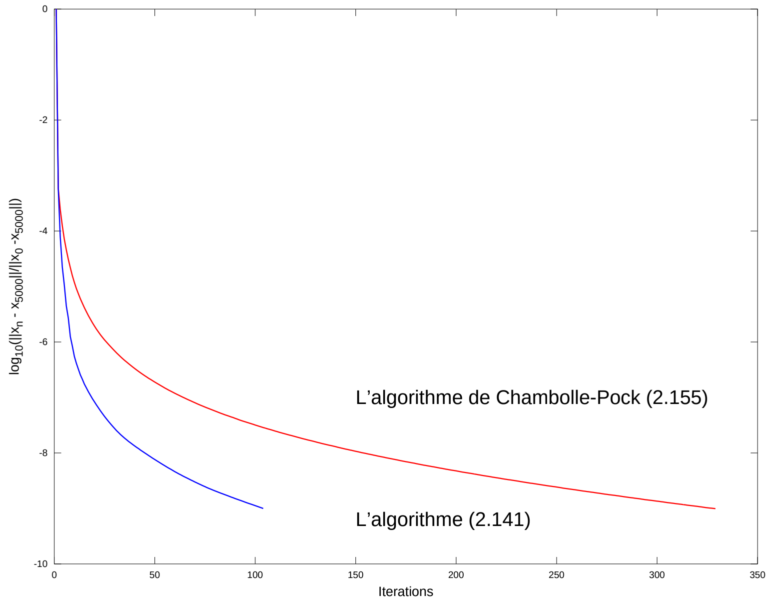

FIGURE 2.7 – Convergence de l'algorithme de Chambolle-Pock (2.155) et de l'algorithme (2.141) pour l'image de Cameraman.

FIGURE 2.8 – L'image originale.

FIGURE 2.9 – L'image bruitée (le bruit blanc de moyenne 0 et SNR $= 26$ dB).

FIGURE 2.10 – L'image débruitée par l'algorithme (2.141).

2.4 Bibliographie

[1] M. Amar and G. Bellettini, A notion of total variation depending on a metric with discontinuous coefficients, *Ann. Inst. H. Poincaré Anal. Non Linéaire,* vol. 11, pp. 91–133, 1994.

[2] H. C. Andrews and B. R. Hunt, *Digital Image Restoration,* Prentice-Hall, Englewood Cliffs, NJ, 1977.

[3] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing,* 2nd ed., Springer-Verlag, New York, 2006.

[4] J.-P. Aubin and H. Frankowska, *Set-Valued Analysis,* Birkhäuser, Boston, MA, 1990.

[5] H. H. Bauschke and P. L. Combettes, A Dykstra-like algorithm for two monotone operators, *Pacific J. Optim.,* vol. 4, pp. 383–391, 2008.

[6] H. H. Bauschke and P. L. Combettes, The Baillon-Haddad theorem revisited, *J. Convex Anal.,* vol. 17, 2010.

[7] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York, 2011.

[8] A. Beck and M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, *SIAM J. Imaging Sci.,* vol. 2, pp. 183–202, 2009.

[9] J. Bect, L. Blanc-Féraud, G. Aubert, and A. Chambolle, A ℓ^1 unified variational framework for image restoration, in *Proc. Eighth Europ. Conf. Comput. Vision,* Prague, 2004, T. Pajdla and J. Matas, eds., Lecture Notes in Comput. Sci. 3024, Springer-Verlag, New York, 2004, pp. 1–13.

[10] A. Ben-Tal, J. M. Borwein, and M. Teboulle, A dual approach to multidimensional L_p spectral estimation problems, *SIAM J. Control Optim.,* vol. 26, pp. 985–996, 1988.

[11] E. van den Berg and M. P. Friedlander, Probing the Pareto frontier for basis pursuit solutions, *SIAM J. Sci. Comput.,* vol. 31, 890–912, 2008.

[12] M. Bertero, C. De Mol, and E. R. Pike, Linear inverse problems with discrete data I – General formulation and singular system analysis, *Inverse Problems,* vol. 1, pp. 301–330, 1985.

[13] J. M. Bioucas-Dias and M. A. Figueiredo, A new TwIST : Two-step iterative shrinkage/thresholding algorithms for image restoration, *IEEE Trans. Image Process.,* vol. 16, pp. 2992–3004, 2007.

[14] J. M. Borwein, A. S. Lewis, and D. Noll, Maximum entropy reconstruction using derivative information. I : Fisher information and convex duality, *Math. Oper. Res.,* vol. 21, pp. 442–468, 1996.

[15] J. M. Borwein and D. R. Luke, Duality and convex programming, in : *Handbook of Imaging,* O. Scherzer Ed. Springer-Verlag, New York, pp. 229–270, 2011.

[16] K. Bredies and D. A. Lorenz, Linear convergence of iterative soft-thresholding, *J. Fourier Anal. Appl.,* vol. 14, pp. 813–837, 2008.

[17] 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.

[18] C. L. Byrne, *Signal Processing – A Mathematical Approach,* A. K. Peters, Wellesley, MA, 2005.

[19] J.-F. Cai, R. H. Chan, L. Shen, and Z. Shen, Convergence analysis of tight framelet approach for missing data recovery, *Adv. Comput. Math.,* vol. 31, pp. 87–113, 2009.

[20] J.-F. Cai, R. H. Chan, and Z. Shen, A framelet-based image inpainting algorithm, *Appl. Comput. Harm. Anal.,* vol. 24, pp. 131–149, 2008.

[21] Y. Censor and T. Elfving, A multiprojection algorithm using Bregman projections in a product space, *Numer. Algorithms,* vol. 8, pp. 221–239, 1994.

[22] Y. Censor and S. A. Zenios, *Parallel Optimization : Theory, Algorithms and Applications,* Oxford University Press, New York, 1997.

[23] A. Chambolle, An algorithm for total variation minimization and applications, *J. Math. Imaging Vision,* vol. 20, pp. 89–97, 2004.

[24] A. Chambolle, Total variation minimization and a class of binary MRF model, *Lecture Notes in Comput. Sci.* 3757, pp 136–152, 2005.

[25] A. Chambolle and T. Pock, A first order primal dual algorithm for convex problems with applications to imaging, *J. Math Imaging Vision*, vol. 40, pp. 120–145, 2011.

[26] T. F. Chan, G. H. Golub, and P. Mulet, A nonlinear primal-dual method for total variation-based image restoration, *SIAM J. Sci. Comput.,* vol. 20, 1964–1977, 1999.

[27] C. Chaux, P. L. Combettes, J.-C. Pesquet, and V. R. Wajs, A variational formulation for frame-based inverse problems, *Inverse Problems,* vol. 23, pp. 1495–1518, 2007.

[28] C. Chaux, J.-C. Pesquet, and N. Pustelnik, Nested iterative algorithms for convex contrained image recovery problems, *SIAM J. Imaging Sci.,* vol. 2, pp. 730–762, 2009.

[29] G. Chen and M. Teboulle, A proximal based decomposition method for minimization problems, *Mathematical programming*, vol. 64, pp. 81-101, 1994.

[30] P. L. Combettes, Signal recovery by best feasible approximation, *IEEE Trans. Image Process.,* vol. 2, pp. 269–271, 1993.

[31] P. L. Combettes, Inconsistent signal feasibility problems : Least-squares solutions in a product space, *IEEE Trans. Signal Process.,* vol. 42, pp. 2955–2966, 1994.

[32] P. L. Combettes, The convex feasibility problem in image recovery, in *Advances in Imaging and Electron Physics* (P. Hawkes, Ed.), vol. 95, pp. 155–270, Academic Press, New York, 1996.

[33] P. L. Combettes, Đinh Dũng, and B. C. Vũ, Dualization of signal recovery problems, preprint, July 2, 2009. http://arxiv.org/abs/0907.0436

[34] P. L. Combettes, Đinh Dũng, and B. C. Vũ, Dualization of signal recovery problems, *Set-Valued Var. Anal.*, vol. 18, pp. 373–404, 2010.

[35] P. L. Combettes and J.-C. Pesquet, Proximal thresholding algorithm for minimization over orthonormal bases, *SIAM J. Optim.,* vol. 18, pp. 1351–1376, 2007.

[36] P. L. Combettes and J.-C. Pesquet, A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery, *IEEE J. Selected Topics Signal Process.,* vol. 1, pp. 564–574, 2007.

[37] P. L. Combettes and J.-C. Pesquet, A proximal decomposition method for solving convex variational inverse problems, *Inverse Problems,* vol. 24, Art. 065014, 27 pp., 2008.

[38] P. L. Combettes and H. J. Trussell, The use of noise properties in set theoretic estimation, *IEEE Trans. Signal Process.,* vol. 39, pp. 1630–1641, 1991.

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

[40] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, PA, 1992.

[41] I. Daubechies, M. Defrise, and C. De Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, *Comm. Pure Appl. Math.,* vol. 57, pp. 1413–1457, 2004.

[42] I. Daubechies, G. Teschke, and L. Vese, Iteratively solving linear inverse problems under general convex constraints, *Inverse Probl. Imaging,* vol. 1, pp. 29–46, 2007.

[43] P. Destuynder, M. Jaoua, and H. Sellami, A dual algorithm for denoising and preserving edges in image processing, *J. Inverse Ill-Posed Probl.,* vol. 15, pp. 149–165, 2007.

[44] D. Donoho and I. Johnstone, Ideal spatial adaptation via wavelet shrinkage, *Biometrika,* vol. 81, pp. 425–455, 1994.

[45] S. Durand and M. Nikolova, Denoising of frame coefficients using ℓ^1 data-fidelity term and edge-preserving regularization, *Multiscale Model. Simul.,* vol. 6, pp. 547–576, 2007.

[46] I. Ekeland and R. Temam, *Analyse Convexe et Problèmes Variationnels,* Dunod, Paris, 1974; *Convex Analysis and Variational Problems,* SIAM, Philadelphia, PA, 1999.

[47] J. Fadili and G. Peyré, Total variation projection with first order schemes, *IEEE Trans., Image Process.,* vol. 20, pp. 657–669, 2011.

[48] W. Fenchel, *Convex Cones, Sets and Functions,* Lecture notes (mimeograph), Princeton University, 1953.

[49] M. Fornasier, Domain decomposition methods for linear inverse problems with sparsity constraints, *Inverse Problems,* vol. 23, pp. 2505–2526, 2007.

[50] R. W. Gerchberg, Super-resolution through error energy reduction, *Optica Acta,* vol. 21, pp. 709–720, 1974.

[51] M. Hintermüller and G. Stadler, An infeasible primal-dual algorithm for total bounded variation-based inf-convolution-type image restoration, *SIAM J. Sci. Comput.,* vol. 28, pp. 1–23, 2006.

[52] Y. Huang, M. K. Ng, and Y.-W. Wen, A fast total variation minimization method for image restoration, *Multiscale Model. Simul.,* vol. 7, pp. 774–795, 2008.

[53] A. N. Iusem and M. Teboulle, A regularized dual-based iterative method for a class of image reconstruction problems, *Inverse Problems,* vol. 9, pp. 679–696, 1993.

[54] T. Kärkkäinen, K. Majava, and M. M. Mäkelä, Comparison of formulations and solution methods for image restoration problems, *Inverse Problems,* vol. 17, pp. 1977–1995, 2001.

[55] R. M. Leahy and C. E. Goutis, An optimal technique for constraint-based image restoration and reconstruction, *IEEE Trans. Acoust., Speech, Signal Process.,* vol. 34, pp. 1629–1642, 1986.

[56] L. Levi, Fitting a bandlimited signal to given points, *IEEE Trans. Inform. Theory,* vol. 11, pp. 372–376, 1965.

[57] S. G. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed. Academic Press, New York, 1999.

[58] B. P. Medoff, Image reconstruction from limited data : Theory and applications in computerized tomography, in : H. Stark (Ed.), *Image Recovery : Theory and Application,* pp. 321–368. Academic Press, San Diego, CA, 1987.

[59] B. Mercier, *Inéquations Variationnelles de la Mécanique* (Publications Mathématiques d'Orsay, no. 80.01). Orsay, France, Université de Paris-XI, 1980.

[60] J.-J. Moreau, Fonctions convexes duales et points proximaux dans un espace hilbertien, *C. R. Acad. Sci. Paris Sér. A Math.,* vol. 255, pp. 2897–2899, 1962.

[61] J.-J. Moreau, Proximité et dualité dans un espace hilbertien, *Bull. Soc. Math. France,* vol. 93, pp. 273-299, 1965.

[62] J.-J. Moreau, *Fonctionnelles Convexes,* Séminaire sur les Équations aux Dérivées Partielles II, Collège de France, Paris, 1966–1967.

[63] A. S. Nemirovsky and D. B. Yudin, *Problem Complexity and Method Efficiency in Optimization,* Wiley, New York, 1983.

[64] Yu. Nesterov, Smooth minimization of non-smooth functions, *Math. Program.,* vol. 103, pp. 127–152, 2005.

[65] D. Noll, Reconstruction with noisy data : An approach via eigenvalue optimization, *SIAM J. Optim.,* vol. 8, pp. 82–104, 1998.

[66] A. Papoulis, A new algorithm in spectral analysis and band-limited extrapolation, *IEEE Trans. Circuits Syst.,* vol. 22, pp. 735–742, 1975.

[67] E. Polak, *Computational Methods in Optimization : A Unified Approach*, Academic Press, New York, 1971.

[68] L. C. Potter and K. S. Arun, A dual approach to linear inverse problems with convex constraints, *SIAM J. Control Optim.,* vol. 31, pp. 1080–1092, 1993.

[69] R. T. Rockafellar, Duality and stability in extremum problems involving convex functions, *Pacific J. Math.,* vol. 21, pp. 167–187, 1967.

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

[71] R. T. Rockafellar, *Conjugate Duality and Optimization,* SIAM, Philadelphia, PA, 1974.

[72] L. I. Rudin, S. Osher, and E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D,* vol. 60, pp. 259–268, 1992.

[73] H. Stark (Ed.), *Image Recovery : Theory and Application,* Academic Press, San Diego, CA, 1987.

[74] A. O. Steinhardt, R. K. Goodrich, and R. A. Roberts, Spectral estimation via minimum energy correlation extension, *IEEE Trans. Acoust., Speech, Signal Process.,* vol. 33, pp. 1509–1515, 1985.

[75] J. A. Tropp, Just relax : Convex programming methods for identifying sparse signals in noise, *IEEE Trans. Inform. Theory,* vol. 52, pp. 1030–1051, 2006.

[76] H. J. Trussell and M. R. Civanlar, The feasible solution in signal restoration, *IEEE Trans. Acoust., Speech, Signal Process.,* vol. 32, pp. 201–212, 1984.

[77] S. Twomey, The application of numerical filtering to the solution of integral equations encountered in indirect sensing measurements, *J. Franklin Inst.,* vol. 279, pp. 95–109, 1965.

[78] J. B. Weaver, Y. Xu, D. M. Healy Jr., and L. D. Cromwell, Filtering noise from images with wavelet transforms, *Magnetic Resonance in Medicine,* vol. 21, pp. 288–295, 1991.

[79] P. Weiss, G. Aubert, and L. Blanc-Féraud, Efficient schemes for total variation minimization under constraints in image processing, *SIAM J. Sci. Comput.*, vol. 31, pp. 2047–2080, 2009.

[80] D. C. Youla, Generalized image restoration by the method of alternating orthogonal projections, *IEEE Trans. Circuits Syst.,* vol. 25, pp. 694–702, 1978.

[81] D. C. Youla and V. Velasco, Extensions of a result on the synthesis of signals in the presence of inconsistent constraints, *IEEE Trans. Circuits Syst.,* vol. 33, pp. 465–468, 1986.

[82] D. C. Youla and H. Webb, Image restoration by the method of convex projections : Part 1 – theory, *IEEE Trans. Medical Imaging,* vol. 1, pp. 81–94, 1982.

[83] C. Zălinescu, *Convex Analysis in General Vector Spaces,* World Scientific, River Edge, NJ, 2002.

Chapitre 3

Proximité pour les sommes de fonctions composites

Nous proposons un algorithme pour calculer l'opérateur proximal d'une somme de fonctions composites. La convergence de l'algorithme est démontrée dans des espaces hilbertiens réels. Des applications sont présentées.

3.1 Description et résultats principaux

On a vu que la suite $(x_n)_{n\in\mathbb{N}}$ engendrée par l'Algorithme 2.3 converge fortement vers $\operatorname{prox}_h z$ avec $h\colon \mathcal{H} \to \left]-\infty,+\infty\right] : x \mapsto f(x)+g(Lx-r)$. Dans ce chapitre, nous traitons le cas où h est une somme de fonctions composites. Nous considérons le problème suivant.

Problème 3.1 Soient $z \in \mathcal{H}$ et $(\omega_i)_{1\leqslant i\leqslant m}$ des réels dans $\left]0,1\right]$ tels que $\sum_{i=1}^m \omega_i = 1$. Pour tout $i \in \{1,\ldots,m\}$, soient $(\mathcal{G}_i, \|\cdot\|_{\mathcal{G}_i})$ un espace hilbertien réel, $r_i \in \mathcal{G}_i$, $g_i \in \Gamma_0(\mathcal{G}_i)$, et $0 \neq L_i \in \mathcal{B}(\mathcal{H},\mathcal{G}_i)$. Le problème est de

$$\underset{x\in\mathcal{H}}{\text{minimiser}} \ \sum_{i=1}^m \omega_i g_i(L_i x - r_i) + \frac{1}{2}\|x-z\|^2. \tag{3.1}$$

Nous supposons que les opérateurs proximaux de $(g_i)_{1\leqslant i\leqslant m}$ sont calculables de manière approchée. Nous cherchons donc une méthode qui permet d'utiliser individuellement les opérateurs proximaux de $(g_i)_{1\leqslant i\leqslant m}$, et d'éclater les structures composites. Nous proposons l'algorithme suivant.

Algorithme 3.2 Pour tout $i \in \{1,\ldots,m\}$, soit $(a_{i,n})_{n\in\mathbb{N}}$ une suite dans $\mathcal{G}_i$.

$$
\begin{array}{l}
\text{Initialisation} \\
\left\lfloor
\begin{array}{l}
\rho = \big(\max_{1\leqslant i\leqslant m}\|L_i\|\big)^{-2} \\
\varepsilon \in \,]0,\min\{1,\rho\}[\\
\text{Pour } i = 1,\ldots,m \\
\left\lfloor\; v_{i,0} \in \mathcal{G}_i \right.
\end{array}
\right. \\
\text{Pour } n = 0,1,\ldots \\
\left\lfloor
\begin{array}{l}
x_n = z - \sum_{i=1}^m \omega_i L_i^* v_{i,n} \\
\gamma_n \in [\varepsilon, 2\rho-\varepsilon] \\
\lambda_n \in [\varepsilon, 1] \\
\text{Pour } i = 1,\ldots,m \\
\left\lfloor\; v_{i,n+1} = v_{i,n} + \lambda_n\Big(\operatorname{prox}_{\gamma_n g_i^*}\big(v_{i,n}+\gamma_n(L_i x_n - r_i)\big) + a_{i,n} - v_{i,n}\Big). \right.
\end{array}
\right.
\end{array}
\tag{3.2}
$$

En utilisant une technique d'espace produit, nous montrons le résultat de convergence suivant.

Théorème 3.3 *Supposons que*

$$(r_i)_{1\leqslant i\leqslant m} \in \operatorname{sri}\big\{(L_i x - y_i)_{1\leqslant i\leqslant m} \;\big|\; x \in \mathcal{H}, (y_i)_{1\leqslant i\leqslant m} \in \times_{i=1}^m \operatorname{dom} g_i\big\} \tag{3.3}$$

et

$$(\forall i \in \{1,\ldots,m\}) \quad \sum_{n\in\mathbb{N}} \|a_{i,n}\|_{\mathcal{G}_i} < +\infty. \tag{3.4}$$

De plus, soient $(x_n)_{n\in\mathbb{N}}$, $(v_{1,n})_{n\in\mathbb{N}}$, ..., $(v_{m,n})_{n\in\mathbb{N}}$ des suites engendrées par l'Algorithme 3.2. Alors, le Problème 3.1 possède une solution unique $\overline{x}$ et on a les résultats suivants.

(i) *Pour tout $i \in \{1,\ldots,m\}$, $(v_{i,n})_{n\in\mathbb{N}}$ converge faiblement vers un point $\overline{v_i} \in \mathcal{G}_i$. De plus, $(\overline{v_i})_{1\leqslant i\leqslant m}$ est une solution du problème dual*

$$\underset{v_1\in\mathcal{G}_1,\ldots,v_m\in\mathcal{G}_m}{\text{minimiser}} \ \frac{1}{2}\left\|z - \sum_{i=1}^m \omega_i L_i^* v_i\right\|^2 + \sum_{i=1}^m \omega_i\big(g_i^*(v_i) + \langle v_i \mid r_i\rangle\big), \tag{3.5}$$

et $\overline{x} = z - \sum_{i=1}^m \omega_i L_i^* \overline{v_i}$.

(ii) *$(x_n)_{n\in\mathbb{N}}$ converge fortement vers $\overline{x}$.*

Exemple 3.4 Soient $\mathcal{H}$ un espace hilbertien réel, et $z \in \mathcal{H}$. Pour tout $i \in \{1,\ldots,m\}$, soient $(\mathcal{G}_i, \|\cdot\|_{\mathcal{G}_i})$ un espace hilbertien réel, $r_i \in \mathcal{G}_i$, $C_i \subset \mathcal{G}_i$ un sous-ensemble convexe fermé non vide, et $0 \neq L_i \in \mathcal{B}(\mathcal{H},\mathcal{G}_i)$. Le problème est de

$$\underset{x\in D}{\text{minimiser}} \ \|x - z\|, \quad \text{où} \quad D = \bigcap_{i=1}^m \big\{x \in \mathcal{H} \;\big|\; L_i x \in r_i + C_i\big\}. \tag{3.6}$$

Cet exemple est un cas particulier du Problème 3.1 avec $(\forall i \in \{1, \dots, m\})$ $g_i = \iota_{C_i} \in \Gamma_0(\mathcal{G}_i)$ et $\omega_i = 1/m$. Alors, remplaçons $\operatorname{prox}_{\gamma_n g_i^*}$ dans (3.2) par $\operatorname{Id} - P_{C_i}$, nous obtenons un algorithme numérique pour résoudre le problème (3.6) (voir Section 3.2.3.1).

Enfin, nous nous intéressons à la résolution de problèmes de traitement du signal dans des espaces hilbertiens. Nous cherchons un signal original $\overline{x} \in \mathcal{H}$ à partir de l'observation de p signaux dégradés,

$$(\forall i \in \{1, \dots, p\}) \quad r_i = T_i \overline{x} + s_i. \tag{3.7}$$

Dans ce modèle, pour tout $i \in \{1, \dots, p\}$, le signal dégradé r_i est dans un espace hilbertien réel $\mathcal{G}_i$, $T_i \in \mathcal{B}(\mathcal{H}, \mathcal{G}_i)$, et $s_i \in \mathcal{G}_i$ est un bruit. Afin de trouver une solution x, nous résoudrons le problème variationnel suivant.

Exemple 3.5 Soient $(\omega_i)_{1 \leqslant i \leqslant p+2}$ des réels dans $]0, 1]$ tels que $\sum_{i=1}^{p+2} \omega_i = 1$, $\Omega \subset \mathbb{R}^2$ un sous-ensemble non vide, borné et ouvert, $\mathcal{H} = \mathrm{H}_0^1(\Omega)$, et $(e_k)_{k \in \mathbb{N}}$ une base orthonormale de $\mathcal{H}$. Pour tout $i \in \{1, \dots, p\}$, soient $(\mathcal{G}_i, \|\cdot\|_{\mathcal{G}_i})$ un espace hilbertien réel, $r_i \in \mathcal{G}_i$, et $0 \neq T_i \in \mathcal{B}(\mathcal{H}, \mathcal{G}_i)$. Le problème est de

$$\underset{x \in \mathcal{H}}{\text{minimiser}} \quad \sum_{i=1}^{p} \omega_i \|T_i x - r_i\|_{\mathcal{G}_i} + \sum_{k \in \mathbb{N}} \left(\omega_{p+1} |\langle x \mid e_k \rangle| + \frac{1}{2} |\langle x \mid e_k \rangle|^2 \right) + \omega_{p+2} \operatorname{tv}_2(x), \tag{3.8}$$

où $\operatorname{tv}_2(x) = \int_\Omega |\nabla x(\omega)|_2 d\omega$ est la variation total de x.

Cet exemple est un cas particulier du Problème 3.1 avec

$$\begin{cases} \mathcal{H} = \mathrm{H}_0^1(\Omega), m = p+2, z = 0; \\ (\forall x \in \mathcal{H}) \quad \|x - z\|^2 = \sum_{k \in \mathbb{N}} |\langle x \mid e_k \rangle|^2; \\ (\forall i \in \{1, \dots, p\}) \quad g_i = \|\cdot\|_{\mathcal{G}_i} \text{ et } L_i = T_i; \\ \mathcal{G}_{p+1} = \ell^2(\mathbb{N}),\ g_{p+1} = \|\cdot\|_{\ell^1},\ r_{p+1} = 0, \text{ et } L_{p+1} \colon x \mapsto (\langle x \mid e_k \rangle)_{k \in \mathbb{N}}; \\ \mathcal{G}_{p+2} = \mathrm{L}^2(\Omega) \oplus \mathrm{L}^2(\Omega),\ g_{p+2} \colon y \mapsto \int_\Omega |y(\omega)|_2 d\omega,\ r_{p+2} = 0, \text{ et } L_{p+2} = \nabla. \end{cases} \tag{3.9}$$

De plus, les opérateurs proximaux des fonctions $(g_i)_{1 \leqslant i \leqslant m}$ sont disponibles (voir Section 3.2.3.2). On peut donc utiliser l'Algorithme 3.2 pour résoudre le problème (3.8) (voir Section 3.2.3.2).

3.2 Article en anglais

PROXIMITY FOR SUMS OF COMPOSITE FUNCTIONS [1]

Abstract : We propose an algorithm for computing the proximity operator of a sum of composite convex functions in Hilbert spaces and investigate its asymptotic behavior. Applications to best approximation and image recovery are described.

3.2.1 Introduction

Let $\mathcal{H}$ be a real Hilbert space with scalar product $\langle \cdot \mid \cdot \rangle$ and associated norm $\|\cdot\|$. The best approximation to a point $z \in \mathcal{H}$ from a nonempty closed convex set $C \subset \mathcal{H}$ is the point $P_C z \in C$ that satisfies $\|P_C z - z\| = \min_{x \in C} \|x - z\|$. The induced best approximation operator $P_C \colon \mathcal{H} \to C$, also called the projector onto C, plays a central role in several branches of applied mathematics [10]. If we designate by ι_C the indicator function of C, i.e.,

$$\iota_C \colon x \mapsto \begin{cases} 0, & \text{if } x \in C; \\ +\infty, & \text{if } x \notin C, \end{cases} \tag{3.10}$$

then $P_C z$ is the solution to the minimization problem

$$\underset{x \in \mathcal{H}}{\text{minimize}} \;\; \iota_C(x) + \frac{1}{2}\|x - z\|^2. \tag{3.11}$$

Now let $\Gamma_0(\mathcal{H})$ be the class of lower semicontinuous convex functions $f \colon \mathcal{H} \to \left]-\infty, +\infty\right]$ such that $\operatorname{dom} f = \big\{x \in \mathcal{H} \;\big|\; f(x) < +\infty\big\} \neq \varnothing$. In [13] Moreau observed that, for every function $f \in \Gamma_0(\mathcal{H})$, the proximal minimization problem

$$\underset{x \in \mathcal{H}}{\text{minimize}} \;\; f(x) + \frac{1}{2}\|x - z\|^2 \tag{3.12}$$

possesses a unique solution, which he denoted by $\operatorname{prox}_f z$. The resulting proximity operator $\operatorname{prox}_f \colon \mathcal{H} \to \mathcal{H}$ therefore extends the notion of a best approximation operator for a convex set. This fruitful concept has become a central tool in mechanics, variational analysis, optimization, and signal processing, e.g., [1, 7, 16].

Though in certain simple cases closed-form expressions are available [7, 8, 14], computing $\operatorname{prox}_f z$ in numerical applications is a challenging task. The objective of this paper is to propose a splitting algorithm to compute proximity operators in the case when f can be decomposed as a sum of composite functions.

1. P. L. Combettes, Đinh Dũng, and B. C. Vũ, Proximity for sums of composite functions, *J. Math. Anal. Appl.*, vol. 380, pp. 680–688, 2011.

Problem 3.6 Let $z \in \mathcal{H}$ and let $(\omega_i)_{1\leqslant i\leqslant m}$ be reals in $]0,1]$ such that $\sum_{i=1}^m \omega_i = 1$. For every $i \in \{1,\dots,m\}$, let $(\mathcal{G}_i, \|\cdot\|_{\mathcal{G}_i})$ be a real Hilbert space, let $r_i \in \mathcal{G}_i$, let $g_i \in \Gamma_0(\mathcal{G}_i)$, and let $L_i\colon \mathcal{H} \to \mathcal{G}_i$ be a nonzero bounded linear operator. The problem is to

$$\underset{x\in\mathcal{H}}{\text{minimize}} \quad \sum_{i=1}^m \omega_i g_i(L_i x - r_i) + \frac{1}{2}\|x-z\|^2. \tag{3.13}$$

The underlying practical assumption we make is that the proximity operators $(\mathrm{prox}_{g_i})_{1\leqslant i\leqslant m}$ are implementable (to within some quantifiable error). We are therefore aiming at devising an algorithm that uses these operators separately. Let us note that such splitting algorithms are already available to solve Problem 3.6 under certain restrictions.

A) Suppose that $\mathcal{G}_1 = \mathcal{H}$, that $L_1 = \mathrm{Id}$, that the functions $(g_i)_{2\leqslant i\leqslant m}$ are differentiable everywhere with a Lipschitz continuous gradient, and that $r_i \equiv 0$. Then (3.13) reduces to the minimization of the sum of $f_1 = g_1 \in \Gamma_0(\mathcal{H})$ and of the smooth function $f_2 = \sum_{i=2}^m \omega_i g_i \circ L_i + \|\cdot - z\|^2/2$, and it can be solved by the forward-backward algorithm [8, 18].

B) The methods proposed in [4] address the case when, for every $i \in \{1,\dots,m\}$, $\mathcal{G}_i = \mathcal{H}$, $L_i = \mathrm{Id}$, and $r_i = 0$.

C) The method proposed in [5] addresses the case when $m = 2$, $\mathcal{G}_1 = \mathcal{H}$, and $L_1 = \mathrm{Id}$, and $r_1 = 0$.

The restrictions imposed in A) are quite stringent since many problems involve at least two nondifferentiable potentials. Let us also observe that since, in general, there is no explicit expression for $\mathrm{prox}_{g_i\circ L_i}$ in terms of prox_{g_i} and L_i, Problem 3.6 cannot be reduced to the setting described in B). On the other hand, using a product space reformulation, we shall show that the setting described in C) can be exploited to solve Problem 3.6 using only approximate implementations of the operators $(\mathrm{prox}_{g_i})_{1\leqslant i\leqslant m}$. Our algorithm is introduced in Section 3.2.2, where we also establish its convergence properties. In Section 3.2.3, our results are applied to best approximation and image recovery problems.

Our notation is standard. $\mathcal{B}(\mathcal{H},\mathcal{G})$ is the space of bounded linear operators from $\mathcal{H}$ to a real Hilbert space $\mathcal{G}$. The adjoint of $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ is denoted by L^*. The conjugate of $f \in \Gamma_0(\mathcal{H})$ is the function $f^* \in \Gamma_0(\mathcal{H})$ defined by $f^*\colon u \mapsto \sup_{x\in\mathcal{H}}(\langle x \mid u\rangle - f(x))$. The projector onto a nonempty closed convex set $C \subset \mathcal{H}$ is denoted by P_C. The strong relative interior of a convex set $C \subset \mathcal{H}$ is

$$\mathrm{sri}\, C = \big\{x \in C \;\big|\; \mathrm{cone}(C-x) = \overline{\mathrm{span}}\,(C-x)\big\}, \quad \text{where} \quad \mathrm{cone}\, C = \bigcup_{\lambda>0} \big\{\lambda x \;\big|\; x \in C\big\}, \tag{3.14}$$

and the relative interior of C is $\mathrm{ri}\, C = \big\{x \in C \;\big|\; \mathrm{cone}(C-x) = \mathrm{span}\,(C-x)\big\}$. We have $\mathrm{int}\, C \subset \mathrm{sri}\, C \subset \mathrm{ri}\, C \subset C$ and, if $\mathcal{H}$ is finite-dimensional, $\mathrm{ri}\, C = \mathrm{sri}\, C$. For background on convex analysis, see [19].

3.2.2 Main result

To solve Problem 3.6, we propose the following algorithm. Its main features are that each function g_i is activated individually by means of its proximity operator, and that the proximity operators can be evaluated simultaneously. It is important to stress that the functions $(g_i)_{1\leqslant i\leqslant m}$ and the operators $(L_i)_{1\leqslant i\leqslant m}$ are used at separate steps in the algorithm, which is thus fully decomposed. In addition, an error $a_{i,n}$ is tolerated in the evaluation of the ith proximity operator at iteration n.

Algorithm 3.7 For every $i \in \{1,\ldots,m\}$, let $(a_{i,n})_{n\in\mathbb{N}}$ be a sequence in $\mathcal{G}_i$.

$$
\begin{array}{l}
\text{Initialization}\\
\left\lfloor
\begin{array}{l}
\rho = \big(\max_{1\leqslant i\leqslant m}\|L_i\|\big)^{-2}\\
\varepsilon \in \left]0, \min\{1,\rho\}\right[\\
\text{For } i = 1,\ldots,m\\
\left\lfloor\; v_{i,0} \in \mathcal{G}_i\right.
\end{array}
\right.\\
\text{For } n = 0,1,\ldots\\
\left\lfloor
\begin{array}{l}
x_n = z - \sum_{i=1}^m \omega_i L_i^* v_{i,n}\\
\gamma_n \in [\varepsilon, 2\rho-\varepsilon]\\
\lambda_n \in [\varepsilon, 1]\\
\text{For } i = 1,\ldots,m\\
\left\lfloor\; v_{i,n+1} = v_{i,n} + \lambda_n\Big(\operatorname{prox}_{\gamma_n g_i^*}\big(v_{i,n} + \gamma_n(L_i x_n - r_i)\big) + a_{i,n} - v_{i,n}\Big).\right.
\end{array}
\right.
\end{array}
\tag{3.15}
$$

Note that an alternative implementation of (3.15) can be obtained via Moreau's decomposition formula in a real Hilbert space $\mathcal{G}$ [8, Lemma 2.10]

$$
(\forall g \in \Gamma_0(\mathcal{G}))(\forall \gamma \in \left]0,+\infty\right[)(\forall v \in \mathcal{G}) \quad \operatorname{prox}_{\gamma g^*} v = v - \gamma \operatorname{prox}_{\gamma^{-1} g}(\gamma^{-1} v). \tag{3.16}
$$

We now describe the asymptotic behavior of Algorithm 3.7.

Theorem 3.8 *Suppose that*

$$
(r_i)_{1\leqslant i\leqslant m} \in \operatorname{sri}\Big\{(L_i x - y_i)_{1\leqslant i\leqslant m} \;\Big|\; x \in \mathcal{H}, (y_i)_{1\leqslant i\leqslant m} \in \textstyle\times_{i=1}^m \operatorname{dom} g_i\Big\} \tag{3.17}
$$

and that

$$
(\forall i \in \{1,\ldots,m\}) \quad \sum_{n\in\mathbb{N}} \|a_{i,n}\|_{\mathcal{G}_i} < +\infty. \tag{3.18}
$$

Furthermore, let $(x_n)_{n\in\mathbb{N}}$, $(v_{1,n})_{n\in\mathbb{N}}$, ..., $(v_{m,n})_{n\in\mathbb{N}}$ be sequences generated by Algorithm 3.7. Then Problem 3.6 possesses a unique solution x and the following hold.

(i) *For every* $i \in \{1,\dots,m\}$, $(v_{i,n})_{n\in\mathbb{N}}$ *converges weakly to a point* $v_i \in \mathcal{G}_i$. *Moreover,* $(v_i)_{1\leqslant i\leqslant m}$ *is a solution to the minimization problem*

$$\underset{v_1\in\mathcal{G}_1,\dots,v_m\in\mathcal{G}_m}{\text{minimize}} \quad \frac{1}{2}\left\|z-\sum_{i=1}^m \omega_i L_i^* v_i\right\|^2 + \sum_{i=1}^m \omega_i\big(g_i^*(v_i)+\langle v_i \mid r_i\rangle\big), \tag{3.19}$$

and $x = z - \sum_{i=1}^m \omega_i L_i^* v_i$.

(ii) $(x_n)_{n\in\mathbb{N}}$ *converges strongly to* x.

Proof. Set $f\colon \mathcal{H}\to\,]-\infty,+\infty] : x\mapsto \sum_{i=1}^m \omega_i g_i(L_i x - r_i)$. The assumptions imply that, for every $i\in\{1,\dots,m\}$, the function $x\mapsto g_i(L_ix-r_i)$ is convex and lower semicontinuous. Hence, f is likewise. On the other hand, it follows from (3.17) that

$$(r_i)_{1\leqslant i\leqslant m} \in \big\{(L_ix-y_i)_{1\leqslant i\leqslant m} \;\big|\; x\in\mathcal{H}, (y_i)_{1\leqslant i\leqslant m}\in \times_{i=1}^m \operatorname{dom} g_i\big\} \tag{3.20}$$

and, therefore, that $\operatorname{dom} f \neq \varnothing$. Thus, $f\in\Gamma_0(\mathcal{H})$ and, as seen in (3.12), Problem 3.6 possesses a unique solution, namely $x = \operatorname{prox}_f z$.

Now let $\boldsymbol{\mathcal{H}}$ be the real Hilbert space obtained by endowing the Cartesian product $\mathcal{H}^m$ with the scalar product $\langle\cdot\mid\cdot\rangle_{\boldsymbol{\mathcal{H}}}\colon (\boldsymbol{x},\boldsymbol{y})\mapsto \sum_{i=1}^m \omega_i\langle x_i \mid y_i\rangle$, where $\boldsymbol{x}=(x_i)_{1\leqslant i\leqslant m}$ and $\boldsymbol{y}=(y_i)_{1\leqslant i\leqslant m}$ denote generic elements in $\boldsymbol{\mathcal{H}}$. The associated norm is

$$\|\cdot\|_{\boldsymbol{\mathcal{H}}}\colon \boldsymbol{x}\mapsto \sqrt{\sum_{i=1}^m \omega_i\|x_i\|^2}. \tag{3.21}$$

Likewise, let $\boldsymbol{\mathcal{G}}$ denote the real Hilbert space obtained by endowing the Cartesian product $\mathcal{G}_1\times\cdots\times\mathcal{G}_m$ with the scalar product and the associated norm respectively defined by

$$\langle\cdot\mid\cdot\rangle_{\boldsymbol{\mathcal{G}}}\colon(\boldsymbol{y},\boldsymbol{z})\mapsto\sum_{i=1}^m\omega_i\langle y_i\mid z_i\rangle_{\mathcal{G}_i} \quad\text{and}\quad \|\cdot\|_{\boldsymbol{\mathcal{G}}}\colon\boldsymbol{y}\mapsto\sqrt{\sum_{i=1}^m\omega_i\|y_i\|_{\mathcal{G}_i}^2}. \tag{3.22}$$

Define

$$\begin{cases}\boldsymbol{f}=\iota_{\boldsymbol{D}}, \quad \text{where} \quad \boldsymbol{D}=\big\{(x,\dots,x)\in\boldsymbol{\mathcal{H}} \;\big|\; x\in\mathcal{H}\big\}\\ \boldsymbol{g}\colon\boldsymbol{\mathcal{G}}\to\,]-\infty,+\infty] : \boldsymbol{y}\mapsto\sum_{i=1}^m\omega_i g_i(y_i)\\ \boldsymbol{L}\colon\boldsymbol{\mathcal{H}}\to\boldsymbol{\mathcal{G}} : \boldsymbol{x}\mapsto(L_ix_i)_{1\leqslant i\leqslant m}\\ \boldsymbol{r}=(r_1,\dots,r_m)\\ \boldsymbol{z}=(z,\dots,z).\end{cases} \tag{3.23}$$

Then $\boldsymbol{f}\in\Gamma_0(\boldsymbol{\mathcal{H}})$, $\boldsymbol{g}\in\Gamma_0(\boldsymbol{\mathcal{G}})$, and $\boldsymbol{L}\in\mathcal{B}(\boldsymbol{\mathcal{H}},\boldsymbol{\mathcal{G}})$. Moreover, $\boldsymbol{D}$ is a closed vector subspace of $\boldsymbol{\mathcal{H}}$ with projector

$$\operatorname{prox}_{\boldsymbol{f}} = P_{\boldsymbol{D}}\colon\boldsymbol{x}\mapsto\left(\sum_{i=1}^m\omega_ix_i,\dots,\sum_{i=1}^m\omega_ix_i\right) \tag{3.24}$$

and

$$\boldsymbol{L}^*\colon \boldsymbol{\mathcal{G}} \to \boldsymbol{\mathcal{H}}\colon \boldsymbol{v} \mapsto \big(L_i^* v_i\big)_{1\leqslant i\leqslant m}. \tag{3.25}$$

Note that (3.22) and (3.21) yield

$$\begin{aligned}(\forall \boldsymbol{x} \in \boldsymbol{\mathcal{H}}) \quad \|\boldsymbol{L}\boldsymbol{x}\|_{\boldsymbol{\mathcal{G}}}^2 &= \sum_{i=1}^m \omega_i \|L_i x_i\|_{\mathcal{G}_i}^2 \\ &\leqslant \sum_{i=1}^m \omega_i \|L_i\|^2 \|x_i\|^2 \\ &\leqslant \Big(\max_{1\leqslant i\leqslant m} \|L_i\|^2\Big) \sum_{i=1}^m \omega_i \|x_i\|^2 \\ &= \Big(\max_{1\leqslant i\leqslant m} \|L_i\|^2\Big) \|\boldsymbol{x}\|_{\boldsymbol{\mathcal{H}}}^2. \end{aligned} \tag{3.26}$$

Therefore,

$$\|\boldsymbol{L}\| \leqslant \max_{1\leqslant i\leqslant m} \|L_i\|. \tag{3.27}$$

We also deduce from (3.17) that

$$\boldsymbol{r} \in \operatorname{sri}\big(\boldsymbol{L}(\operatorname{dom} \boldsymbol{f}) - \operatorname{dom} \boldsymbol{g}\big). \tag{3.28}$$

Furthermore, in view of (3.21) and (3.23), in the space $\boldsymbol{\mathcal{H}}$, (3.13) is equivalent to

$$\underset{\boldsymbol{x}\in\boldsymbol{\mathcal{H}}}{\text{minimize}} \;\; \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{L}\boldsymbol{x} - \boldsymbol{r}) + \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{z}\|_{\boldsymbol{\mathcal{H}}}^2. \tag{3.29}$$

Next, we derive from [5, Proposition 3.3] that the dual problem of (3.29) is to

$$\underset{\boldsymbol{v}\in\boldsymbol{\mathcal{G}}}{\text{minimize}} \;\; \widetilde{\boldsymbol{f}^*}(\boldsymbol{z} - \boldsymbol{L}^*\boldsymbol{v}) + \boldsymbol{g}^*(\boldsymbol{v}) + \langle \boldsymbol{v} \mid \boldsymbol{r}\rangle_{\boldsymbol{\mathcal{G}}}, \tag{3.30}$$

where $\widetilde{\boldsymbol{f}^*}\colon \boldsymbol{u} \mapsto \inf_{\boldsymbol{w}\in\boldsymbol{\mathcal{H}}} \big(\boldsymbol{f}^*(\boldsymbol{w}) + (1/2)\|\boldsymbol{u} - \boldsymbol{w}\|_{\boldsymbol{\mathcal{H}}}^2\big)$ is the Moreau envelope of $\boldsymbol{f}^*$. Since $\boldsymbol{f} = \iota_{\boldsymbol{D}}$, we have $\boldsymbol{f}^* = \iota_{\boldsymbol{D}^\perp}$. Hence, (3.21) and (3.24) yield

$$(\forall \boldsymbol{u} \in \boldsymbol{\mathcal{H}}) \quad \widetilde{\boldsymbol{f}^*}(\boldsymbol{u}) = \frac{1}{2}\|\boldsymbol{u} - P_{\boldsymbol{D}^\perp}\boldsymbol{u}\|_{\boldsymbol{\mathcal{H}}}^2 = \frac{1}{2}\|P_{\boldsymbol{D}}\boldsymbol{u}\|_{\boldsymbol{\mathcal{H}}}^2 = \frac{1}{2}\left\|\sum_{i=1}^m \omega_i u_i\right\|^2. \tag{3.31}$$

On the other hand, (3.22) and (3.23) yield

$$(\forall \boldsymbol{v} \in \boldsymbol{\mathcal{G}}) \quad \boldsymbol{g}^*(\boldsymbol{v}) = \sum_{i=1}^m \omega_i g_i^*(v_i) \quad \text{and} \quad \operatorname{prox}_{\boldsymbol{g}^*} \boldsymbol{v} = \big(\operatorname{prox}_{g_i^*} v_i\big)_{1\leqslant i\leqslant m}. \tag{3.32}$$

Altogether, it follows from (3.25), (3.31), (3.32), and (3.22), that

$$\text{(3.30) is equivalent to (3.19).} \tag{3.33}$$

Now define

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} \boldsymbol{x}_n = (x_n, \ldots, x_n) \\ \boldsymbol{v}_n = (v_{1,n}, \ldots, v_{m,n}) \\ \boldsymbol{a}_n = (a_{1,n}, \ldots, a_{m,n}). \end{cases} \tag{3.34}$$

Then, in view of (3.23), (3.24), (3.25), (3.27), and (3.32), (3.15) is a special case of the following routine.

$$\begin{array}{l} \text{Initialization} \\ \left\lfloor \begin{array}{l} \rho = \|\boldsymbol{L}\|^{-2} \\ \varepsilon \in \left]0, \min\{1, \rho\}\right[\\ \boldsymbol{v}_0 \in \boldsymbol{\mathcal{G}} \end{array} \right. \\ \text{For } n = 0, 1, \ldots \\ \left\lfloor \begin{array}{l} \boldsymbol{x}_n = \operatorname{prox}_{\boldsymbol{f}}(\boldsymbol{z} - \boldsymbol{L}^*\boldsymbol{v}_n) \\ \gamma_n \in [\varepsilon, 2\rho - \varepsilon] \\ \lambda_n \in [\varepsilon, 1] \\ \boldsymbol{v}_{n+1} = \boldsymbol{v}_n + \lambda_n\big(\operatorname{prox}_{\gamma_n \boldsymbol{g}^*}(\boldsymbol{v}_n + \gamma_n(\boldsymbol{L}\boldsymbol{x}_n - \boldsymbol{r})) + \boldsymbol{a}_n - \boldsymbol{v}_n\big). \end{array} \right. \end{array} \tag{3.35}$$

Moreover, (3.18) implies that $\sum_{n\in\mathbb{N}} \|\boldsymbol{a}_n\|_{\boldsymbol{\mathcal{G}}} < +\infty$. Hence, it follows from (3.28) and [5, Theorem 3.7] that the following hold, where $\boldsymbol{x}$ is the solution to (3.29).

(a) $(\boldsymbol{v}_n)_{n\in\mathbb{N}}$ converges weakly to a solution $\boldsymbol{v}$ to (3.30) and $\boldsymbol{x} = \operatorname{prox}_{\boldsymbol{f}}(\boldsymbol{z} - \boldsymbol{L}^*\boldsymbol{v})$.

(b) $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ converges strongly to $\boldsymbol{x}$.

In view of (3.21), (3.22), (3.23), (3.24), (3.25), (3.33), and (3.34), items (a) and (b) provide respectively items (i) and (ii). □

Remark 3.9 Let us consider Problem 3.6 in the special case when $(\forall i \in \{1, \ldots, m\})$ $\mathcal{G}_i = \mathcal{H}$, $L_i = \operatorname{Id}$, and $r_i = 0$. Then (3.13) reduces to

$$\underset{x\in\mathcal{H}}{\text{minimize}} \quad \sum_{i=1}^{m} \omega_i g_i(x) + \frac{1}{2}\|x - z\|^2. \tag{3.36}$$

Now let us implement Algorithm 3.7 with $\gamma_n \equiv 1$, $\lambda_n \equiv 1$, $a_{i,n} \equiv 0$, and $v_{i,0} \equiv 0$. The iteration process resulting from (3.15) can be written as

$$\begin{array}{l} \text{Initialization} \\ \left\lfloor \begin{array}{l} x_0 = z \\ \text{For } i = 1, \ldots, m \\ \left\lfloor \; v_{i,0} = 0 \right. \end{array} \right. \\ \text{For } n = 0, 1, \ldots \\ \left\lfloor \begin{array}{l} \text{For } i = 1, \ldots, m \\ \left\lfloor \; v_{i,n+1} = \operatorname{prox}_{g_i^*}(x_n + v_{i,n}). \right. \\ x_{n+1} = z - \sum_{i=1}^{m} \omega_i v_{i,n+1}. \end{array} \right. \end{array} \tag{3.37}$$

For every $i \in \{1, \dots, m\}$ and $n \in \mathbb{N}$, set $z_{i,n} = x_n + v_{i,n}$. Then (3.37) yields

$$
\begin{array}{l}
\text{Initialization} \\
\left\lfloor
\begin{array}{l}
x_0 = z \\
\text{For } i = 1, \dots, m \\
\left\lfloor z_{i,0} = z \right.
\end{array}
\right. \\
\text{For } n = 0, 1, \dots \\
\left\lfloor
\begin{array}{l}
x_{n+1} = z - \sum_{i=1}^m \omega_i \mathrm{prox}_{g_i^*} z_{i,n} \\
\text{For } i = 1, \dots, m \\
\left\lfloor z_{i,n+1} = x_{n+1} + \mathrm{prox}_{g_i^*} z_{i,n}. \right.
\end{array}
\right.
\end{array}
\tag{3.38}
$$

Next we observe that $(\forall n \in \mathbb{N})\ \sum_{i=1}^m \omega_i z_{i,n} = z$. Indeed, the identity is clearly satisfied for $n = 0$ and, for every $n \in \mathbb{N}$, (3.38) yields $\sum_{i=1}^m \omega_i z_{i,n+1} = x_{n+1} + \sum_{i=1}^m \omega_i \mathrm{prox}_{g_i^*} z_{i,n} = (z - \sum_{i=1}^m \omega_i \mathrm{prox}_{g_i^*} z_{i,n}) + \sum_{i=1}^m \omega_i \mathrm{prox}_{g_i^*} z_{i,n} = z$. Thus, invoking (3.16) with $\gamma = 1$, we can rewrite (3.38) as

$$
\begin{array}{l}
\text{Initialization} \\
\left\lfloor
\begin{array}{l}
x_0 = z \\
\text{For } i = 1, \dots, m \\
\left\lfloor z_{i,0} = z \right.
\end{array}
\right. \\
\text{For } n = 0, 1, \dots \\
\left\lfloor
\begin{array}{l}
x_{n+1} = \sum_{i=1}^m \omega_i \mathrm{prox}_{g_i} z_{i,n} \\
\text{For } i = 1, \dots, m \\
\left\lfloor z_{i,n+1} = x_{n+1} + z_{i,n} - \mathrm{prox}_{g_i} z_{i,n}. \right.
\end{array}
\right.
\end{array}
\tag{3.39}
$$

This is precisely the Dykstra-like algorithm proposed in [4, Theorem 4.2] for computing $\mathrm{prox}_{\sum_{i=1}^m \omega_i g_i} z$ (which itself extends the classical parallel Dykstra algorithm for projecting z onto an intersection of closed convex sets [2, 11]). Hence, Algorithm 3.7 can be viewed as an extension of this algorithm, which was derived and analyzed with different techniques in [4].

3.2.3 Applications

As noted in the Introduction, special cases of Problem 3.6 have already been considered in the literature under certain restrictions on the number m of composite functions, the complexity of the linear operators $(L_i)_{1 \leqslant i \leqslant m}$, and/or the smoothness of the potentials $(g_i)_{1 \leqslant i \leqslant m}$ (one will find specific applications in [3, 5, 7, 8, 9, 15] and the references therein). The proposed framework makes it possible to remove these restrictions simultaneously. In this section, we provide two illustrations.

3.2.3.1 Best approximation from an intersection of composite convex sets

In this section, we consider the problem of finding the best approximation $P_D z$ to a point $z \in \mathcal{H}$ from a closed convex subset D of $\mathcal{H}$ defined as an intersection of affine inverse images of closed convex sets.

Problem 3.10 Let $z \in \mathcal{H}$ and, for every $i \in \{1, \ldots, m\}$, let $(\mathcal{G}_i, \|\cdot\|_{\mathcal{G}_i})$ be a real Hilbert space, let $r_i \in \mathcal{G}_i$, let C_i be a nonempty closed convex subset of $\mathcal{G}_i$, and let $0 \neq L_i \in \mathcal{B}(\mathcal{H}, \mathcal{G}_i)$. The problem is to

$$\underset{x \in D}{\text{minimize}}\ \|x - z\|, \quad \text{where} \quad D = \bigcap_{i=1}^{m} \big\{x \in \mathcal{H} \ \big|\ L_i x \in r_i + C_i\big\}. \tag{3.40}$$

In view of (3.10), Problem 3.10 is a special case of Problem 3.6, where $(\forall i \in \{1, \ldots, m\})$ $g_i = \iota_{C_i}$ and $\omega_i = 1/m$. It follows that, for every $i \in \{1, \ldots, m\}$ and every $\gamma \in \left]0, +\infty\right[$, $\operatorname{prox}_{\gamma g_i}$ reduces to the projector P_{C_i} onto C_i. Hence, using (3.16), we can rewrite Algorithm 3.7 in the following form, where we have set $c_{i,n} = -\gamma_n^{-1} a_{i,n}$ for simplicity.

Algorithm 3.11 For every $i \in \{1, \ldots, m\}$, let $(c_{i,n})_{n \in \mathbb{N}}$ be a sequence in $\mathcal{G}_i$.

$$
\begin{array}{l}
\text{Initialization}\\
\left\lfloor
\begin{array}{l}
\rho = \big(\max_{1 \leqslant i \leqslant m} \|L_i\|\big)^{-2}\\
\varepsilon \in \left]0, \min\{1, \rho\}\right[\\
\text{For } i = 1, \ldots, m\\
\left\lfloor\ v_{i,0} \in \mathcal{G}_i \right.
\end{array}
\right.\\
\text{For } n = 0, 1, \ldots\\
\left\lfloor
\begin{array}{l}
x_n = z - \sum_{i=1}^{m} \omega_i L_i^* v_{i,n}\\
\gamma_n \in [\varepsilon, 2\rho - \varepsilon]\\
\lambda_n \in [\varepsilon, 1]\\
\text{For } i = 1, \ldots, m\\
\left\lfloor\ v_{i,n+1} = v_{i,n} + \gamma_n \lambda_n \Big(L_i x_n - r_i - P_{C_i}\big(\gamma_n^{-1} v_{i,n} + L_i x_n - r_i\big) - c_{i,n}\Big). \right.
\end{array}
\right.
\end{array}
\tag{3.41}
$$

In the light of the above, we obtain the following application of Theorem 3.8(ii).

Corollary 3.12 *Suppose that*

$$(r_i)_{1 \leqslant i \leqslant m} \in \operatorname{sri}\Big\{(L_i x - y_i)_{1 \leqslant i \leqslant m} \ \Big|\ x \in \mathcal{H}, (y_i)_{1 \leqslant i \leqslant m} \in \bigtimes_{i=1}^{m} C_i\Big\} \tag{3.42}$$

and that $(\forall i \in \{1, \ldots, m\})$ $\sum_{n \in \mathbb{N}} \|c_{i,n}\|_{\mathcal{G}_i} < +\infty$. *Then every sequence* $(x_n)_{n \in \mathbb{N}}$ *generated by Algorithm 3.11 converges strongly to the solution* $P_D z$ *to Problem 3.10.*

3.2.3.2 Nonsmooth image recovery

A wide range of signal and image recovery problems can be modeled as instances of Problem 3.6. In this section, we focus on the problem of recovering an image $\overline{x} \in \mathcal{H}$ from p noisy measurements

$$r_i = T_i\overline{x} + s_i, \quad 1 \leqslant i \leqslant p. \tag{3.43}$$

In this model, the ith measurement r_i lies in a Hilbert space $\mathcal{G}_i$, $T_i \in \mathcal{B}(\mathcal{H}, \mathcal{G}_i)$ is the data formation operator, and $s_i \in \mathcal{G}_i$ is the realization of a noise process. A typical data fitting potential in such models is the function

$$x \mapsto \sum_{i=1}^{p} \omega_i g_i(T_i x - r_i), \quad \text{where} \quad 0 \leqslant g_i \in \Gamma_0(\mathcal{G}_i) \quad \text{and } g_i \text{ vanishes only at } 0. \tag{3.44}$$

The proposed framework can handle $p \geqslant 1$ nondifferentiable functions $(g_i)_{1\leqslant i\leqslant p}$ as well as the incorporation of additional potential functions to model prior knowledge on the original image $\overline{x}$. In the illustration we provide below, the following is assumed.

– The image space is $\mathcal{H} = \mathrm{H}_0^1(\Omega)$, where Ω is a nonempty bounded open domain in $\mathbb{R}^2$.
– $\overline{x}$ admits a sparse decomposition in an orthonormal basis $(e_k)_{k\in\mathbb{N}}$ of $\mathcal{H}$. As discussed in [9, 20] this property can be promoted by the “elastic net” potential $x \mapsto \sum_{k\in\mathbb{N}} \phi_k(\langle x \mid e_k\rangle)$, where $(\forall k \in \mathbb{N})\ \phi_k \colon \xi \mapsto \alpha|\xi| + \beta|\xi|^2$, with $\alpha > 0$ and $\beta > 0$. More general choices of suitable functions $(\phi_k)_{k\in\mathbb{N}}$ are available [6].
– $\overline{x}$ is piecewise smooth. This property is promoted by the total variation potential $\mathrm{tv}(x) = \int_\Omega |\nabla x(\omega)|_2 d\omega$, where $|\cdot|_2$ denotes the Euclidean norm on $\mathbb{R}^2$ [17].

Upon setting $g_i \equiv \|\cdot\|_{\mathcal{G}_i}$ in (3.44), these considerations lead us to the following formulation (see [5, Example 2.10] for more general nonsmooth potentials).

Problem 3.13 Let $\mathcal{H} = \mathrm{H}_0^1(\Omega)$, where $\Omega \subset \mathbb{R}^2$ is nonempty, bounded, and open, let $(\omega_i)_{1\leqslant i\leqslant p+2}$ be reals in $]0,1]$ such that $\sum_{i=1}^{p+2} \omega_i = 1$, and let $(e_k)_{k\in\mathbb{N}}$ be an orthonormal basis of $\mathcal{H}$. For every $i \in \{1,\ldots,p\}$, let $0 \neq T_i \in \mathcal{B}(\mathcal{H}, \mathcal{G}_i)$, where $(\mathcal{G}_i, \|\cdot\|_{\mathcal{G}_i})$ is a real Hilbert space, and let $r_i \in \mathcal{G}_i$. The problem is to

$$\underset{x\in\mathcal{H}}{\text{minimize}} \quad \sum_{i=1}^{p} \omega_i \|T_i x - r_i\|_{\mathcal{G}_i} + \sum_{k\in\mathbb{N}} \left(\omega_{p+1}|\langle x \mid e_k\rangle| + \frac{1}{2}|\langle x \mid e_k\rangle|^2\right) + \omega_{p+2}\,\mathrm{tv}(x). \tag{3.45}$$

It follows from Parseval’s identity that Problem 3.13 is a special case of Problem 3.6 in $\mathcal{H} = \mathrm{H}_0^1(\Omega)$ with $m = p+2$, $z = 0$, and

$$\begin{cases} g_i = \|\cdot\|_{\mathcal{G}_i} \text{ and } L_i = T_i, \text{ if } 1 \leqslant i \leqslant p; \\ \mathcal{G}_{p+1} = \ell^2(\mathbb{N}),\ g_{p+1} = \|\cdot\|_{\ell^1},\ r_{p+1} = 0, \text{ and } L_{p+1} \colon x \mapsto (\langle x \mid e_k\rangle)_{k\in\mathbb{N}}; \\ \mathcal{G}_{p+2} = \mathrm{L}^2(\Omega) \oplus \mathrm{L}^2(\Omega),\ g_{p+2} \colon y \mapsto \int_\Omega |y(\omega)|_2 d\omega,\ r_{p+2} = 0, \text{ and } L_{p+2} = \nabla. \end{cases} \tag{3.46}$$

To implement Algorithm 3.7, it suffices to note that $L_{p+1}^*\colon (\nu_k)_{k\in\mathbb{N}} \mapsto \sum_{k\in\mathbb{N}} \nu_k e_k$ and $L_{p+2}^* = -\operatorname{div}$, and to specify the proximity operators of the functions $(\gamma g_i^*)_{1\leqslant i\leqslant m}$, where $\gamma \in \left]0,+\infty\right[$. First, let $i \in \{1,\ldots,p\}$. Then $g_i = \|\cdot\|_{\mathcal{G}_i}$ and therefore $g_i^* = \iota_{B_i}$, where B_i is the closed unit ball of $\mathcal{G}_i$. Hence $\operatorname{prox}_{\gamma g_i^*} = P_{B_i}$. Next, it follows from (3.16) and [8, Example 2.20] that $\operatorname{prox}_{\gamma g_{p+1}^*}\colon (\xi_k)_{k\in\mathbb{N}} \mapsto (P_{[-1,1]}\xi_k)_{k\in\mathbb{N}}$. Finally, since g_{p+2} is the support function of the set [12]

$$K = \big\{ y \in \mathcal{G}_{p+2} \;\big|\; |y|_2 \leqslant 1 \text{ a.e.} \big\}, \tag{3.47}$$

$g_{p+2}^* = \iota_K$ and therefore $\operatorname{prox}_{\gamma g_{p+2}^*} = P_K$, which is straightforward to compute. Altogether, as $\|L_{p+1}\| = 1$ and $\|L_{p+2}\| \leqslant 1$, Algorithm 3.7 assumes the following form (since all the proximity operators can be implemented with simple projections, we dispense with the errors terms).

Algorithm 3.14

$$
\begin{array}{l}
\text{Initialization}\\
\left\lfloor
\begin{array}{l}
\rho = \big(\max\{1,\|T_1\|,\ldots,\|T_p\|\}\big)^{-2}\\
\varepsilon \in \left]0,\min\{1,\rho\}\right[\\
\text{For } i=1,\ldots,p\\
\left\lfloor\; v_{i,0} \in \mathcal{G}_i \right.\\
v_{p+1,0} = (\nu_{k,0})_{k\in\mathbb{N}} \in \ell^2(\mathbb{N})\\
v_{p+2,0} \in \mathrm{L}^2(\Omega)\oplus \mathrm{L}^2(\Omega)
\end{array}\right.\\
\text{For } n=0,1,\ldots\\
\left\lfloor
\begin{array}{l}
x_n = z - \sum_{i=1}^p \omega_i T_i^* v_{i,n} - \omega_{p+1}\sum_{k\in\mathbb{N}} \nu_{k,n} e_k + \omega_{p+2}\operatorname{div} v_{p+2,n}\\
\gamma_n \in [\varepsilon, 2\rho-\varepsilon]\\
\lambda_n \in [\varepsilon,1]\\
\text{For } i=1,\ldots,p\\
\left\lfloor\; v_{i,n+1} = v_{i,n} + \lambda_n\left(\dfrac{v_{i,n}+\gamma_n(T_i x_n - r_i)}{\max\{1,\|v_{i,n}+\gamma_n(T_i x_n - r_i)\|_{\mathcal{G}_i}\}} - v_{i,n}\right)\right.\\
\text{For every } k\in\mathbb{N},\; \nu_{k,n+1} = \nu_{k,n} + \lambda_n\left(\dfrac{\nu_{k,n}+\gamma_n\langle x_n \mid e_k\rangle}{\max\{1,|\nu_{k,n}+\gamma_n\langle x_n \mid e_k\rangle|\}} - \nu_{k,n}\right)\\
\text{For almost every } \omega\in\Omega,\\
\qquad v_{p+2,n+1}(\omega) = v_{p+2,n}(\omega) + \lambda_n\left(\dfrac{v_{p+2,n}(\omega)+\gamma_n\nabla x_n(\omega)}{\max\{1,|v_{p+2,n}(\omega)+\gamma_n\nabla x_n(\omega)|_2\}} - v_{p+2,n}(\omega)\right).
\end{array}\right.
\end{array}
\tag{3.48}
$$

Let us establish the main convergence property of this algorithm.

Corollary 3.15 *Every sequence $(x_n)_{n\in\mathbb{N}}$ generated by Algorithm 3.14 converges strongly to the solution to Problem 3.13.*

Proof. In view of the above discussion and of Theorem 3.8(ii), it remains to check that (3.17) is satisfied. Set $S = \big\{(L_i x - y_i)_{1\leqslant i\leqslant m} \;\big|\; x\in\mathcal{H}, (y_i)_{1\leqslant i\leqslant m} \in \times_{i=1}^m \mathrm{dom}\, g_i\big\}$. We have $\mathrm{dom}\, g_i = \mathcal{G}_i$ for every $i \in \{1,\ldots,p\}$, $\mathrm{dom}\, g_{p+1} = \ell^1(\mathbb{N})$, and $\mathrm{dom}\, g_{p+2} = \mathrm{L}^2(\Omega)\oplus\mathrm{L}^2(\Omega)$. Consequently,

$$\begin{aligned} S &= \Big\{(T_1x - y_1,\ldots,T_px-y_p,(\langle x \mid e_k\rangle - \eta_k)_{k\in\mathbb{N}}, \nabla x - y_{p+2} \;\Big| \\ &\qquad\qquad x\in\mathcal{H}, (y_i)_{1\leqslant i\leqslant p}\in \times_{i=1}^p\mathcal{G}_i, (\eta_k)_{k\in\mathbb{N}}\in\ell^1(\mathbb{N}), y_{p+2}\in \mathrm{L}^2(\Omega)\oplus\mathrm{L}^2(\Omega)\Big\} \\ &= \big(\times_{i=1}^p \mathcal{G}_i\big)\times\ell^2(\mathbb{N})\times\big(\mathrm{L}^2(\Omega)\oplus\mathrm{L}^2(\Omega)\big) \\ &= \times_{i=1}^m \mathcal{G}_i. \end{aligned} \tag{3.49}$$

Hence, we trivially have $(r_1,\ldots,r_p,0,0)\in \mathrm{sri}\, S$. □

Let us emphasize that a novelty of the above variational framework is to perform total variation image recovery in the presence of several nondifferentiable composite terms, with guaranteed strong convergence to the solution to the problem, and with elementary steps in the form of simple projections. The finite-dimensional version of the algorithm can easily be obtained by discretizing the operators ∇ and div as in [3] (see also [5, Section 4.4] for variants of the total variation potential).

3.3 Résultats numériques

Dans cette section, nous illustrons une application de l'Algorithme 3.2 en traitement de l'image.

3.3.1 Débruitage par variation totale sous contrainte

Nous nous intéressons au problème (2.145). On utilise deux contraintes au lieu d'une, comme dans (2.147). Le problème est de

$$\underset{x\in C_1\cap C_2}{\text{minimiser}}\; \frac{1}{3}\mu\|\nabla x\|_{2,1} + \frac{1}{2}\|x-z\|^2, \tag{3.50}$$

où μ est un paramètre strictement positif, C_1 est donné par (2.149) et

$$C_2 = \big\{x\in\mathbb{R}^{N\times N} \;\big|\; \big(\forall (i,j)\in\{1,\ldots,N\}^2\big)\; 0\leqslant x(i,j)\leqslant 1\big\}. \tag{3.51}$$

Le problème (3.50) correspond à un cas particulier du Problème 3.1, où

$$\begin{cases} \mathcal{H} = \mathcal{G}_1 = \mathcal{G}_2 = \mathbb{R}^{N\times N}, \mathcal{G}_3 = \mathcal{H}\times\mathcal{H}, \\ L_1 = L_2 = \mathrm{Id}\,, L_3 = \mu\nabla, r_1 = r_2 = r_3 = 0, \\ \omega_1=\omega_2=\omega_3 = 1/3, g_1 = \iota_{C_1}, g_2 = \iota_{C_2}, g_3 = \|\cdot\|_{2,1}. \end{cases}$$

Nous utiliserons l'Algorithme 3.2 avec les paramètres

$$\begin{cases} N = 256, v_{1,0} = 0, v_{2,0} = 0, v_{3,0} = 0, \varepsilon = 10^{-5}, \\ \omega_1 = \omega_2 = \omega_3 = 1/3, \mu = 0.5, \rho = \max\{2\mu\sqrt{2}, 1\}^2, \\ (\forall n \in \mathbb{N}) \quad \lambda_n = 1, \ \gamma_n = 2/\rho - \varepsilon, \\ (\forall n \in \mathbb{N}) \quad a_{1,n} = 0, a_{2,n} = 0, a_{3,n} = 0. \end{cases} \tag{3.52}$$

On comparera également la solution à celle obtenue pour le problème (3.50) avec l'algorithme (2.141) pour résoudre le problème (2.147) avec les paramètres

$$\begin{cases} N = 256, p = 2, \varepsilon = 10^{-3}, v_0 = 0, \mu = 0.5, \\ (\forall n \in \mathbb{N}) \quad \lambda_n = 1, \tau_n = \mu^{-1}/4 - \varepsilon, b_n = 0, \\ (\forall n \in \mathbb{N})(\forall (k,l) \in \{1, \ldots, N\}^2) \quad \alpha_{n,k,l}^{(1)} = 0, \alpha_{n,k,l}^{(2)} = 0. \end{cases} \tag{3.53}$$

Nous obtenons les résultats présentés dans le tableau suivant et la figure 3.3 :

$n = 100$ itérations	Image bruitée	Méthode (2.141)	Méthode 3.2
Rapport signal-sur-bruit	24.7	28.6	33

On voit que le modèle (3.50) avec deux contraintes nous donne un meilleur résultat que le modèle (2.147) avec une contrainte.

FIGURE 3.1 – L'image originale.

FIGURE 3.2 – L'image bruitée (le bruit blanc de moyenne 0 et SNR $=24.7$ dB).

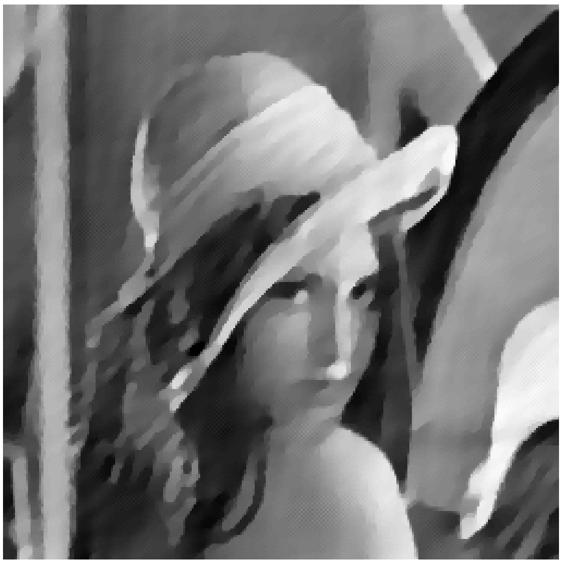

FIGURE 3.3 – L'image débruitée avec la contrainte C_1 (problème (2.147)).

FIGURE 3.4 – L'image débruitée avec deux contraintes C_1 et C_2 (problème (3.50)).

3.3.2 Restauration à partir d'observations multiples

Nous cherchons une image originale $\overline{x} \in \mathbb{R}^{N\times N}$ à partir de deux images dégradées $r_1 \in \mathbb{R}^{m_1}$ et $r_2 \in \mathbb{R}^{m_2}$ selon le modèle

$$r_1 = L_1\overline{x} + s_1 \text{ et } r_2 = L_2\overline{x} + s_2, \tag{3.54}$$

où $L_1 \colon \mathbb{R}^{N\times N} \to \mathbb{R}^{m_1}$ et $L_1 \colon \mathbb{R}^{N\times N} \to \mathbb{R}^{m_2}$ sont deux opérateurs linéaires, $s_1 \in \mathbb{R}^{m_1}$ et $s_2 \in \mathbb{R}^{m_2}$ sont des bruits additifs. Nous utilisons le modèle suivant

$$\underset{x\in C}{\text{minimiser}}\ \omega_1\|L_1x - r_1\| + \omega_2\|L_2x - r_2\| + \frac{\alpha}{2}\|x\|^2, \tag{3.55}$$

où α, ω_1 et ω_2 sont des paramètres strictement positifs, $C \subset \mathbb{R}^{N\times N}$ est un sous-ensemble convexe fermé non vide. Nous utilisons $C = C_2$ (cf. (3.51)), et L_1 et L_2 sont les opérateurs de convolution definis respectivement par $L_1 \colon x \mapsto h_1 * x$ où h_1 est un noyau de taille 10×10, et $L_2 \colon x \mapsto h_2 * x$ où h_2 est un noyau de taille 15×15. De plus, s_1 et s_2 sont des réalisations d'un bruit blanc de moyenne 0 et de variance 0.001. Nous utilisons l'Algorithme 3.2 avec les paramètres

$$\omega_2 = \omega_3 = 1/3, N = 256, \alpha = 10^{-5}, \gamma_n \equiv 1.99, \lambda_n \equiv 1, a_{1,n} \equiv 0, a_{2,n} \equiv 0, a_{3,n} \equiv 0. \tag{3.56}$$

Nous obtenons les résultats dans le tableau suivant et les figures 3.5, 3.6 et 3.7 :

$n = 40$ itérations	L'image dégradée 1	L'image dégradée 2	Résultat
Rapport signal-sur-bruit	22.85	20.63	28.2

FIGURE 3.5 – L'image observée 1.

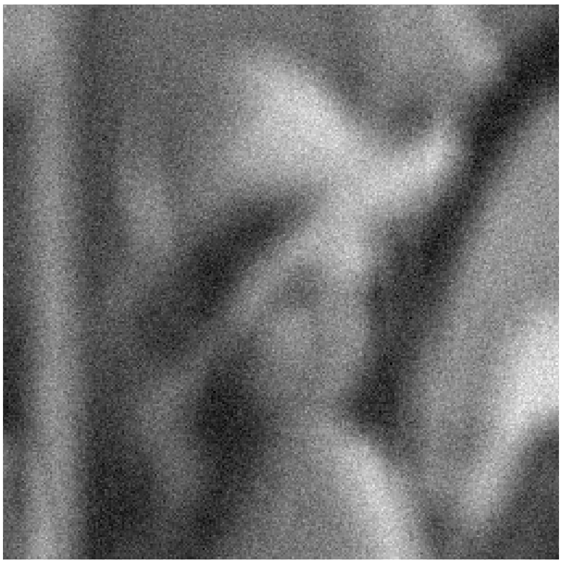

FIGURE 3.6 – L'image observée 2.

FIGURE 3.7 – L'image restaurée.

3.4 Bibliographie

[1] P. Alart, O. Maisonneuve, and R. T. Rockafellar (Eds.), *Nonsmooth Mechanics and Analysis – Theoretical and Numerical Advances.* Springer-Verlag, New York, 2006.

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

[3] A. Chambolle, Total variation minimization and a class of binary MRF model, *Lecture Notes in Comput. Sci.,* vol. 3757, pp 136–152, 2005.

[4] P. L. Combettes, Iterative construction of the resolvent of a sum of maximal monotone operators, *J. Convex Anal.,* vol. 16, pp. 727–748, 2009.

[5] P. L. Combettes, Đinh Dũng, and B. C. Vũ, Dualization of signal recovery problems, *Set-Valued Var. Anal.*, vol. 18, pp. 373–404, 2010.

[6] P. L. Combettes and J.-C. Pesquet, Proximal thresholding algorithm for minimization over orthonormal bases, *SIAM J. Optim.,* vol. 18, pp. 1351–1376, 2007.

[7] P. L. Combettes and J.-C. Pesquet, Proximal splitting methods in signal processing, in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering,* (H. H. Bauschke et al., eds), pp. 185–212. Springer, New York, 2011.

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

[9] C. De Mol, E. De Vito, and L. Rosasco, Elastic-net regularization in learning theory, *J. Complexity,* vol. 25, pp. 201–230, 2009.

[10] F. Deutsch, *Best Approximation in Inner Product Spaces.* Springer-Verlag, New York, 2001.

[11] N. Gaffke and R. Mathar, A cyclic projection algorithm via duality, *Metrika,* vol. 36, pp. 29–54, 1989.

[12] B. Mercier, *Inéquations Variationnelles de la Mécanique* (Publications Mathématiques d'Orsay, no. 80.01). Orsay, France, Université de Paris-XI, 1980.

[13] J.-J. Moreau, Fonctions convexes duales et points proximaux dans un espace hilbertien, *C. R. Acad. Sci. Paris Sér. A Math.,* vol. 255, pp. 2897–2899, 1962.

[14] J.-J. Moreau, Proximité et dualité dans un espace hilbertien, *Bull. Soc. Math. France,* vol. 93, pp. 273-299, 1965.

[15] L. C. Potter and K. S. Arun, A dual approach to linear inverse problems with convex constraints, *SIAM J. Control Optim.,* vol. 31, pp. 1080–1092, 1993.

[16] R. T. Rockafellar and R. J. B. Wets, *Variational Analysis,* 3rd printing. Springer-Verlag, New York, 2009.

[17] L. I. Rudin, S. Osher, and E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D,* vol. 60, pp. 259–268, 1992.

[18] P. Tseng, Applications of a splitting algorithm to decomposition in convex programming and variational inequalities, *SIAM J. Control Optim.,* vol. 29, pp. 119–138, 1991.

[19] C. Zălinescu, *Convex Analysis in General Vector Spaces,* World Scientific, River Edge, NJ, 2002.

[20] H. Zou and T. Hastie, Regularization and variable selection via the elastic net, *J. R. Stat. Soc. Ser. B Stat. Methodol.,* vol. 67, pp. 301–320, 2005.

Chapitre 4

Résolution d'inclusions monotones impliquant des opérateurs cocoercifs

Nous proposons un cadre général pour résoudre des inclusions composites faisant intérieur des opérateurs maximalement monotones et des opérateurs cocoercifs. La méthode proposée unifie les méthodes primales-duales de [10, 18, 22].

4.1 Description et résultats principaux

Nous nous intéressons à la résolution d'inclusions monotones associées des opérateurs cocoercifs dans des espaces hilbertiens réels.

Problème 4.1 Soient $\mathcal{H}$ un espace hilbertien réel, $z \in \mathcal{H}$, m un entier strictement positif, $(\omega_i)_{1\leqslant i\leqslant m}$ des nombres réels dans $]0,1]$ tels que $\sum_{i=1}^m \omega_i = 1$, $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ un opérateur maximalement monotone, $\mu \in\,]0,+\infty[$, et $C\colon \mathcal{H} \to \mathcal{H}$ un opérateur μ-cocoercif. Pour tout $i \in \{1,\ldots,m\}$, soient $\mathcal{G}_i$ un espace hilbertien réel, $r_i \in \mathcal{G}_i$, $\nu_i \in\,]0,+\infty[$, $B_i\colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ un opérateur maximalement monotone, $D_i\colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ un opérateur maximalement monotone et ν_i-fortement monotone, et $0 \neq L_i \in \mathcal{B}(\mathcal{H},\mathcal{G}_i)$. Le problème est de résoudre l'inclusion primale

$$\text{trouver } \overline{x} \in \mathcal{H} \text{ tel que } z \in A\overline{x} + \sum_{i=1}^m \omega_i L_i^*\big((B_i \,\square\, D_i)(L_i\overline{x} - r_i)\big) + C\overline{x}, \tag{4.1}$$

et l'inclusion duale

$$\text{trouver } \overline{v}_1 \in \mathcal{G}_1,\ldots,\overline{v}_m \in \mathcal{G}_m \text{ tels que}$$
$$(\exists\, x \in \mathcal{H}) \quad \begin{cases} z - \sum_{i=1}^m \omega_i L_i^*\overline{v}_i \in Ax + Cx \\ \big(\forall i \in \{1,\ldots,m\}\big)\ \overline{v}_i \in (B_i \,\square\, D_i)(L_i x - r_i). \end{cases} \tag{4.2}$$

Nous notons par $\mathcal{P}$ et $\mathcal{D}$ des ensembles de solutions de problèmes (4.1) et (4.2), respectivement.

Plusieurs cas particuliers du Problème 4.1 sont présentés dans [16]. Dans le cas où C et $(D_i)_{1\leqslant i\leqslant m}$ sont lipschitziens, le Problème 4.1 est étudié dans [16] mais ses auteurs n'exploitent pas de cocoercivité de ces opérateurs. En utilisant la méthode explicite-implicite dans la somme hilbertienne directe $\mathcal{K} = \mathcal{H} \oplus \mathcal{G}_1 \oplus \ldots \oplus \mathcal{G}_m$, nous obtenons le résultat suivant.

Théorème 4.2 *Dans le Problème 4.1, supposons que*

$$z \in \operatorname{ran}\Big(A + \sum_{i=1}^{m} \omega_i L_i^*\big((B_i \,\square\, D_i)(L_i \cdot - r_i)\big) + C\Big). \tag{4.3}$$

Soient τ et $(\sigma_i)_{1\leqslant i\leqslant m}$ des réels strictement positifs tels que

$$2\rho \min\{\mu, \nu_1, \ldots, \nu_m\} > 1, \text{où } \rho = \min\Big\{\tau^{-1}, \sigma_1^{-1}, \ldots, \sigma_m^{-1}\Big\}\Bigg(1 - \sqrt{\tau \sum_{i=1}^{m} \sigma_i \omega_i \|L_i\|^2}\Bigg). \tag{4.4}$$

Soient $\varepsilon \in \,]0,1[$, $(\lambda_n)_{n\in\mathbb{N}}$ une suite dans $[\varepsilon, 1]$, $x_0 \in \mathcal{H}$, $(a_{1,n})_{n\in\mathbb{N}}$ et $(a_{2,n})_{n\in\mathbb{N}}$ des suites absolument sommables dans $\mathcal{H}$. Pour tout $i \in \{1, \ldots, m\}$, soient $v_{i,0} \in \mathcal{G}_i$, $(b_{i,n})_{n\in\mathbb{N}}$ et $(c_{i,n})_{n\in\mathbb{N}}$ des suites absolument sommables dans $\mathcal{G}_i$. Soient $(x_n)_{n\in\mathbb{N}}$ et $(v_{1,n}, \ldots, v_{m,n})_{n\in\mathbb{N}}$ des suites engendrées comme suit.

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} p_n = J_{\tau A}\Big(x_n - \tau\Big(\sum_{i=1}^{m} \omega_i L_i^* v_{i,n} + C x_n + a_{1,n} - z\Big)\Big) + a_{2,n} \\ y_n = 2p_n - x_n \\ x_{n+1} = x_n + \lambda_n(p_n - x_n) \\ \text{pour } i = 1, \ldots, m \\ \left\lfloor \begin{array}{l} q_{i,n} = J_{\sigma_i B_i^{-1}}\Big(v_{i,n} + \sigma_i\Big(L_i y_n - D_i^{-1} v_{i,n} - c_{i,n} - r_i\Big)\Big) + b_{i,n} \\ v_{i,n+1} = v_{i,n} + \lambda_n(q_{i,n} - v_{i,n}). \end{array} \right. \end{array} \right. \tag{4.5}$$

Alors pour un point $\overline{x} \in \mathcal{P}$ et un point$(\overline{v}_1, \ldots, \overline{v}_m) \in \mathcal{D}$, nous avons ce qui suit.

(i) $x_n \rightharpoonup \overline{x}$ *et* $(v_{1,n}, \ldots, v_{m,n}) \rightharpoonup (\overline{v}_1, \ldots, \overline{v}_m)$.

(ii) *Supposons que C soit uniformément monotone en $\overline{x}$. Alors $x_n \to \overline{x}$.*

(iii) *Supposons que D_j^{-1} soit uniformément monotone en $\overline{v}_j$ pour un indice $j \in \{1, \ldots, m\}$. Alors $v_{j,n} \to \overline{v}_j$.*

Dans cette méthode, on obtient un éclatement de tous les opérateurs puisque $(L_i)_{1\leqslant i\leqslant m}$, C, $(D_i^{-1})_{1\leqslant i\leqslant m}$, A et $(B_i^{-1})_{1\leqslant i\leqslant m}$ sont utilisés individuellement à chaque itération. De plus, la méthode tolère des erreurs dans l'évaluation de chaque opérateur impliqué. On voit que les opérateurs univoques tels que $(L_i)_{1\leqslant i\leqslant m}$, $(L_i^*)_{1\leqslant i\leqslant m}$, C, $(D_i^{-1})_{1\leqslant i\leqslant m}$

sont utilisés dans des étapes explicites, pourtant les opérateurs multivoques comme A et $(B_i^{-1})_{1\leqslant i\leqslant m}$ sont utilisés dans des étapes implicites, c'est à dire que (4.5) admet une structure de la méthode explicite-implicite. Donce, il diffère de la méthode de Combettes-Pesquet dans [16].

Nous appliquons l'algorithme (4.5) au problème variationnel [16, Problem 4.1].

Problème 4.3 Soient $\mathcal{H}$ un espace hilbertien réel, $z \in \mathcal{H}$, m un entier strictement positif, $(\omega_i)_{1\leqslant i\leqslant m}$ des nombres réels dans $]0,1]$ tels que $\sum_{i=1}^m \omega_i = 1$, $f \in \Gamma_0(\mathcal{H})$, et $h\colon \mathcal{H} \to \mathbb{R}$ une fonction convexe différentiable telle que son gradient est μ^{-1}-lipschitzien avec $\mu \in]0,+\infty[$. Pour tout $i \in \{1,\ldots,m\}$, soient $\mathcal{G}_i$ un espace hilbertien réel, $r_i \in \mathcal{G}_i$, $g_i \in \Gamma_0(\mathcal{G}_i)$, $\ell_i \in \Gamma_0(\mathcal{G}_i)$ une fonction ν_i-fortement convexe avec $\nu_i \in]0,+\infty[$, et $0 \neq L_i\colon \mathcal{H} \to \mathcal{G}_i$. Le problème primal est de

$$\underset{x\in\mathcal{H}}{\text{minimiser}}\ f(x) + \sum_{i=1}^m \omega_i (g_i \,\square\, \ell_i)(L_i x - r_i) + h(x) - \langle x \mid z\rangle, \tag{4.6}$$

et le problème dual est de

$$\underset{v_1\in\mathcal{G}_1,\ldots,v_m\in\mathcal{G}_m}{\text{minimiser}}\ (f^* \,\square\, h^*)\Big(z - \sum_{i=1}^m \omega_i L_i^* v_i\Big) + \sum_{i=1}^m \omega_i \big(g_i^*(v_i) + \ell_i^*(v_i) + \langle v_i \mid r_i\rangle_{\mathcal{G}_i}\big). \tag{4.7}$$

Nous notons par $\mathcal{P}_1$ et $\mathcal{D}_1$ des ensembles de solutions de problèmes (4.6) et (4.7), respectivement.

Corollaire 4.4 *Dans le Problème 4.3, supposons que*

$$z \in \operatorname{ran}\Big(\partial f + \sum_{i=1}^m \omega_i L_i^* \big((\partial g_i \,\square\, \partial \ell_i)(L_i \cdot - r_i)\big) + \nabla h\Big). \tag{4.8}$$

Soient τ et $(\sigma_i)_{1\leqslant i\leqslant m}$ des nombres réels strictement positifs tels que

$$2\rho \min\{\mu,\nu_1,\ldots,\nu_m\} > 1, \text{où } \rho = \min\Big\{\tau^{-1},\sigma_1^{-1},\ldots,\sigma_m^{-1}\Big\}\Big(1 - \sqrt{\tau\sum_{i=1}^m \sigma_i\omega_i\|L_i\|^2}\Big). \tag{4.9}$$

Soient $\varepsilon \in]0,1[$, $(\lambda_n)_{n\in\mathbb{N}}$ une suite dans $[\varepsilon,1]$, $x_0 \in \mathcal{H}$, $(a_{1,n})_{n\in\mathbb{N}}$ et $(a_{2,n})_{n\in\mathbb{N}}$ des suites absolument sommables dans $\mathcal{H}$. Pour tout $i \in \{1,\ldots,m\}$, soient $v_{i,0} \in \mathcal{G}_i$, $(b_{i,n})_{n\in\mathbb{N}}$ et $(c_{i,n})_{n\in\mathbb{N}}$ des suites absolument sommables dans $\mathcal{G}_i$. Soient $(x_n)_{n\in\mathbb{N}}$ et $(v_{1,n},\ldots,v_{m,n})_{n\in\mathbb{N}}$ des *suites engendrées comme suit.*

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} p_n = \operatorname{prox}_{\tau f}\Big(x_n - \tau\Big(\sum_{i=1}^m \omega_i L_i^* v_{i,n} + \nabla h(x_n) + a_{1,n} - z\Big)\Big) + a_{2,n} \\ y_n = 2p_n - x_n \\ x_{n+1} = x_n + \lambda_n(p_n - x_n) \\ \text{pour } i = 1, \ldots, m \\ \left\lfloor \begin{array}{l} q_{i,n} = \operatorname{prox}_{\sigma_i g_i^*}\Big(v_{i,n} + \sigma_i\Big(L_i y_n - \nabla \ell_i^*(v_{i,n}) + c_{i,n} - r_i\Big)\Big) + b_{i,n} \\ v_{i,n+1} = v_{i,n} + \lambda_n(q_{i,n} - v_{i,n}). \end{array} \right. \end{array} \right. \tag{4.10}$$

Alors pour un point $\overline{x} \in \mathcal{P}_1$ et un point $(\overline{v}_1, \ldots, \overline{v}_m) \in \mathcal{D}_1$, nous avons ce qui suit.

(i) $x_n \rightharpoonup \overline{x}$ *et* $(v_{1,n}, \ldots, v_{m,n}) \rightharpoonup (\overline{v}_1, \ldots, \overline{v}_m)$.

(ii) *Supposons que h soit uniformément convexe en $\overline{x}$. Alors $x_n \to \overline{x}$.*

(iii) *Supposons que ℓ_j^* soit uniformément convexe en $\overline{v}_j$ pour un indice $j \in \{1, \ldots, m\}$. Alors $v_{j,n} \to \overline{v}_j$.*

Des liens avec des méthodes existantes et des cas particuliers de ces résultats sont présentés dans la Section 4.2.

4.2 Article en anglais

A SPLITTING ALGORITHM FOR DUAL MONOTONE INCLUSIONS INVOLVING COCOERCIVE OPERATORS [1]

Abstract : We consider the problem of solving dual monotone inclusions involving sums of composite parallel-sum type operators. A feature of this work is to exploit explicitly the cocoercivity of some of the operators appearing in the model. Several splitting algorithms recently proposed in the literature are recovered as special cases.

4.2.1 Introduction

Monotone operator splitting methods have found many applications in applied mathematics, e.g., evolution inclusions [2], partial differential equations [1, 20, 23], mechanics [21], variational inequalities [6, 19], Nash equilibria [8], and various optimization problems [7, 9, 10, 14, 15, 17, 25, 29]. In such formulations, cocoercivity often plays a central role ; see for instance [2, 6, 11, 13, 19, 20, 21, 23, 28, 29, 30]. Recall that

1. B. C. Vũ, A splitting algorithm for dual monotone inclusions involving cocoercive operators, *Adv. Comput. Math.*, vol. 38, pp. 667–681, 2013.

an operator $C\colon \mathcal{H} \to \mathcal{H}$ is cocoercive with constant $\beta \in \left]0, +\infty\right[$ if its inverse is β-strongly monotone, that is,

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \langle x - y \mid Cx - Cy \rangle \geqslant \beta \|Cx - Cy\|^2. \tag{4.11}$$

In this paper, we revisit a general primal-dual splitting framework proposed in [16] in the presence Lipschitzian operators in the context of cocoercive operators. This will lead to a new type of splitting technique and provide a unifying framework for some algorithms recently proposed in the literature. The problem under investigation is the following, where the parallel sum operation is denoted by $\Box$ (see (4.22)).

Problem 4.5 Let $\mathcal{H}$ be a real Hilbert space, let $z \in \mathcal{H}$, let m be a strictly positive integer, let $(\omega_i)_{1\leqslant i\leqslant m}$ be real numbers in $\left]0, 1\right]$ such that $\sum_{i=1}^m \omega_i = 1$, let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone, and let $C\colon \mathcal{H} \to \mathcal{H}$ be μ-cocoercive for some $\mu \in \left]0, +\infty\right[$. For every $i \in \{1, \ldots, m\}$, let $\mathcal{G}_i$ be a real Hilbert space, let $r_i \in \mathcal{G}_i$, let $B_i\colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ be maximally monotone, let $D_i\colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ be maximally monotone and ν_i-strongly monotone for some $\nu_i \in \left]0, +\infty\right[$, and suppose that $L_i\colon \mathcal{H} \to \mathcal{G}_i$ is a nonzero bounded linear operator. The problem is to solve the primal inclusion

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } z \in A\overline{x} + \sum_{i=1}^{m} \omega_i L_i^*\big((B_i \,\Box\, D_i)(L_i\overline{x} - r_i)\big) + C\overline{x}, \tag{4.12}$$

together with the dual inclusion

$$\begin{aligned}&\text{find } \overline{v}_1 \in \mathcal{G}_1, \ldots, \overline{v}_m \in \mathcal{G}_m \text{ such that}\\ &\qquad (\exists\, x \in \mathcal{H}) \begin{cases} z - \sum_{i=1}^m \omega_i L_i^*\overline{v}_i \in Ax + Cx \\ \big(\forall i \in \{1, \ldots, m\}\big)\ \overline{v}_i \in (B_i \,\Box\, D_i)(L_i x - r_i). \end{cases}\end{aligned} \tag{4.13}$$

We denote by $\mathcal{P}$ and $\mathcal{D}$ the sets of solutions to (4.12) and (4.13), respectively.

In the case when $(D_i^{-1})_{1\leqslant i\leqslant m}$ and C are general monotone Lipschitzian operators, Problem 4.5 was investigated in [16]. Here are a couple of special cases of Problem 4.5.

Example 4.6 In Problem 4.5, set $z = 0$ and

$$\big(\forall i \in \{1, \ldots, m\}\big) \quad B_i\colon v \mapsto \{0\} \quad \text{and} \quad D_i\colon v \mapsto \{0\}. \tag{4.14}$$

The primal inclusion (4.12) reduces to

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } 0 \in A\overline{x} + C\overline{x}. \tag{4.15}$$

This problem is studied in [2, 11, 13, 17, 23, 28, 29].

Example 4.7 Suppose that in Problem 4.5,

$$A\colon x \mapsto \{0\}, \quad C\colon x \mapsto 0, \quad \text{and} \quad \big(\forall i \in \{1,\ldots,m\}\big) \quad D_i\colon v \mapsto \begin{cases} \mathcal{G}_i & \text{if } v = 0, \\ \varnothing & \text{if } v \neq 0. \end{cases} \tag{4.16}$$

Then we obtain the primal-dual pair

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } z \in \sum_{i=1}^{m} \omega_i L_i^* \big(B_i(L_i\overline{x} - r_i)\big), \tag{4.17}$$

and

$$\text{find } \overline{v}_1 \in \mathcal{G}_1, \ldots, \overline{v}_m \in \mathcal{G}_m \text{ such that } \begin{cases} \sum_{i=1}^{m} \omega_i L_i^* \overline{v}_i = z, \\ (\exists\, x \in \mathcal{H})\big(\forall i \in \{1,\ldots,m\}\big)\ \overline{v}_i \in B_i(L_i x - r_i). \end{cases} \tag{4.18}$$

This framework is considered in [7], where further special cases will be found. In particular, it contains the classical Fenchel-Rockafellar [27] and Mosco [24] duality settings, as well as that of [3].

The paper is organized as follows. Section 4.2.2 is devoted to notation and background. In Section 4.2.3, we present our algorithm, prove its convergence, and compare it to existing work. Applications to minimization problems are provided in Section 4.2.4, where further connections with the state-of-the-art are made.

4.2.2 Notation and background

We recall some notation and background from convex analysis and monotone operator theory (see [6] for a detailed account).

Throughout, $\mathcal{H}$, $\mathcal{G}$, and $(\mathcal{G}_i)_{1\leqslant i\leqslant m}$ are real Hilbert spaces. The scalars product and the associated norms of both $\mathcal{H}$ and $\mathcal{G}$ are denoted respectively by $\langle \cdot \mid \cdot \rangle$ and $\|\cdot\|$. For every $i \in \{1,\ldots,m\}$, the scalar product and associated norm of $\mathcal{G}_i$ are denoted respectively by $\langle \cdot \mid \cdot \rangle_{\mathcal{G}_i}$ and $\|\cdot\|_{\mathcal{G}_i}$. We denote by $\mathcal{B}(\mathcal{H},\mathcal{G})$ the space of all bounded linear operators from $\mathcal{H}$ to $\mathcal{G}$. The symbols $\rightharpoonup$ and $\to$ denote respectively weak and strong convergence. Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a set-valued operator. The domain and the graph of A are respectively defined by $\operatorname{dom} A = \big\{x \in \mathcal{H} \bigm| Ax \neq \varnothing\big\}$ and $\operatorname{gra} A = \big\{(x,u) \in \mathcal{H} \times \mathcal{H} \bigm| u \in Ax\big\}$. We denote by $\operatorname{zer} A = \big\{x \in \mathcal{H} \bigm| 0 \in Ax\big\}$ the set of zeros of A, and by $\operatorname{ran} A = \big\{u \in \mathcal{H} \bigm| (\exists\, x \in \mathcal{H})\ u \in Ax\big\}$ the range of A. The inverse of A is $A^{-1}\colon \mathcal{H} \mapsto 2^{\mathcal{H}}\colon u \mapsto \big\{x \in \mathcal{H} \bigm| u \in Ax\big\}$. The resolvent of A is

$$J_A = (\operatorname{Id} + A)^{-1}, \tag{4.19}$$

where Id denotes the identity operator on $\mathcal{H}$. Moreover, A is monotone if

$$(\forall(x,y)\in\mathcal{H}\times\mathcal{H})\;(\forall(u,v)\in Ax\times Ay)\quad \langle x-y\mid u-v\rangle\geqslant 0, \tag{4.20}$$

and maximally monotone if it is monotone and there exists no monotone operator $B\colon\mathcal{H}\to 2^{\mathcal{H}}$ such that gra$B$ properly contains gra$A$. We say that A is uniformly monotone at $x\in\operatorname{dom}A$ if there exists an increasing function $\phi\colon\,[0,+\infty[\,\to[0,+\infty]$ vanishing only at 0 such that

$$\big(\forall u\in Ax\big)\big(\forall(y,v)\in \operatorname{gra}A\big)\quad \langle x-y\mid u-v\rangle\geqslant\phi(\|x-y\|). \tag{4.21}$$

If $A-\alpha\,\mathrm{Id}$ is monotone for some $\alpha\in\,]0,+\infty[$, then A is said to be α-strongly monotone. The parallel sum of two set-valued operators A and B from $\mathcal{H}$ to $2^{\mathcal{H}}$ is

$$A\,\square\,B=(A^{-1}+B^{-1})^{-1}. \tag{4.22}$$

The class of all lower semicontinuous convex functions $f\colon\mathcal{H}\to\,]-\infty,+\infty]$ such that $\operatorname{dom}f=\big\{x\in\mathcal{H}\;\big|\;f(x)<+\infty\big\}\neq\varnothing$ is denoted by $\Gamma_0(\mathcal{H})$. Now, let $f\in\Gamma_0(\mathcal{H})$. The conjugate of f is the function $f^*\in\Gamma_0(\mathcal{H})$ defined by $f^*\colon u\mapsto\sup_{x\in\mathcal{H}}(\langle x\mid u\rangle-f(x))$, and the subdifferential of $f\in\Gamma_0(\mathcal{H})$ is the maximally monotone operator

$$\partial f\colon\mathcal{H}\to 2^{\mathcal{H}}\colon x\mapsto\big\{u\in\mathcal{H}\;\big|\;(\forall y\in\mathcal{H})\quad\langle y-x\mid u\rangle+f(x)\leqslant f(y)\big\} \tag{4.23}$$

with inverse given by

$$(\partial f)^{-1}=\partial f^*. \tag{4.24}$$

Moreover, the proximity operator of f is

$$\operatorname{prox}_f\colon\mathcal{H}\to\mathcal{H}\colon x\mapsto\underset{y\in\mathcal{H}}{\operatorname{argmin}}\;f(y)+\frac{1}{2}\|x-y\|^2. \tag{4.25}$$

We have

$$J_{\partial f}=\operatorname{prox}_f. \tag{4.26}$$

The infimal convolution of two functions f and g from $\mathcal{H}$ to $]-\infty,+\infty]$ is

$$f\,\square\,g\colon\mathcal{H}\to\,]-\infty,+\infty]\,:x\mapsto\inf_{y\in\mathcal{H}}(f(x)+g(x-y)). \tag{4.27}$$

Finally, the relative interior of a subset S of $\mathcal{H}$, i.e., the set of points $x\in S$ such that the cone generated by $-x+S$ is a vector subspace of $\mathcal{H}$, is denoted by $\operatorname{ri}S$.

4.2.3 Algorithm and convergence

Our main result is the following theorem, in which we introduce our splitting algorithm and prove its convergence.

Theorem 4.8 *In Problem 4.5, suppose that*

$$z \in \operatorname{ran}\Big(A + \sum_{i=1}^{m} \omega_i L_i^*\big((B_i \,\square\, D_i)(L_i \cdot - r_i)\big) + C\Big). \tag{4.28}$$

Let τ and $(\sigma_i)_{1\leqslant i\leqslant m}$ be strictly positive numbers such that

$$2\rho \min\{\mu, \nu_1, \ldots, \nu_m\} > 1, \text{where } \rho = \min\Big\{\tau^{-1}, \sigma_1^{-1}, \ldots, \sigma_m^{-1}\Big\}\Bigg(1 - \sqrt{\tau \sum_{i=1}^{m} \sigma_i \omega_i \|L_i\|^2}\Bigg). \tag{4.29}$$

Let $\varepsilon \in \,]0,1[$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon,1]$, let $x_0 \in \mathcal{H}$, let $(a_{1,n})_{n\in\mathbb{N}}$ and $(a_{2,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{H}$. For every $i \in \{1,\ldots,m\}$, let $v_{i,0} \in \mathcal{G}_i$, and let $(b_{i,n})_{n\in\mathbb{N}}$ and $(c_{i,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{G}_i$. Let $(x_n)_{n\in\mathbb{N}}$ and $(v_{1,n},\ldots,v_{m,n})_{n\in\mathbb{N}}$ be sequences generated by the following routine

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} p_n = J_{\tau A}\Big(x_n - \tau\Big(\sum_{i=1}^{m} \omega_i L_i^* v_{i,n} + C x_n + a_{1,n} - z\Big)\Big) + a_{2,n} \\ y_n = 2p_n - x_n \\ x_{n+1} = x_n + \lambda_n(p_n - x_n) \\ \text{for } i = 1,\ldots,m \\ \left\lfloor \begin{array}{l} q_{i,n} = J_{\sigma_i B_i^{-1}}\Big(v_{i,n} + \sigma_i\Big(L_i y_n - D_i^{-1} v_{i,n} - c_{i,n} - r_i\Big)\Big) + b_{i,n} \\ v_{i,n+1} = v_{i,n} + \lambda_n(q_{i,n} - v_{i,n}). \end{array} \right. \end{array} \right. \tag{4.30}$$

Then the following hold for some $\overline{x} \in \mathcal{P}$ and $(\overline{v}_1,\ldots,\overline{v}_m) \in \mathcal{D}$.

(i) $x_n \rightharpoonup \overline{x}$ *and* $(v_{1,n},\ldots,v_{m,n}) \rightharpoonup (\overline{v}_1,\ldots,\overline{v}_m)$.

(ii) *Suppose that C is uniformly monotone at $\overline{x}$. Then $x_n \to \overline{x}$.*

(iii) *Suppose that D_j^{-1} is uniformly monotone at $\overline{v}_j$ for some $j \in \{1,\ldots,m\}$. Then $v_{j,n} \to \overline{v}_j$.*

Proof. We define $\boldsymbol{\mathcal{G}}$ as the real Hilbert space obtained by endowing the Cartesian product $\mathcal{G}_1 \times \ldots \times \mathcal{G}_m$ with the scalar product and the associated norm respectively defined by

$$\langle \cdot \mid \cdot \rangle_{\boldsymbol{\mathcal{G}}} \colon (\boldsymbol{v}, \boldsymbol{w}) \mapsto \sum_{i=1}^{m} \omega_i \langle v_i \mid w_i \rangle_{\mathcal{G}_i} \quad \text{and} \quad \|\cdot\|_{\boldsymbol{\mathcal{G}}} \colon \boldsymbol{v} \mapsto \sqrt{\sum_{i=1}^{m} \omega_i \|v_i\|_{\mathcal{G}_i}^2}, \tag{4.31}$$

where $\boldsymbol{v} = (v_1, \dots, v_m)$ and $\boldsymbol{w} = (w_1, \dots, w_m)$ denote generic elements in $\boldsymbol{\mathcal{G}}$. Next, we let $\boldsymbol{\mathcal{K}}$ be the Hilbert direct sum

$$\boldsymbol{\mathcal{K}} = \mathcal{H} \oplus \boldsymbol{\mathcal{G}}. \tag{4.32}$$

Thus, the scalar product and the norm of $\boldsymbol{\mathcal{K}}$ are respectively defined by

$$\langle \cdot \mid \cdot \rangle_{\boldsymbol{\mathcal{K}}} \colon \big((x, \boldsymbol{v}), (y, \boldsymbol{w})\big) \mapsto \langle x \mid y \rangle + \langle \boldsymbol{v} \mid \boldsymbol{w} \rangle_{\boldsymbol{\mathcal{G}}} \quad \text{and} \quad \| \cdot \|_{\boldsymbol{\mathcal{K}}} \colon (x, \boldsymbol{v}) \mapsto \sqrt{\|x\|^2 + \|\boldsymbol{v}\|^2_{\boldsymbol{\mathcal{G}}}}. \tag{4.33}$$

Let us set

$$\begin{aligned} \boldsymbol{M} \colon \boldsymbol{\mathcal{K}} &\to 2^{\boldsymbol{\mathcal{K}}} \\ (x, v_1, \dots, v_m) &\mapsto \big(-z + Ax\big) \times \big(r_1 + B_1^{-1} v_1\big) \times \dots \times \big(r_m + B_m^{-1} v_m\big). \end{aligned} \tag{4.34}$$

Since the operators A and $(B_i)_{1 \leqslant i \leqslant m}$ are maximally monotone, $\boldsymbol{M}$ is maximally monotone [6, Propositions 20.22 and 20.23]. We also introduce

$$\boldsymbol{S} \colon \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}} \tag{4.35}$$

$$(x, v_1, \dots, v_m) \mapsto \left(\sum_{i=1}^m \omega_i L_i^* v_i, -L_1 x, \dots, -L_m x \right). \tag{4.36}$$

Note that $\boldsymbol{S}$ is linear, bounded, and skew (i.e, $\boldsymbol{S}^* = -\boldsymbol{S}$). Hence, $\boldsymbol{S}$ is maximally monotone [6, Example 20.30]. Moreover, since $\operatorname{dom} \boldsymbol{S} = \boldsymbol{\mathcal{K}}$, $\boldsymbol{M} + \boldsymbol{S}$ is maximally monotone [6, Corollary 24.24(i)]. Since, for every $i \in \{1, \dots, m\}$, D_i is ν_i-strongly monotone, D_i^{-1} is ν_i-cocoercive. Let us prove that

$$\begin{aligned} \boldsymbol{Q} \colon \boldsymbol{\mathcal{K}} &\to \boldsymbol{\mathcal{K}} \\ (x, v_1, \dots, v_m) &\mapsto \big(Cx, D_1^{-1} v_1, \dots, D_m^{-1} v_m\big) \end{aligned} \tag{4.37}$$

is β-cocoercive with

$$\beta = \min\{\mu, \nu_1, \dots, \nu_m\}. \tag{4.38}$$

For every $(x, v_1, \dots, v_m)$ and every $(y, w_1, \dots, w_m)$ in $\boldsymbol{\mathcal{K}}$, we have

$$\begin{aligned} &\langle (x, v_1, \dots, v_m) - (y, w_1, \dots, w_m) \mid \boldsymbol{Q}(x, v_1, \dots, v_m) - \boldsymbol{Q}(y, w_1, \dots, w_m) \rangle_{\boldsymbol{\mathcal{K}}} \\ &= \langle x - y \mid Cx - Cy \rangle + \sum_{i=1}^m \omega_i \big\langle v_i - w_i \mid D_i^{-1} v_i - D_i^{-1} w_i \big\rangle_{\mathcal{G}_i} \\ &\geqslant \mu \|Cx - Cy\|^2 + \sum_{i=1}^m \nu_i \omega_i \|D_i^{-1} v_i - D_i^{-1} w_i\|^2_{\mathcal{G}_i} \\ &\geqslant \beta \left(\|Cx - Cy\|^2 + \sum_{i=1}^m \omega_i \|D_i^{-1} v_i - D_i^{-1} w_i\|^2_{\mathcal{G}_i} \right) \\ &= \beta \|\boldsymbol{Q}(x, v_1, \dots, v_m) - \boldsymbol{Q}(y, w_1, \dots, w_m)\|^2_{\boldsymbol{\mathcal{K}}}. \end{aligned} \tag{4.39}$$

Therefore, by (4.11), $\boldsymbol{Q}$ is β-cocoercive. It is shown in [16, Eq. (3.12)] that under the condition (4.28), $\text{zer}(\boldsymbol{M}+\boldsymbol{S}+\boldsymbol{Q}) \neq \varnothing$. Moreover, [16, Eq. (3.21)] and [16, Eq. (3.22)] yield

$$(\overline{x}, \overline{\boldsymbol{v}}) \in \text{zer}(\boldsymbol{M}+\boldsymbol{S}+\boldsymbol{Q}) \Rightarrow \overline{x} \in \mathcal{P} \quad \text{and} \quad \overline{\boldsymbol{v}} \in \mathcal{D}. \tag{4.40}$$

Now, define

$$\begin{aligned} \boldsymbol{V} \colon \boldsymbol{\mathcal{K}} &\to \boldsymbol{\mathcal{K}} \\ (x, v_1, \ldots, v_m) &\mapsto \left(\tau^{-1}x - \sum_{i=1}^{m} \omega_i L_i^* v_i, \sigma_1^{-1} v_1 - L_1 x, \ldots, \sigma_m^{-1} v_m - L_m x\right). \end{aligned} \tag{4.41}$$

Then $\boldsymbol{V}$ is self-adjoint. Let us check that $\boldsymbol{V}$ is ρ-strongly positive. To this end, define

$$\boldsymbol{T} \colon \mathcal{H} \to \boldsymbol{\mathcal{G}} \colon x \mapsto \left(\sqrt{\sigma_1} L_1 x, \ldots, \sqrt{\sigma_m} L_m x\right). \tag{4.42}$$

Then

$$(\forall x \in \mathcal{H}) \quad \|\boldsymbol{T}x\|^2_{\boldsymbol{\mathcal{G}}} = \sum_{i=1}^{m} \omega_i \sigma_i \|L_i x\|^2_{\mathcal{G}_i} \leqslant \|x\|^2 \sum_{i=1}^{m} \omega_i \sigma_i \|L_i\|^2, \tag{4.43}$$

which implies that

$$\|\boldsymbol{T}\|^2 \leqslant \sum_{i=1}^{m} \omega_i \sigma_i \|L_i\|^2. \tag{4.44}$$

Now set

$$\delta = \left(\sqrt{\tau \sum_{i=1}^{m} \sigma_i \omega_i \|L_i\|^2}\right)^{-1} - 1. \tag{4.45}$$

Then it follows from (4.29) that $\delta > 0$. Moreover, (4.44) and (4.45) yield

$$\tau \|\boldsymbol{T}\|^2 (1+\delta) \leqslant \tau(1+\delta) \sum_{i=1}^{m} \omega_i \sigma_i \|L_i\|^2 = (1+\delta)^{-1}. \tag{4.46}$$

For every $\boldsymbol{x} = (x, v_1, \ldots, v_m)$ in $\boldsymbol{\mathcal{K}}$, by using (4.46), we obtain

$$
\begin{aligned}
\langle \boldsymbol{x} \mid \boldsymbol{V}\boldsymbol{x}\rangle_{\boldsymbol{\mathcal{K}}} &= \tau^{-1}\|x\|^2 + \sum_{i=1}^m \sigma_i^{-1}\omega_i\|v_i\|^2_{\mathcal{G}_i} - 2\sum_{i=1}^m \omega_i\langle L_i x \mid v_i\rangle_{\mathcal{G}_i} \\
&= \tau^{-1}\|x\|^2 + \sum_{i=1}^m \sigma_i^{-1}\omega_i\|v_i\|^2_{\mathcal{G}_i} - 2\sum_{i=1}^m \omega_i\Big\langle \sqrt{\sigma_i}L_i x \mid \sqrt{\sigma_i}^{-1} v_i\Big\rangle_{\mathcal{G}_i} \\
&= \tau^{-1}\|x\|^2 + \sum_{i=1}^m \sigma_i^{-1}\omega_i\|v_i\|^2_{\mathcal{G}_i} - 2\Big\langle \boldsymbol{T}x \mid (\sqrt{\sigma_1}^{-1}v_1, \ldots, \sqrt{\sigma_m}^{-1}v_m)\Big\rangle_{\boldsymbol{\mathcal{G}}} \\
&\geqslant \tau^{-1}\|x\|^2 + \sum_{i=1}^m \sigma_i^{-1}\omega_i\|v_i\|^2_{\mathcal{G}_i} \\
&\qquad - \left(\frac{\|\boldsymbol{T}x\|^2_{\boldsymbol{\mathcal{G}}}}{\tau(1+\delta)\|\boldsymbol{T}\|^2} + \tau(1+\delta)\|\boldsymbol{T}\|^2\sum_{i=1}^m \sigma_i^{-1}\omega_i\|v_i\|^2_{\mathcal{G}_i}\right) \\
&\geqslant \Big(1-(1+\delta)^{-1}\Big)\left(\tau^{-1}\|x\|^2 + \sum_{i=1}^m \sigma_i^{-1}\omega_i\|v_i\|^2_{\mathcal{G}_i}\right) \\
&\geqslant \Big(1-(1+\delta)^{-1}\Big)\min\{\tau^{-1}, \sigma_1^{-1}, \ldots, \sigma_m^{-1}\}\|\boldsymbol{x}\|^2_{\boldsymbol{\mathcal{K}}} \\
&= \rho\|\boldsymbol{x}\|^2_{\boldsymbol{\mathcal{K}}}.
\end{aligned} \tag{4.47}
$$

Therefore, $\boldsymbol{V}$ is ρ-strongly positive. Furthermore, it follows from (4.47) that

$$
\boldsymbol{V}^{-1} \text{ exists and } \|\boldsymbol{V}^{-1}\| \leqslant \rho^{-1}. \tag{4.48}
$$

(i) : We first observe that (4.30) is equivalent to

$$
(\forall n \in \mathbb{N}) \quad
\left\lfloor
\begin{array}{l}
\tau^{-1}(x_n - p_n) - \sum_{i=1}^m \omega_i L_i^* v_{i,n} - Cx_n \in \\
\qquad\qquad -z + A(p_n - a_{2,n}) + a_{1,n} - \tau^{-1}a_{2,n} \\
x_{n+1} = x_n + \lambda_n(p_n - x_n) \\
\text{for } i = 1, \ldots, m \\
\left\lfloor
\begin{array}{l}
\sigma_i^{-1}(v_{i,n} - q_{i,n}) - L_i(x_n - p_n) - D_i^{-1}v_{i,n} \in \\
\qquad\qquad r_i + B_i^{-1}(q_{i,n} - b_{i,n}) - L_i p_n + c_{i,n} - \sigma_i^{-1}b_{i,n} \\
v_{i,n+1} = v_{i,n} + \lambda_n(q_{i,n} - v_{i,n}).
\end{array}
\right.
\end{array}
\right. \tag{4.49}
$$

Now set

$$
(\forall n \in \mathbb{N}) \quad
\begin{cases}
\boldsymbol{x}_n = (x_n, v_{1,n}, \ldots, v_{m,n}) \\
\boldsymbol{y}_n = (p_n, q_{1,n}, \ldots, q_{m,n}) \\
\boldsymbol{a}_n = (a_{2,n}, b_{1,n}, \ldots, b_{m,n}) \\
\boldsymbol{c}_n = (a_{1,n}, c_{1,n}, \ldots, c_{m,n}) \\
\boldsymbol{d}_n = (\tau^{-1}a_{2,n}, \sigma_1^{-1}b_{1,n}, \ldots, \sigma_m^{-1}b_{m,n}).
\end{cases} \tag{4.50}
$$

We have

$$\sum_{n\in\mathbb{N}} \|\boldsymbol{a}_n\|_{\boldsymbol{\mathcal{K}}} < +\infty, \quad \sum_{n\in\mathbb{N}} \|\boldsymbol{c}_n\|_{\boldsymbol{\mathcal{K}}} < +\infty, \quad \text{and} \quad \sum_{n\in\mathbb{N}} \|\boldsymbol{d}_n\|_{\boldsymbol{\mathcal{K}}} < +\infty. \tag{4.51}$$

Furthermore, (4.49) yields

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} \boldsymbol{V}(\boldsymbol{x}_n - \boldsymbol{y}_n) - \boldsymbol{Q}\boldsymbol{x}_n \in (\boldsymbol{M} + \boldsymbol{S})(\boldsymbol{y}_n - \boldsymbol{a}_n) + \boldsymbol{S}\boldsymbol{a}_n + \boldsymbol{c}_n - \boldsymbol{d}_n \\ \boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda_n(\boldsymbol{y}_n - \boldsymbol{x}_n). \end{array} \right. \tag{4.52}$$

Next, we set

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{b}_n = \boldsymbol{V}^{-1}\big((\boldsymbol{S} + \boldsymbol{V})\boldsymbol{a}_n + \boldsymbol{c}_n - \boldsymbol{d}_n\big). \tag{4.53}$$

Then (4.51) implies that

$$\sum_{n\in\mathbb{N}} \|\boldsymbol{b}_n\|_{\boldsymbol{\mathcal{K}}} < +\infty. \tag{4.54}$$

Moreover, using (4.48) and (4.53), we have

$$\begin{aligned}
&(\forall n \in \mathbb{N}) \quad \boldsymbol{V}(\boldsymbol{x}_n - \boldsymbol{y}_n) - \boldsymbol{Q}\boldsymbol{x}_n \in (\boldsymbol{M} + \boldsymbol{S})(\boldsymbol{y}_n - \boldsymbol{a}_n) + \boldsymbol{S}\boldsymbol{a}_n + \boldsymbol{c}_n - \boldsymbol{d}_n \\
\Leftrightarrow\ &(\forall n \in \mathbb{N}) \quad (\boldsymbol{V} - \boldsymbol{Q})\boldsymbol{x}_n \in (\boldsymbol{M} + \boldsymbol{S} + \boldsymbol{V})(\boldsymbol{y}_n - \boldsymbol{a}_n) + (\boldsymbol{S} + \boldsymbol{V})\boldsymbol{a}_n + \boldsymbol{c}_n - \boldsymbol{d}_n \\
\Leftrightarrow\ &(\forall n \in \mathbb{N}) \quad \boldsymbol{y}_n = \big(\boldsymbol{M} + \boldsymbol{S} + \boldsymbol{V}\big)^{-1}\Big((\boldsymbol{V} - \boldsymbol{Q})\boldsymbol{x}_n - (\boldsymbol{S} + \boldsymbol{V})\boldsymbol{a}_n - \boldsymbol{c}_n + \boldsymbol{d}_n\Big) + \boldsymbol{a}_n \\
\Leftrightarrow\ &(\forall n \in \mathbb{N}) \quad \boldsymbol{y}_n = \Big(\mathrm{Id} + \boldsymbol{V}^{-1}(\boldsymbol{M} + \boldsymbol{S})\Big)^{-1}\Big((\mathrm{Id} - \boldsymbol{V}^{-1}\boldsymbol{Q})\boldsymbol{x}_n - \boldsymbol{b}_n\Big) + \boldsymbol{a}_n.
\end{aligned} \tag{4.55}$$

We derive from (4.52) that, for every $n \in \mathbb{N}$,

$$\begin{aligned}
\boldsymbol{x}_{n+1} &= \boldsymbol{x}_n + \lambda_n\Big(\big(\mathrm{Id} + \boldsymbol{V}^{-1}(\boldsymbol{M} + \boldsymbol{S})\big)^{-1}\big(\boldsymbol{x}_n - \boldsymbol{V}^{-1}\boldsymbol{Q}\boldsymbol{x}_n - \boldsymbol{b}_n\big) + \boldsymbol{a}_n - \boldsymbol{x}_n\Big) \\
&= \boldsymbol{x}_n + \lambda_n\Big(J_{\boldsymbol{A}}\big(\boldsymbol{x}_n - \boldsymbol{B}\boldsymbol{x}_n - \boldsymbol{b}_n\big) + \boldsymbol{a}_n - \boldsymbol{x}_n\Big),
\end{aligned} \tag{4.56}$$

where

$$\boldsymbol{A} = \boldsymbol{V}^{-1}(\boldsymbol{M} + \boldsymbol{S}) \quad \text{and} \quad \boldsymbol{B} = \boldsymbol{V}^{-1}\boldsymbol{Q}. \tag{4.57}$$

Algorithm (4.56) has the structure of the forward-backward splitting algorithm [13]. Hence, it is sufficient to check the convergence conditions of the forward-backward splitting algorithm [13, Corollary 6.5] to prove our claims. To this end, let us introduce the real Hilbert space $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$ with scalar product and norm defined by

$$\big(\forall(\boldsymbol{x}, \boldsymbol{y}) \in \boldsymbol{\mathcal{K}} \times \boldsymbol{\mathcal{K}}\big) \quad \langle \boldsymbol{x} \mid \boldsymbol{y} \rangle_{\boldsymbol{V}} = \langle \boldsymbol{x} \mid \boldsymbol{V}\boldsymbol{y} \rangle_{\boldsymbol{\mathcal{K}}} \quad \text{and} \quad \|\boldsymbol{x}\|_{\boldsymbol{V}} = \sqrt{\langle \boldsymbol{x} \mid \boldsymbol{V}\boldsymbol{x} \rangle_{\boldsymbol{\mathcal{K}}}}, \tag{4.58}$$

respectively. Since $\boldsymbol{V}$ is a bounded linear operator, it follows from (4.51) and (4.54) that

$$\sum_{n\in\mathbb{N}} \|\boldsymbol{a}_n\|_{\boldsymbol{V}} < +\infty \quad \text{and} \quad \sum_{n\in\mathbb{N}} \|\boldsymbol{b}_n\|_{\boldsymbol{V}} < +\infty. \tag{4.59}$$

Moreover, since $\boldsymbol{M}+\boldsymbol{S}$ is monotone on $\boldsymbol{\mathcal{K}}$, we have

$$\big(\forall(\boldsymbol{x},\boldsymbol{y})\in\boldsymbol{\mathcal{K}}\times\boldsymbol{\mathcal{K}}\big)\quad \langle\boldsymbol{x}-\boldsymbol{y}\mid\boldsymbol{A}\boldsymbol{x}-\boldsymbol{A}\boldsymbol{y}\rangle_{\boldsymbol{V}}=\langle\boldsymbol{x}-\boldsymbol{y}\mid\boldsymbol{V}\boldsymbol{A}\boldsymbol{x}-\boldsymbol{V}\boldsymbol{A}\boldsymbol{y}\rangle_{\boldsymbol{\mathcal{K}}}$$
$$=\langle\boldsymbol{x}-\boldsymbol{y}\mid(\boldsymbol{M}+\boldsymbol{S})\boldsymbol{x}-(\boldsymbol{M}+\boldsymbol{S})\boldsymbol{y}\rangle_{\boldsymbol{\mathcal{K}}}\tag{4.60}$$
$$\geqslant 0.\tag{4.61}$$

Hence, $\boldsymbol{A}$ is monotone on $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$. Likewise, $\boldsymbol{B}$ is monotone on $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$. Since $\boldsymbol{V}$ is strongly positive, and since $\boldsymbol{M}+\boldsymbol{S}$ is maximally monotone on $\boldsymbol{\mathcal{K}}$, $\boldsymbol{A}$ is maximally monotone on $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$. Next, let us show that $\boldsymbol{B}$ is $(\beta\rho)$-cocoercive on $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$. Using (4.39), (4.47) and (4.48), we have, $\forall(\boldsymbol{x},\boldsymbol{y})\in\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}\times\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$,

$$\begin{aligned}\langle\boldsymbol{x}-\boldsymbol{y}\mid\boldsymbol{B}\boldsymbol{x}-\boldsymbol{B}\boldsymbol{y}\rangle_{\boldsymbol{V}}&=\langle\boldsymbol{x}-\boldsymbol{y}\mid\boldsymbol{V}\boldsymbol{B}\boldsymbol{x}-\boldsymbol{V}\boldsymbol{B}\boldsymbol{y}\rangle_{\boldsymbol{\mathcal{K}}}\\&=\langle\boldsymbol{x}-\boldsymbol{y}\mid\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\rangle_{\boldsymbol{\mathcal{K}}}\\&\geqslant\beta\|\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\|^2_{\boldsymbol{\mathcal{K}}}\\&=\beta\|\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\|_{\boldsymbol{\mathcal{K}}}\|\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\|_{\boldsymbol{\mathcal{K}}}\\&=\beta\|\boldsymbol{V}^{-1}\|^{-1}\|\boldsymbol{V}^{-1}\|\|\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\|_{\boldsymbol{\mathcal{K}}}\|\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\|_{\boldsymbol{\mathcal{K}}}\\&\geqslant\beta\|\boldsymbol{V}^{-1}\|^{-1}\|\boldsymbol{V}^{-1}\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{V}^{-1}\boldsymbol{Q}\boldsymbol{y}\|_{\boldsymbol{\mathcal{K}}}\|\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\|_{\boldsymbol{\mathcal{K}}}\\&\geqslant\beta\|\boldsymbol{V}^{-1}\|^{-1}\big\langle\boldsymbol{V}^{-1}\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{V}^{-1}\boldsymbol{Q}\boldsymbol{y}\mid\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\big\rangle_{\boldsymbol{\mathcal{K}}}\\&=\beta\|\boldsymbol{V}^{-1}\|^{-1}\langle\boldsymbol{B}\boldsymbol{x}-\boldsymbol{B}\boldsymbol{y}\mid\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\rangle_{\boldsymbol{\mathcal{K}}}\\&=\beta\|\boldsymbol{V}^{-1}\|^{-1}\|\boldsymbol{B}\boldsymbol{x}-\boldsymbol{B}\boldsymbol{y}\|^2_{\boldsymbol{V}}\\&\geqslant\beta\rho\|\boldsymbol{B}\boldsymbol{x}-\boldsymbol{B}\boldsymbol{y}\|^2_{\boldsymbol{V}}.\end{aligned}\tag{4.62}$$

Hence, by (4.11), $\boldsymbol{B}$ is $(\beta\rho)$-cocoercive on $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$. Moreover, it follows from our assumption that $2\beta\rho>1$. Altogether, by [13, Corollary 6.5] the sequence $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ converges weakly in $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$ to some $\overline{\boldsymbol{x}}=(\overline{x},\overline{v}_1,\ldots,\overline{v}_m)\in\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})=\operatorname{zer}(\boldsymbol{M}+\boldsymbol{S}+\boldsymbol{Q})$. Since $\boldsymbol{V}$ is self-adjoint and $\boldsymbol{V}^{-1}$ exists, the weak convergence of the sequence $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ to $\overline{\boldsymbol{x}}$ in $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$ is equivalent to the weak convergence of $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ to $\overline{\boldsymbol{x}}$ in $\boldsymbol{\mathcal{K}}$. Hence, $\boldsymbol{x}_n\rightharpoonup\overline{\boldsymbol{x}}\in\operatorname{zer}(\boldsymbol{M}+\boldsymbol{S}+\boldsymbol{Q})$. It follows from (4.40) that $\overline{x}\in\mathcal{P}$ and $(\overline{v}_1,\ldots,\overline{v}_m)\in\mathcal{D}$. This proves (i).

(ii)&(iii) : It follows from [13, Remark 3.4] that

$$\sum_{n\in\mathbb{N}}\|\boldsymbol{B}\boldsymbol{x}_n-\boldsymbol{B}\overline{\boldsymbol{x}}\|^2_{\boldsymbol{V}}<+\infty.\tag{4.63}$$

On the other hand, from (4.47) and (4.63) yield $\boldsymbol{B}\boldsymbol{x}_n-\boldsymbol{B}\overline{\boldsymbol{x}}=\boldsymbol{V}^{-1}(\boldsymbol{Q}\boldsymbol{x}_n-\boldsymbol{Q}\overline{\boldsymbol{x}})\to 0$, which implies that $\boldsymbol{Q}\boldsymbol{x}_n-\boldsymbol{Q}\overline{\boldsymbol{x}}\to 0$. Hence,

$$Cx_n\to C\overline{x}\quad\text{and}\quad\big(\forall i\in\{1,\ldots,m\}\big)\quad D_i^{-1}v_{i,n}\to D_i^{-1}\overline{v}_i.\tag{4.64}$$

If C is uniformly monotone at $\overline{x}$, then there exists an increasing function $\phi_C\colon\ [0,+\infty[\,\to[0,+\infty]$ vanishing only at 0 such that

$$\phi_C(\|x_n-\overline{x}\|)\leqslant\langle x_n-\overline{x}\mid Cx_n-C\overline{x}\rangle\leqslant\|x_n-\overline{x}\|\,\|Cx_n-C\overline{x}\|.\tag{4.65}$$

Since $(x_n - \overline{x})_{n\in\mathbb{N}}$ is bounded, it follows from (4.64) and (4.65) that $x_n \to \overline{x}$. This proves (ii), and (iii) is proved in a similar fashion. □

Remark 4.9 Here are some remarks concerning the connections between our framework and existing work.

(i) The strategy used in the proof of Theorem 4.8 is to reformulate algorithm (4.30) as a forward-backward splitting algorithm in a real Hilbert space endowed with a suitable norm. Such a renorming technique was used in [22] for a minimization problem in finite-dimensional spaces. The same strategy is also used in the primal-dual minimization problem of [18], which will be further discussed in Remark 4.13(iii) below.

(ii) Consider the special case when $z = 0$, and $(B_i)_{1\leqslant i\leqslant m}$ and $(D_i)_{1\leqslant i\leqslant m}$ are as in (4.14). Then algorithm (4.30) reduces to

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n\Big(J_{\tau A}\Big(x_n - \tau(Cx_n + a_{1,n})\Big) + a_{2,n} - x_n\Big), \tag{4.66}$$

which is the standard forward-backward splitting algorithm [13, Algorithm 6.4] where the sequence $(\gamma_n)_{n\in\mathbb{N}}$ in [13, Eq. (6.3)] is constant.

(iii) Problems (4.17) and (4.18) in Example 4.7 can also be solved by [7, Theorem 3.8]. However, the algorithm resulting from (4.30) in this special case is different from that of [7, Theorem 3.8].

(iv) In Problem 4.5, since C and $(D_i^{-1})_{1\leqslant i\leqslant m}$ are cocoercive, they are Lipschitzian. Hence, Problem 4.5 can also be solved by the algorithm proposed in [16, Theorem 3.1], which has a different structure from that of the present algorithm.

(v) Consider the special case when $z = 0$ and $(\forall i \in \{1,\ldots,m\})$ $\mathcal{G}_i = \mathcal{H}, L_i = \mathrm{Id}$, $D_i^{-1} = 0$, and $r_i = 0$. Then the primal inclusion (4.12) reduces to

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } 0 \in A\overline{x} + \sum_{i=1}^{m} \omega_i B_i\overline{x} + C\overline{x}. \tag{4.67}$$

An alternative algorithm to solve this problem is proposed in [26], which provides only primal solution.

Remark 4.10 Suppose that C and $(D_i)_{1\leqslant i\leqslant m}$ are as in (4.16). Then it can be shown that Theorem 4.8(i) remains valid for any sequence $(\lambda_n)_{n\in\mathbb{N}}$ in $[\varepsilon, 2-\varepsilon]$ and with condition (4.29) replaced by

$$\tau\sum_{i=1}^{m} \omega_i\sigma_i\|L_i\|^2 < 1. \tag{4.68}$$

4.2.4 Application to minimization problems

We provide an application of the algorithm (4.30) to minimization problems, by revisiting [16, Problem 4.1].

Problem 4.11 Let $\mathcal{H}$ be a real Hilbert space, let $z \in \mathcal{H}$, let m be a strictly positive integer, let $(\omega_i)_{1\leqslant i\leqslant m}$ be real numbers in $]0,1]$ such that $\sum_{i=1}^m \omega_i = 1$, let $f \in \Gamma_0(\mathcal{H})$, and let $h\colon \mathcal{H} \to \mathbb{R}$ be convex and differentiable with a μ^{-1}-Lipschitzian gradient for some $\mu \in\,]0,+\infty[$. For every $i \in \{1,\ldots,m\}$, let $\mathcal{G}_i$ be a real Hilbert space, let $r_i \in \mathcal{G}_i$, let $g_i \in \Gamma_0(\mathcal{G}_i)$, let $\ell_i \in \Gamma_0(\mathcal{G}_i)$ be ν_i-strongly convex, for some $\nu_i \in\,]0,+\infty[$, and suppose that $L_i\colon \mathcal{H} \to \mathcal{G}_i$ is a nonzero bounded linear operator. Consider the primal problem

$$\underset{x\in\mathcal{H}}{\text{minimize}}\; f(x) + \sum_{i=1}^m \omega_i (g_i \,\square\, \ell_i)(L_i x - r_i) + h(x) - \langle x \mid z\rangle, \tag{4.69}$$

and the dual problem

$$\underset{v_1\in\mathcal{G}_1,\ldots,v_m\in\mathcal{G}_m}{\text{minimize}}\; (f^* \,\square\, h^*)\Big(z - \sum_{i=1}^m \omega_i L_i^* v_i\Big) + \sum_{i=1}^m \omega_i \big(g_i^*(v_i) + \ell_i^*(v_i) + \langle v_i \mid r_i\rangle_{\mathcal{G}_i}\big). \tag{4.70}$$

We denote by $\mathcal{P}_1$ and $\mathcal{D}_1$ the sets of solutions to (4.69) and (4.70), respectively.

Corollary 4.12 *In Problem 4.11, suppose that*

$$z \in \operatorname{ran}\Big(\partial f + \sum_{i=1}^m \omega_i L_i^*\big((\partial g_i \,\square\, \partial \ell_i)(L_i \cdot - r_i)\big) + \nabla h\Big). \tag{4.71}$$

Let τ and $(\sigma_i)_{1\leqslant i\leqslant m}$ be strictly positive numbers such that

$$2\rho \min\{\mu,\nu_1,\ldots,\nu_m\} > 1, \textit{where } \rho = \min\big\{\tau^{-1},\sigma_1^{-1},\ldots,\sigma_m^{-1}\big\}\Big(1 - \sqrt{\tau\sum_{i=1}^m \sigma_i\omega_i\|L_i\|^2}\Big). \tag{4.72}$$

Let $\varepsilon \in\,]0,1[$ and let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon,1]$, let $x_0 \in \mathcal{H}$, let $(a_{1,n})_{n\in\mathbb{N}}$ and $(a_{2,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{H}$. For every $i \in \{1,\ldots,m\}$, let $v_{i,0} \in \mathcal{G}_i$, and let $(b_{i,n})_{n\in\mathbb{N}}$ and $(c_{i,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{G}_i$. Let $(x_n)_{n\in\mathbb{N}}$ and $(v_{1,n},\ldots,v_{m,n})_{n\in\mathbb{N}}$ be sequences generated by the following routine

$$(\forall n\in\mathbb{N})\quad \left\lfloor \begin{array}{l} p_n = \operatorname{prox}_{\tau f}\Big(x_n - \tau\Big(\sum_{i=1}^m \omega_i L_i^* v_{i,n} + \nabla h(x_n) + a_{1,n} - z\Big)\Big) + a_{2,n}\\ y_n = 2p_n - x_n\\ x_{n+1} = x_n + \lambda_n(p_n - x_n)\\ \text{for } i = 1,\ldots,m\\ \left\lfloor \begin{array}{l} q_{i,n} = \operatorname{prox}_{\sigma_i g_i^*}\Big(v_{i,n} + \sigma_i\Big(L_i y_n - \nabla \ell_i^*(v_{i,n}) + c_{i,n} - r_i\Big)\Big) + b_{i,n}\\ v_{i,n+1} = v_{i,n} + \lambda_n(q_{i,n} - v_{i,n}). \end{array}\right. \end{array}\right.$$

$$ \tag{4.73} $$

Then the following hold for some $\overline{x} \in \mathcal{P}_1$ and $(\overline{v}_1, \ldots, \overline{v}_m) \in \mathcal{D}_1$.

(i) $x_n \rightharpoonup \overline{x}$ *and* $(v_{1,n}, \ldots, v_{m,n}) \rightharpoonup (\overline{v}_1, \ldots, \overline{v}_m)$.

(ii) *Suppose that h is uniformly convex at $\overline{x}$. Then $x_n \to \overline{x}$.*

(iii) *Suppose that ℓ_j^* is uniformly convex at $\overline{v}_j$ for some $j \in \{1, \ldots, m\}$. Then $v_{j,n} \to \overline{v}_j$.*

Proof. The connection between Problem 4.11 and Problem 4.5 is established in the proof of [16, Theorem 4.2]. Since ∇h is μ^{-1}-Lipschitz continuous, by the Baillon-Haddad Theorem [4, 5], it is μ-cocoercive. Moreover since, for every $i \in \{1, \ldots, m\}$, ℓ_i is ν_i-strongly convex, $\partial \ell_i$ is ν_i-strongly monotone. Hence, by applying Theorem 4.8(i) with $A = \partial f$, $J_{\tau A} = \text{prox}_{\tau f}$, $C = \nabla h$ and for every $i \in \{1, \ldots, m\}$, $D_i^{-1} = \nabla \ell_i^*$, $B_i = \partial g_i$, $J_{\sigma_i B_i^{-1}} = \text{prox}_{\sigma_i g_i^*}$, we obtain that the sequence $(x_n)_{n\in\mathbb{N}}$ converges weakly to some $\overline{x} \in \mathcal{H}$ such that

$$ z \in \partial f(\overline{x}) + \sum_{i=1}^{m} \omega_i L_i^* \big((\partial g_i \,\square\, \partial \ell_i)(L_i \overline{x} - r_i) \big) + \nabla h(\overline{x}), \tag{4.74} $$

and the sequence $((v_{1,n}, \ldots, v_{m,n}))_{n\in\mathbb{N}}$ converges weakly to some $(\overline{v}_1, \ldots, \overline{v}_m)$ such that

$$ (\exists\, x \in \mathcal{H}) \quad \begin{cases} z - \sum_{i=1}^{m} \omega_i L_i^* \overline{v}_i \in \partial f(x) + \nabla h(x) \\ (\forall i \in \{1, \ldots, m\}) \quad \overline{v}_i \in (\partial g_i \,\square\, \partial \ell_i)(L_i x - r_i). \end{cases} \tag{4.75} $$

As shown in the proof of [16, Theorem 4.2], $\overline{x} \in \mathcal{P}_1$ and $(\overline{v}_1, \ldots, \overline{v}_m) \in \mathcal{D}_1$. This proves (i). Now, if h is uniformly convex at $\overline{x}$, then ∇h is uniformly monotone at $\overline{x}$. Hence, (ii) follows from Theorem 4.8(ii). Similarly, (iii) follows from Theorem 4.8(iii). ∎

Remark 4.13 Here are some observations on the above results.

(i) If a function $\varphi \colon \mathcal{H} \to \mathbb{R}$ is convex and differentiable function with a β^{-1}-Lipschitzian gradient, then $\nabla \varphi$ is β-cocoercive [4, 5]. Hence, in the context of convex minimization problems, the restriction of cocoercivity made in Problem 4.5 with respect to the problem considered in [16] disappears. Yet, the algorithm we obtain is quite different from that proposed in [16, Theorem 4.2].

(ii) Sufficient conditions which ensure that (4.71) is satisfied are provided in [16, Proposition 4.3]. For instance, if (4.69) has at least one solution, if $\mathcal{H}$ and $(\mathcal{G}_i)_{1\leqslant i\leqslant m}$ are finite-dimensional, and if there exists $x \in \text{ri}\,\text{dom}\, f$ such that

$$ (\forall i \in \{1, \ldots, m\}) \quad L_i x - r_i \in \text{ri}\,\text{dom}\, g_i + \text{ri}\,\text{dom}\, \ell_i, \tag{4.76} $$

then (4.71) holds.

(iii) Consider the special case when $z = 0$ and, for every $i \in \{1,\ldots,m\}$, $r_i = 0, \sigma_i = \sigma \in \,]0,+\infty[$, and

$$\ell_i\colon v \mapsto \begin{cases} 0 & \text{if } v = 0, \\ +\infty & \text{otherwise.} \end{cases} \tag{4.77}$$

Then (4.73) reduces to

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} p_n = \operatorname{prox}_{\tau f}\Big(x_n - \tau\Big(\sum_{i=1}^m \omega_i L_i^* v_{i,n} + \nabla h(x_n) + a_{1,n}\Big)\Big) + a_{2,n} \\ y_n = 2p_n - x_n \\ x_{n+1} = x_n + \lambda_n(p_n - x_n) \\ \text{for } i = 1,\ldots,m \\ \left\lfloor \begin{array}{l} q_{i,n} = \operatorname{prox}_{\sigma g_i^*}\Big(v_{i,n} + \sigma\big(L_i y_n + c_{i,n}\big)\Big) + b_{i,n} \\ v_{i,n+1} = v_{i,n} + \lambda_n(q_{i,n} - v_{i,n}), \end{array} \right. \end{array} \right. \tag{4.78}$$

which is the method proposed in [18, Eq. (36)]. However, in this setting, the conditions (4.71) and (4.72) are different from the conditions [18, Eq. (39)] and [18, Eq. (38)], respectively.

(iv) In finite-dimensional spaces, with exact implementation of the operators, and with the additional restrictions that $m = 1$, $h\colon x \mapsto 0$, ℓ_1 is as in (4.77), $r_1 = 0$, and $z = 0$, (4.73) remains convergent if $\lambda_n \equiv \lambda \in \,]0,2[$ under the same condition presented here [22, Remark 5.4]. If we further impose the restriction $\lambda_n \equiv 1$, then (4.73) reduces to the method proposed in [10, Algorithm 1]. An alternative primal-dual algorithm is proposed in [12].

Acknowledgement. I thank Professor Patrick L. Combettes for bringing this problem to my attention and for helpful discussions.

4.3 Bibliographie

[1] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran, Alternating proximal algorithms for weakly coupled convex minimization problems – Applications to dynamical games and PDE's, *J. Convex Anal.,* vol. 15, pp. 485–506, 2008.

[2] H. Attouch, L. M. Briceño-Arias, and P. L. Combettes, A parallel splitting method for coupled monotone inclusions, *SIAM J. Control Optim.*, vol. 48, pp. 3246–3270, 2010.

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

[4] J.-B. Baillon and G. Haddad, Quelques propriétés des opérateurs angle-bornés et n-cycliquement monotones, *Israel J. Math.,* vol. 26, pp. 137–150, 1977.

[5] H. H. Bauschke and P. L. Combettes, The Baillon-Haddad theorem revisited, *J. Convex Anal.*, vol. 17, pp. 781–787, 2010.

[6] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York, 2011.

[7] 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.

[8] L. M. Briceño-Arias and P. L. Combettes, Monotone operator methods for Nash equilibria in non-potential games, in *Computational and Analytical Mathematics,* (D. Bailey, H. H. Bauschke, P. Borwein, F. Garvan, M. Théra, J. Vanderwerff, and H. Wolkowicz, eds.). Springer, New York, 2013.

[9] J.-F. Cai, R. H. Chan, and Z. Shen, Simultaneous cartoon and texture inpainting, *Inverse Probl. Imaging,* vol. 4, pp. 379–395, 2010.

[10] A. Chambolle and T. Pock, A first order primal dual algorithm for convex problems with applications to imaging, *J. Math. Imaging Vision*, vol. 40, pp. 120–145, 2011.

[11] G. H-G. Chen and R. T. Rockafellar, Convergence rates in forward-backward splitting, *SIAM J. Optim.,* vol. 7, pp. 421–444, 1997.

[12] G. Chen and M. Teboulle, A proximal based decomposition method for minimization problems, *Math. Program.*, vol. 64, pp. 81–101, 1994.

[13] P. L. Combettes, Solving monotone inclusions via compositions of nonexpansive averaged operators, *Optimization*, vol. 53, pp. 475–504, 2004.

[14] P. L. Combettes, Đinh Dũng, and B. C. Vũ, Proximity for sums of composite functions, *J. Math. Anal. Appl.*, vol. 380, pp. 680–688, 2011.

[15] P. L. Combettes and J.-C. Pesquet, Proximal splitting methods in signal processing, in : *Fixed-Point Algorithms for Inverse Problems in Science and Engineering,* H. H. Bauschke *et al.* eds. Springer, New York, pp. 185–212, 2011.

[16] P. L. Combettes and J.-C. Pesquet, Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators, *Set-Valued Var. Anal.*, vol. 20, pp. 307–330, 2012.

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

[18] L. Condat, A generic first-order primal-dual method for convex optimization involving Lipschitzian, proximable and linear composite terms, 2011. http://hal.archives-ouvertes.fr/hal-00609728/fr/

[19] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer-Verlag, New York, 2003.

[20] D. Gabay, Applications of the method of multipliers to variational inequalities, in : M. Fortin and R. Glowinski (eds.), *Augmented Lagrangian Methods : Applications to the Numerical Solution of Boundary Value Problems,* pp. 299–331. North-Holland, Amsterdam, 1983.

[21] R. Glowinski and P. Le Tallec, *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*, SIAM, Philadelphia, 1989.

[22] B. He and X. Yuan, Convergence analysis of primal-dual algorithms for a saddle-point problem : from contraction perspective, *SIAM J. Imaging Sci,* vol.5, pp. 119–149, 2012.

[23] B. Mercier, *Topics in Finite Element Solution of Elliptic Problems* (Lectures on Mathematics, no. 63). Tata Institute of Fundamental Research, Bombay, 1979.

[24] U. Mosco, Dual variational inequalities, *J. Math. Anal. Appl.,* vol. 40, pp. 202–206, 1972.

[25] J.-C. Pesquet and N. Pustelnik, A parallel inertial proximal optimization method, *Pacific Journal of Optimization, vol. 8,* no. 2, pp. 273–305, 2012.

[26] H. Raguet, J. Fadili, and G. Peyré, Generalized forward-backward splitting, 2011. http://arxiv.org/abs/1108.4404

[27] R. T. Rockafellar, Duality and stability in extremum problems involving convex functions, *Pacific J. Math.,* vol. 21, pp. 167–187, 1967.

[28] P. Tseng, Further applications of a splitting algorithm to decomposition in variational inequalities and convex programming, *Math. Programming*, vol. 48, pp. 249–263, 1990.

[29] P. Tseng, Applications of a splitting algorithm to decomposition in convex programming and variational inequalities, *SIAM J. Control Optim.*, vol. 29, pp. 119–138, 1991.

[30] D. L. Zhu and P. Marcotte, Co-coercivity and its role in the convergence of iterative schemes for solving variational inequalities, *SIAM J. Optim.*, vol. 6, pp. 714–726, 1996.

Chapitre 5

Suites quasi-fejériennes à métrique variable

Nous introduisons la notion de suite quasi-fejérienne à métrique variable et analysons ses propriétés asymptotiques dans des espaces hilbertiens. On déduit de ces résultats la convergence de nouveaux algorithmes avec métrique variable.

5.1 Description et résultats principaux

Définition 5.1 Soient $\alpha \in \left]0,+\infty\right[$, $\phi\colon \left[0,+\infty\right[\to \left[0,+\infty\right[$, $(W_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{P}_\alpha(\mathcal{H})$, C un sous-ensemble non vide de $\mathcal{H}$, et $(x_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{H}$. Alors $(x_n)_{n\in\mathbb{N}}$ est :

(i) une suite ϕ-quasi-fejérienne monotone par rapport à C relativement à $(W_n)_{n\in\mathbb{N}}$ si

$$\big(\exists\,(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})\big)\big(\forall z \in C\big)\big(\exists\,(\varepsilon_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})\big)(\forall n \in \mathbb{N})$$
$$\phi(\|x_{n+1}-z\|_{W_{n+1}}) \leqslant (1+\eta_n)\phi(\|x_n - z\|_{W_n}) + \varepsilon_n; \quad (5.1)$$

(ii) une suite stationnairement ϕ-quasi-fejérienne monotone par rapport à C relativement à $(W_n)_{n\in\mathbb{N}}$ si

$$\big(\exists\,(\varepsilon_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})\big)\big(\exists\,(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})\big)(\forall z \in C)(\forall n \in \mathbb{N})$$
$$\phi(\|x_{n+1}-z\|_{W_{n+1}}) \leqslant (1+\eta_n)\phi(\|x_n - z\|_{W_n}) + \varepsilon_n. \quad (5.2)$$

Dans le cas où $W_n \equiv \mathrm{Id}$, $\eta_n \equiv 0$, et $\phi = |\cdot|$ ou $\phi = |\cdot|^2$, on obtient la notion de suite quasi–fejérienne par rapport à C étudiée dans [12]. De plus, si $\varepsilon_n \equiv 0$, (5.1) correspond aux cas classiques étudiés dans [20, 27, 30]. La notion de suite quasi-fejérienne est un outil fondamental pour analyser la convergence de méthodes numériques [2, 5, 12, 13, 14, 18, 19, 30, 31, 35].

Nous présentons quelques propriétés élémentaires dans la proposition suivante.

Proposition 5.2 *Soit $\alpha \in]0,+\infty[$, soit $\phi\colon [0,+\infty[\to [0,+\infty[$ une fonction strictement croissante telle que $\lim_{t\to+\infty}\phi(t) = +\infty$, soit $(W_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{P}_\alpha(\mathcal{H})$, soit C un sous-ensemble non vide de $\mathcal{H}$, et soit $(x_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{H}$ telle que (5.1) est vérifiée. Nous avons les propriétés suivantes.*

(i) *Soit $z \in C$. Alors $(\|x_n - z\|_{W_n})_{n\in\mathbb{N}}$ converge.*

(ii) *$(x_n)_{n\in\mathbb{N}}$ est bornée.*

Nous avons le résultat suivant pour la convergence faible.

Théorème 5.3 *Soit $\alpha \in]0,+\infty[$, soit $\phi\colon [0,+\infty[\to [0,+\infty[$ une fonction strictement croissante et telle que $\lim_{t\to+\infty}\phi(t) = +\infty$, soient $(W_n)_{n\in\mathbb{N}}$ et W des opérateurs dans $\mathcal{P}_\alpha(\mathcal{H})$ tels que $W_n \to W$ ponctuellement, soit C un sous-ensemble non vide de $\mathcal{H}$, et soit $(x_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{H}$ telle que (5.1) est vérifiée. Alors $(x_n)_{n\in\mathbb{N}}$ converge faiblement vers un point de C si et seulement si tous les points d'accumulation faible de $(x_n)_{n\in\mathbb{N}}$ sont dans C.*

Nous présentons les caractérisations de la convergence forte.

Proposition 5.4 *Soit $\alpha \in]0,+\infty[$, soit $\chi \in [1,+\infty[$, et soit $\phi\colon [0,+\infty[\to [0,+\infty[$ une fonction croissante semi-continue supérieurement volatilisée seulement en 0 telle que*

$$\big(\forall(\xi_1,\xi_2) \in [0,+\infty[^2\big) \quad \phi(\xi_1+\xi_2) \leqslant \chi\big(\phi(\xi_1)+\phi(\xi_2)\big). \tag{5.3}$$

Soit $(W_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{P}_\alpha(\mathcal{H})$ telle que $\mu = \sup_{n\in\mathbb{N}}\|W_n\| < +\infty$, soit C un sous-ensemble fermé non vide de $\mathcal{H}$, et soit $(x_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{H}$ telle que (5.2) est vérifiée. Alors $(x_n)_{n\in\mathbb{N}}$ converge fortement vers un point dans C si et seulement si $\underline{\lim}\, d_C(x_n) = 0$.

Proposition 5.5 *Soit $\alpha \in]0,+\infty[$, soit $(\nu_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$, et soit $(W_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{P}_\alpha(\mathcal{H})$ telle que*

$$\mu = \sup_{n\in\mathbb{N}}\|W_n\| < +\infty \quad \text{et} \quad (\forall n \in \mathbb{N}) \quad (1+\nu_n)W_{n+1} \succcurlyeq W_n. \tag{5.4}$$

De plus, soit C un sous-ensemble de $\mathcal{H}$ tel que $\operatorname{int} C \neq \varnothing$, soient $z \in C$ et $\rho \in]0,+\infty[$ tels que $B(z;\rho) \subset C$, et soit $(x_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{H}$ telle que

$$\big(\exists\,(\varepsilon_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})\big)\big(\exists\,(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})\big)(\forall x \in B(z;\rho))(\forall n \in \mathbb{N})$$
$$\|x_{n+1}-x\|_{W_{n+1}}^2 \leqslant (1+\eta_n)\|x_n - x\|_{W_n}^2 + \varepsilon_n. \tag{5.5}$$

Alors $(x_n)_{n\in\mathbb{N}}$ converge fortement.

Nous nous penchons à présent sur le cas où $\phi = |\cdot|^2$.

Proposition 5.6 *Soit $\alpha \in]0,+\infty[$, soit $(\eta_n)_{n\in\mathbb{N}}$ une suite dans $\ell_+^1(\mathbb{N})$, soit $(W_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{P}_\alpha(\mathcal{H})$ telle que*

$$\mu = \sup_{n\in\mathbb{N}} \|W_n\| < +\infty \quad \textit{et} \quad (\forall n \in \mathbb{N}) \quad (1+\eta_n)W_n \succcurlyeq W_{n+1}. \tag{5.6}$$

Soit C un sous-ensemble non vide de $\mathcal{H}$ et soit $(x_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{H}$ telle que

$$\big(\exists\,(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})\big)\big(\forall z \in C\big)\big(\exists\,(\varepsilon_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})\big)(\forall n \in \mathbb{N})$$
$$\|x_{n+1} - z\|^2_{W_{n+1}} \leqslant (1+\eta_n)\|x_n - z\|^2_{W_n} + \varepsilon_n. \tag{5.7}$$

Alors, nous avons les propriétés suivantes.

(i) *$(x_n)_{n\in\mathbb{N}}$ est $|\cdot|^2$-quasi-fejérienne par rapport à $\operatorname{conv} C$ relativement à $(W_n)_{n\in\mathbb{N}}$.*

(ii) *Pour tout $y \in \overline{\operatorname{conv}}\, C$, $(\|x_n - y\|_{W_n})_{n\in\mathbb{N}}$ converge.*

Dans le cas où $W_n \equiv \mathrm{Id}\,, \eta_n \equiv 0, \varepsilon \equiv 0$, et $\phi = |\cdot|$ ou $\phi = |\cdot|^2$, et le sous-ensemble C dans (5.2) est convexe fermé, la suite de projections $(P_C x_n)_{n\in\mathbb{N}}$ converge fortement ; voir [2, Theorem 2.16(iv)], [32, Remark 1], et [12, Proposition 3.6(iv)]. Nous montrons qu'il est encore vrai dans le contexte de métrique variable.

Proposition 5.7 *Soit $\alpha \in]0,+\infty[$, soit $(\eta_n)_{n\in\mathbb{N}}$ une suite dans $\ell_+^1(\mathbb{N})$, soit $(W_n)_{n\in\mathbb{N}}$ une suite uniformément bornée dans $\mathcal{P}_\alpha(\mathcal{H})$, soit C un sous-ensemble convexe fermé non vide de $\mathcal{H}$, et soit $(x_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{H}$ telle que*

$$\big(\exists\,(\varepsilon_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})\big)\big(\exists\,(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})\big)(\forall z \in C)(\forall n \in \mathbb{N})$$
$$\|x_{n+1} - z\|^2_{W_{n+1}} \leqslant (1+\eta_n)\|x_n - z\|^2_{W_n} + \varepsilon_n. \tag{5.8}$$

Alors $(P_C^{W_n} x_n)_{n\in\mathbb{N}}$ converge fortement.

En fin, nous proposons un nouvel algorithme pour résoudre les problèmes d'admissibilité convexe.

Théorème 5.8 *Soit $(C_i)_{i\in I}$ une famille finie ou infinie dénombrable de sous-ensembles convexes fermés de $\mathcal{H}$ telle que $C = \bigcap_{i\in I} C_i \neq \varnothing$, soit $(a_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{H}$ telle que $\sum_{n\in\mathbb{N}} \|a_n\| < +\infty$, soit $\alpha \in]0,+\infty[$, soit $(\eta_n)_{n\in\mathbb{N}}$ une suite dans $\ell_+^1(\mathbb{N})$, et soit $(W_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{P}_\alpha(\mathcal{H})$ telle que*

$$\mu = \sup_{n\in\mathbb{N}} \|W_n\| < +\infty \quad \textit{et} \quad (\forall n \in \mathbb{N}) \quad (1+\eta_n)W_n \succcurlyeq W_{n+1}. \tag{5.9}$$

Soit $\mathrm{i}\colon \mathbb{N} \to I$ telle que

$$(\forall j \in I)(\exists\, M_j \in \mathbb{N} \smallsetminus \{0\})(\forall n \in \mathbb{N}) \quad j \in \{\mathrm{i}(n), \ldots, \mathrm{i}(n+M_j-1)\}. \tag{5.10}$$

Pour tout $i \in I$, soit $(T_{i,n})_{n\in\mathbb{N}}$ une suite d'opérateurs telle que

$$(\forall n \in \mathbb{N}) \quad T_{i,n} \in \mathfrak{T}(W_n) \quad \textit{et} \quad \operatorname{Fix} T_{i,n} = C_i. \tag{5.11}$$

Soient $\varepsilon \in \,]0,1[$*,* $x_0 \in \mathcal{H}$*,* $(\lambda_n)_{n\in\mathbb{N}}$ *une suite dans* $[\varepsilon, 2-\varepsilon]$*, et*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n\big(T_{\mathrm{i}(n),n}x_n + a_n - x_n\big). \tag{5.12}$$

Supposons que, pour toute suite strictement croissante $(p_n)_{n\in\mathbb{N}}$ *dans* $\mathbb{N}$*, pour tout* $x \in \mathcal{H}$*, et tout* $j \in I$*,*

$$\begin{cases} x_{p_n} \rightharpoonup x \\ T_{j,p_n}x_{p_n} - x_{p_n} \to 0 \\ (\forall n \in \mathbb{N}) \;\; j = \mathrm{i}(p_n) \end{cases} \quad \Rightarrow \quad x \in C_j. \tag{5.13}$$

Alors, pour un point $\overline{x} \in C$*, nous avons les propriétés suivantes.*

(i) $x_n \rightharpoonup \overline{x}$.

(ii) *Supposons que* $\operatorname{int} C \neq \varnothing$ *et qu'il existe* $(\nu_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$ *telle que* $(\forall n \in \mathbb{N})$ $(1+\nu_n)W_{n+1} \succcurlyeq W_n$*. Alors* $x_n \to \overline{x}$*.*

(iii) *Supposons que* $\varliminf d_C(x_n) = 0$*. Alors* $x_n \to \overline{x}$*.*

(iv) *Supposons qu'il existe un index* $j \in I$ *de demicompact régularité : pour toute suite strictement croissante* $(p_n)_{n\in\mathbb{N}}$ *in* $\mathbb{N}$

$$\begin{cases} \sup_{n\in\mathbb{N}} \|x_{p_n}\| < +\infty \\ T_{j,p_n}x_{p_n} - x_{p_n} \to 0 \\ (\forall n \in \mathbb{N}) \;\; j = \mathrm{i}(p_n) \end{cases} \quad \Rightarrow \quad (x_{p_n})_{n\in\mathbb{N}} \text{ admet un point d'accumulation forte.} \tag{5.14}$$

Alors $x_n \to \overline{x}$*.*

Nous présentons une application du Théorème 5.8 à la méthode des projections périodiques à métrique variable.

Corollaire 5.9 *Soit* m *un entier strictement positif, soit* $I = \{1,\ldots,m\}$*, soit* $(C_i)_{i\in I}$ *une famille de sous-ensembles convexes fermés non vides de* $\mathcal{H}$ *telle que* $C = \bigcap_{i\in I} C_i \neq \varnothing$*, soit* $x_0 \in \mathcal{H}$*, soit* $\alpha \in \,]0,+\infty[$*, soit* $(\eta_n)_{n\in\mathbb{N}}$ *une suite dans* $\ell_+^1(\mathbb{N})$*, et soit* $(W_n)_{n\in\mathbb{N}}$ *une suite dans* $\mathcal{P}_\alpha(\mathcal{H})$ *telle que* $\sup_{n\in\mathbb{N}} \|W_n\| < +\infty$ *et* $(\forall n \in \mathbb{N})$ $(1+\eta_n)W_n \succcurlyeq W_{n+1}$*. Posons*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = P_{C_{1+\mathrm{rem}(n,m)}}^{W_n} x_n, \tag{5.15}$$

où $\mathrm{rem}(n,m) = n(\operatorname{mod} m) + 1$*. Alors, pour un point* $\overline{x} \in C$*, nous avons les propriétés suivantes.*

(i) $x_n \rightharpoonup \overline{x}$.

(ii) *Supposons que* $\operatorname{int} C \neq \varnothing$ *et qu'il existe* $(\nu_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$ *telle que* $(\forall n \in \mathbb{N})$ $(1+\nu_n)W_{n+1} \succcurlyeq W_n$*. Alors* $x_n \to \overline{x}$*.*

(iii) *Supposons qu'il existe $j \in I$ tel que C_j est bornément compact, i.e., son intersection avec toute boule fermée de $\mathcal{H}$ est compacte. Alors $x_n \to \overline{x}$.*

Corollaire 5.10 *Soit $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ un opérateur maximalement monotone tel que $C = \{z \in \mathcal{H} \mid 0 \in Az\} \neq \varnothing$, soit $\alpha \in \left]0,+\infty\right[$, soit $(a_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{H}$ telle que $\sum_{n\in\mathbb{N}} \|a_n\| < +\infty$, soit $(\eta_n)_{n\in\mathbb{N}}$ une suite dans $\ell^1_+(\mathbb{N})$, et soit $(W_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{P}_\alpha(\mathcal{H})$ telle que $\mu = \sup_{n\in\mathbb{N}} \|W_n\| < +\infty$ et $(\forall n \in \mathbb{N})\ (1+\eta_n)W_n \succcurlyeq W_{n+1}$. Soient $\varepsilon \in \left]0,1\right[$, $x_0 \in \mathcal{H}$, $(\lambda_n)_{n\in\mathbb{N}}$ une suite dans $[\varepsilon, 2-\varepsilon]$, $(\gamma_n)_{n\in\mathbb{N}}$ une suite dans $\left[\varepsilon,+\infty\right[$. Posons*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n\big(J^{W_n}_{\gamma_n A} x_n + a_n - x_n\big). \tag{5.16}$$

Alors, pour un point $\overline{x} \in C$, nous avons les propriétés suivantes.

(i) $x_n \rightharpoonup \overline{x}$.

(ii) *Supposons que $\operatorname{int} C \neq \varnothing$ et il existe $(\nu_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N})$ telle que $(\forall n \in \mathbb{N})\ (1+\nu_n)W_{n+1} \succcurlyeq W_n$. Alors $x_n \to \overline{x}$.*

(iii) *Supposons que A soit uniformément monotone en $\overline{x}$. Alors $x_n \to \overline{x}$.*

Nous présentons une application de l'algorithme proximal à métrique variable (5.16) à un problème inverse.

Corollaire 5.11 *Soit $f \in \Gamma_0(\mathcal{H})$ et soit I un ensemble fini d'index non vide. Pour tout $i \in I$, soit $(\mathcal{G}_i, \|\cdot\|_i)$ un espace hilbertien réel, soit $0 \neq L_i \in \mathcal{B}(\mathcal{H},\mathcal{G}_i)$, soit $r_i \in \mathcal{G}_i$, et soit $\mu_i \in \left]0,+\infty\right[$. Considérons le problème*

$$\underset{x\in\mathcal{H}}{\text{minimiser}}\ f(x) + \frac{1}{2}\sum_{i\in I} \mu_i \|L_i x - r_i\|_i^2. \tag{5.17}$$

Soit $\varepsilon \in \left]0, 1/(1+\sum_{i\in I}\mu_i\|L_i\|^2)\right[$, soit $(a_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{H}$ telle que $\sum_{n\in\mathbb{N}}\|a_n\| < +\infty$, soit $(\eta_n)_{n\in\mathbb{N}}$ une suite dans $\ell^1_+(\mathbb{N})$, et soit $(\gamma_n)_{n\in\mathbb{N}}$ une suite dans $\mathbb{R}$ telle que

$$(\forall n \in \mathbb{N}) \quad \varepsilon \leqslant \gamma_n \leqslant \frac{1-\varepsilon}{\sum_{i\in I}\mu_i\|L_i\|^2} \quad \text{et} \quad (1+\eta_n)\gamma_n - \gamma_{n+1} \leqslant \frac{\eta_n}{\sum_{i\in I}\mu_i\|L_i\|^2}. \tag{5.18}$$

De plus, soient C un ensemble de solutions du problème (5.17), $x_0 \in \mathcal{H}$, $(\lambda_n)_{n\in\mathbb{N}}$ une suite dans $[\varepsilon, 2-\varepsilon]$, et posons

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n\left(\operatorname{prox}_{\gamma_n f}\left(x_n + \gamma_n \sum_{i\in I}\mu_i L_i^*(r_i - L_i x_n)\right) + a_n - x_n\right). \tag{5.19}$$

Alors on a les résultats suivants pour un point $\overline{x} \in C$.

(i) *Supposons que*

$$\lim_{\|x\|\to+\infty} f(x) + \frac{1}{2}\sum_{i\in I}\mu_i\|L_i x - r_i\|_i^2 = +\infty. \tag{5.20}$$

Alors $x_n \rightharpoonup \overline{x}$.

(ii) *Supposons qu'il existe $j \in I$ tel que l'opérateur L_j vérifie*

$$(\exists\, \beta \in\,]0,+\infty[)(\forall x \in \mathcal{H}) \quad \|L_j x\|_j \geqslant \beta\|x\|. \tag{5.21}$$

Alors $C = \{\overline{x}\}$ et $x_n \to \overline{x}$.

5.2 Article en anglais

VARIABLE METRIC QUASI-FEJÉR MONOTONICITY [1]

Abstract : The notion of quasi-Fejér monotonicity has proven to be an efficient tool to simplify and unify the convergence analysis of various algorithms arising in applied nonlinear analysis. In this paper, we extend this notion in the context of variable metric algorithms, whereby the underlying norm is allowed to vary at each iteration. Applications to convex feasibility problems are demonstrated.

Let C be a nonempty closed subset of the Euclidean space $\mathbb{R}^N$ and let y be a point in its complement. In 1922, Fejér [21] considered the problem of finding a point $x \in \mathbb{R}^N$ such that $(\forall z \in C)\ \|x - z\| < \|y - z\|$. Based on this work, the term Fejér-monotonicity was coined in [27] in connection with sequences $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}^N$ that satisfy

$$(\forall z \in C)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - z\| \leqslant \|x_n - z\|. \tag{5.22}$$

This concept was later broadened to that of quasi-Fejér monotonicity in [20] by relaxing (5.22) to

$$(\forall z \in C)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - z\|^2 \leqslant \|x_n - z\|^2 + \varepsilon_n, \tag{5.23}$$

where $(\varepsilon_n)_{n\in\mathbb{N}}$ is a summable sequence in $[0,+\infty[$. These notions have proven to be remarkably useful in simplifying and unifying the convergence analysis of a large collection of algorithms arising in hilbertian nonlinear analysis, see for instance [2, 5, 12, 13, 14, 18, 19, 30, 31, 35] and the references therein. In recent years, there have been attempts to generalize standard algorithms such as those discussed in the above references by allowing the underlying metric to vary over the course of the iterations, e.g., [7, 10, 11, 16, 26, 29]. In order to better understand the convergence properties of such algorithms and lay the ground for further developments, we extend in the present paper the notion of quasi-Fejér monotonicity to the context of variable metric iterations in general Hilbert spaces and investigate its properties.

Our notation and preliminary results are presented in Section 5.3. The notion of variable metric quasi-Fejér monotonicity is introduced in Section 5.4, where weak and

1. P. L. Combettes and B. C. Vũ, Variable Metric Quasi-Fejér Monotonicity, *Nonlinear Analysis : Theory, Methods, and Applications*, vol. 78, pp. 17–31, 2013.

strong convergence results are also established. In Section 5.5, we focus on the special case when, as in (5.23), monotonicity is with respect to the squared norms. Finally, we illustrate the potential of these tools in the analysis of variable metric convex feasibility algorithms in Section 5.6 and in the design of algorithms for solving inverse problems in Section 5.7.

5.3 Notation and technical facts

Throughout, $\mathcal{H}$ is a real Hilbert space, $\langle\cdot\mid\cdot\rangle$ is its scalar product and $\|\cdot\|$ the associated norm. The symbols $\rightharpoonup$ and $\to$ denote respectively weak and strong convergence, Id denotes the identity operator, and $B(z;\rho)$ denotes the closed ball of center $z\in\mathcal{H}$ and radius $\rho\in\,]0,+\infty[$; $\mathcal{S}(\mathcal{H})$ is the space of self-adjoint bounded linear operators from $\mathcal{H}$ to $\mathcal{H}$. The Loewner partial ordering on $\mathcal{S}(\mathcal{H})$ is defined by

$$(\forall L_1\in\mathcal{S}(\mathcal{H}))(\forall L_2\in\mathcal{S}(\mathcal{H}))\quad L_1\succcurlyeq L_2\quad\Leftrightarrow\quad(\forall x\in\mathcal{H})\quad\langle L_1x\mid x\rangle\geqslant\langle L_2x\mid x\rangle.\tag{5.24}$$

Now let $\alpha\in[0,+\infty[$, set

$$\mathcal{P}_\alpha(\mathcal{H})=\big\{L\in\mathcal{S}(\mathcal{H})\;\big|\;L\succcurlyeq\alpha\,\mathrm{Id}\big\},\tag{5.25}$$

and fix $W\in\mathcal{P}_\alpha(\mathcal{H})$. We define a semi-scalar product and a semi-norm (a scalar product and a norm if $\alpha>0$) by

$$(\forall x\in\mathcal{H})(\forall y\in\mathcal{H})\quad\langle x\mid y\rangle_W=\langle Wx\mid y\rangle\quad\text{and}\quad\|x\|_W=\sqrt{\langle Wx\mid x\rangle}.\tag{5.26}$$

Let C be a nonempty subset of $\mathcal{H}$, let $\alpha\in\,]0,+\infty[$, and let $W\in\mathcal{P}_\alpha(\mathcal{H})$. The interior of C is $\operatorname{int}C$, the distance function of C is d_C, and the convex envelope of C is $\operatorname{conv}C$, with closure $\overline{\operatorname{conv}}\,C$. If C is closed and convex, the projection operator onto C relative to the metric induced by W in (5.26) is

$$P_C^W\colon\mathcal{H}\to C\colon x\mapsto\underset{y\in C}{\operatorname{argmin}}\;\|x-y\|_W.\tag{5.27}$$

We write $P_C^{\mathrm{Id}}=P_C$. Finally, $\ell_+^1(\mathbb{N})$ denotes the set of summable sequences in $[0,+\infty[$.

Lemma 5.12 *Let $\alpha\in\,]0,+\infty[$, let $\mu\in\,]0,+\infty[$, and let A and B be operators in $\mathcal{S}(\mathcal{H})$ such that $\mu\,\mathrm{Id}\succcurlyeq A\succcurlyeq B\succcurlyeq\alpha\,\mathrm{Id}$. Then the following hold.*

(i) $\alpha^{-1}\,\mathrm{Id}\succcurlyeq B^{-1}\succcurlyeq A^{-1}\succcurlyeq\mu^{-1}\,\mathrm{Id}$.

(ii) $(\forall x\in\mathcal{H})\;\langle A^{-1}x\mid x\rangle\geqslant\|A\|^{-1}\|x\|^2$.

(iii) $\|A^{-1}\|\leqslant\alpha^{-1}$.

Proof. These facts are known [24, Section VI.2.6]. We provide a simple convex-analytic proof.

(i) : It suffices to show that $B^{-1} \succcurlyeq A^{-1}$. Set $(\forall x \in \mathcal{H})\ f(x) = \langle Ax \mid x\rangle/2$ and $g(x) = \langle Bx \mid x\rangle/2$. The conjugate of f is $f^*\colon \mathcal{H} \to [-\infty,+\infty] : u \mapsto \sup_{x\in\mathcal{H}} \big(\langle x \mid u\rangle - f(x)\big) = \langle A^{-1}u \mid u\rangle/2$ [5, Proposition 17.28]. Likewise, $g^*\colon \mathcal{H} \to [-\infty,+\infty] : u \mapsto \langle B^{-1}u \mid u\rangle/2$. Since, $f \geqslant g$, we have $g^* \geqslant f^*$, hence the result.

(ii) : Since $\|A\|\,\mathrm{Id} \succcurlyeq A$, (i) yields $A^{-1} \succcurlyeq \|A\|^{-1}\,\mathrm{Id}$.

(iii) : We have $A^{-1} \in \mathcal{S}(\mathcal{H})$ and, by (i), $(\forall x \in \mathcal{H})\ \|x\|^2/\alpha \geqslant \langle A^{-1}x \mid x\rangle$. Hence, upon taking the supremum over $B(0;1)$, we obtain $1/\alpha \geqslant \|A^{-1}\|$. □

Lemma 5.13 [30, Lemma 2.2.2] *Let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence in $[0,+\infty[$, let $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$, and let $(\varepsilon_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$ be such that $(\forall n \in \mathbb{N})\ \alpha_{n+1} \leqslant (1+\eta_n)\alpha_n + \varepsilon_n$. Then $(\alpha_n)_{n\in\mathbb{N}}$ converges.*

The following lemma extends the classical property that a uniformly bounded monotone sequence of operators in $\mathcal{S}(\mathcal{H})$ converges pointwise [33, Théorème 104.1].

Lemma 5.14 *Let $\alpha \in\,]0,+\infty[$, let $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$, and let $(W_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{H})$ such that $\mu = \sup_{n\in\mathbb{N}} \|W_n\| < +\infty$. Suppose that one of the following holds.*

(i) $(\forall n \in \mathbb{N})\ (1+\eta_n)W_n \succcurlyeq W_{n+1}$.

(ii) $(\forall n \in \mathbb{N})\ (1+\eta_n)W_{n+1} \succcurlyeq W_n$.

Then there exists $W \in \mathcal{P}_\alpha(\mathcal{H})$ such that $W_n \to W$ pointwise.

Proof. (i) : Set $\tau = \prod_{n\in\mathbb{N}}(1+\eta_n)$, $\tau_0 = 1$, and, for every $n \in \mathbb{N} \smallsetminus \{0\}$, $\tau_n = \prod_{k=0}^{n-1}(1+\eta_k)$. Then $\tau_n \to \tau < +\infty$ [25, Theorem 3.7.3] and

$$(\forall n \in \mathbb{N}) \quad \mu\,\mathrm{Id} \succcurlyeq W_n \succcurlyeq \alpha\,\mathrm{Id} \quad \text{and} \quad \tau_{n+1} = \tau_n(1+\eta_n). \tag{5.28}$$

Now define

$$(\forall n \in \mathbb{N})(\forall m \in \mathbb{N}) \quad W_{n,m} = \frac{1}{\tau_n}W_n - \frac{1}{\tau_{n+m}}W_{n+m}. \tag{5.29}$$

Then we derive from (5.28) that $(\forall n \in \mathbb{N})(\forall m \in \mathbb{N}\smallsetminus\{0\})(\forall x \in \mathcal{H})$

$$\begin{aligned} 0 &= \frac{1}{\tau_n}\langle W_n x \mid x\rangle - \frac{1}{\tau_{n+m}}\prod_{k=n}^{n+m-1}(1+\eta_k)\langle W_n x \mid x\rangle \\ &\leqslant \frac{1}{\tau_n}\langle W_n x \mid x\rangle - \frac{1}{\tau_{n+m}}\langle W_{n+m} x \mid x\rangle \\ &= \langle W_{n,m} x \mid x\rangle \\ &\leqslant \frac{1}{\tau_n}\langle W_n x \mid x\rangle \\ &\leqslant \langle W_n x \mid x\rangle \\ &\leqslant \mu\|x\|^2. \end{aligned} \tag{5.30}$$

Therefore

$$(\forall n \in \mathbb{N})(\forall m \in \mathbb{N}) \quad W_{n,m} \in \mathcal{P}_0(\mathcal{H}) \quad \text{and} \quad \|W_{n,m}\| \leqslant \mu. \tag{5.31}$$

Let us fix $x \in \mathcal{H}$. By assumption, $(\forall n \in \mathbb{N})$ $\|x\|^2_{W_{n+1}} \leqslant (1+\eta_n)\|x\|^2_{W_n}$. Hence, by Lemma 5.13, $(\|x\|^2_{W_n})_{n\in\mathbb{N}}$ converges. In turn, $(\tau_n^{-1}\|x\|^2_{W_n})_{n\in\mathbb{N}}$ converges, which implies that

$$\|x\|^2_{W_{n,m}} = \langle W_{n,m}x \mid x\rangle = \frac{1}{\tau_n}\|x\|^2_{W_n} - \frac{1}{\tau_{n+m}}\|x\|^2_{W_{n+m}} \to 0 \quad \text{as} \quad n, m \to +\infty. \tag{5.32}$$

Therefore, using (5.31), Cauchy-Schwarz for the semi-norms $(\|\cdot\|_{W_{n,m}})_{(n,m)\in\mathbb{N}^2}$, and (5.32), we obtain

$$\begin{aligned} \|W_{n,m}x\|^4 &= \langle x \mid W_{n,m}x\rangle^2_{W_{n,m}} \\ &\leqslant \|x\|^2_{W_{n,m}} \|W_{n,m}x\|^2_{W_{n,m}} \\ &\leqslant \|x\|^2_{W_{n,m}} \mu^3\|x\|^2 \\ &\to 0 \quad \text{as} \quad n, m \to +\infty. \end{aligned} \tag{5.33}$$

Thus, we derive from (5.29) that $(\tau_n^{-1}W_nx)_{n\in\mathbb{N}}$ is a Cauchy sequence. Hence, it converges strongly, and so does $(W_nx)_{n\in\mathbb{N}}$. If we call Wx the limit of $(W_nx)_{n\in\mathbb{N}}$, the above construction yields the desired operator $W \in \mathcal{P}_\alpha(\mathcal{H})$.

(ii) : Set $(\forall n \in \mathbb{N})$ $L_n = W_n^{-1}$. It follows from Lemma 5.12(i) et (iii) that $(L_n)_{n\in\mathbb{N}}$ lies in $\mathcal{P}_{1/\mu}(\mathcal{H})$, $\sup_{n\in\mathbb{N}}\|L_n\| \leqslant 1/\alpha$, and $(\forall n \in \mathbb{N})$ $(1+\eta_n)L_n \succcurlyeq L_{n+1}$. Hence, appealing to (i), there exists $L \in \mathcal{P}_{1/\mu}(\mathcal{H})$ such that $\|L\| \leqslant 1/\alpha$ and $L_n \to L$ pointwise. Now let $x \in \mathcal{H}$, and set $W = L^{-1}$ and $(\forall n \in \mathbb{N})$ $x_n = L_n(Wx)$. Then $W \in \mathcal{P}_\alpha(\mathcal{H})$ and $x_n \to L(Wx) = x$. Moreover, $\|W_nx - Wx\| = \|W_n(x - x_n)\| \leqslant \mu\|x_n - x\| \to 0$. □

5.4 Variable metric quasi-Fejér monotone sequences

Our paper hinges on the following extension of (5.23).

Definition 5.15 Let $\alpha \in \left]0, +\infty\right[$, let $\phi\colon \left[0, +\infty\right[\to \left[0, +\infty\right[$, let $(W_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{H})$, let C be a nonempty subset of $\mathcal{H}$, and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$. Then $(x_n)_{n\in\mathbb{N}}$ is :

(i) ϕ-quasi-Fejér monotone with respect to the target set C relative to $(W_n)_{n\in\mathbb{N}}$ if

$$\big(\exists\,(\eta_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N})\big)\big(\forall z \in C\big)\big(\exists\,(\varepsilon_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N})\big)(\forall n \in \mathbb{N}) \\ \phi(\|x_{n+1} - z\|_{W_{n+1}}) \leqslant (1+\eta_n)\phi(\|x_n - z\|_{W_n}) + \varepsilon_n; \tag{5.34}$$

(ii) stationarily ϕ-quasi-Fejér monotone with respect to the target set C relative to $(W_n)_{n\in\mathbb{N}}$ if

$$\big(\exists\,(\varepsilon_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})\big)\big(\exists\,(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})\big)(\forall z \in C)(\forall n \in \mathbb{N}) \\ \phi(\|x_{n+1} - z\|_{W_{n+1}}) \leqslant (1 + \eta_n)\phi(\|x_n - z\|_{W_n}) + \varepsilon_n. \quad (5.35)$$

We start with basic properties.

Proposition 5.16 *Let $\alpha \in \left]0, +\infty\right[$, let $\phi\colon \left[0, +\infty\right[\to \left[0, +\infty\right[$ be strictly increasing and such that $\lim_{t\to+\infty} \phi(t) = +\infty$, let $(W_n)_{n\in\mathbb{N}}$ be in $\mathcal{P}_\alpha(\mathcal{H})$, let C be a nonempty subset of $\mathcal{H}$, and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that (5.34) is satisfied. Then the following hold.*

(i) *Let $z \in C$. Then $(\|x_n - z\|_{W_n})_{n\in\mathbb{N}}$ converges.*

(ii) *$(x_n)_{n\in\mathbb{N}}$ is bounded.*

Proof. (i) : Set $(\forall n \in \mathbb{N})\ \xi_n = \|x_n - z\|_{W_n}$. It follows from (5.34) and Lemma 5.13 that $(\phi(\xi_n))_{n\in\mathbb{N}}$ converges, say $\phi(\xi_n) \to \lambda$. In turn, since $\lim_{t\to+\infty} \phi(t) = +\infty$, $(\xi_n)_{n\in\mathbb{N}}$ is bounded and, to show that it converges, it suffices to show that it cannot have two distinct cluster points. Suppose to the contrary that we can extract two subsequences $(\xi_{k_n})_{n\in\mathbb{N}}$ and $(\xi_{l_n})_{n\in\mathbb{N}}$ such that $\xi_{k_n} \to \eta$ and $\xi_{l_n} \to \zeta > \eta$, and fix $\varepsilon \in \left]0, (\zeta - \eta)/2\right[$. Then, for n sufficiently large, $\xi_{k_n} \leqslant \eta + \varepsilon < \zeta - \varepsilon \leqslant \xi_{l_n}$ and, since ϕ is strictly increasing, $\phi(\xi_{k_n}) \leqslant \phi(\eta + \varepsilon) < \phi(\zeta - \varepsilon) \leqslant \phi(\xi_{l_n})$. Taking the limit as $n \to +\infty$ yields $\lambda \leqslant \phi(\eta + \varepsilon) < \phi(\zeta - \varepsilon) \leqslant \lambda$, which is impossible.

(ii) : Let $z \in C$. Since $(W_n)_{n\in\mathbb{N}}$ lies in $\mathcal{P}_\alpha(\mathcal{H})$, we have

$$(\forall n \in \mathbb{N}) \quad \alpha\|x_n - z\|^2 \leqslant \langle x_n - z \mid W_n(x_n - z)\rangle = \|x_n - z\|^2_{W_n}. \quad (5.36)$$

Hence, since (i) asserts that $(\|x_n - z\|_{W_n})_{n\in\mathbb{N}}$ is bounded, so is $(x_n)_{n\in\mathbb{N}}$. ☐

The next result concerns weak convergence. In the case of standard Fejér monotonicity (5.22), it appears in [9, Lemma 6] and, in the case of quasi-Fejér monotonicity (5.23), it appears in [1, Proposition 1.3].

Theorem 5.17 *Let $\alpha \in \left]0, +\infty\right[$, let $\phi\colon \left[0, +\infty\right[\to \left[0, +\infty\right[$ be strictly increasing and such that $\lim_{t\to+\infty} \phi(t) = +\infty$, let $(W_n)_{n\in\mathbb{N}}$ and W be operators in $\mathcal{P}_\alpha(\mathcal{H})$ such that $W_n \to W$ pointwise, let C be a nonempty subset of $\mathcal{H}$, and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that (5.34) is satisfied. Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in C if and only if every weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ is in C.*

Proof. Necessity is clear. To show sufficiency, suppose that every weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ is in C, and let x and y be two such points, say $x_{k_n} \rightharpoonup x$ and $x_{l_n} \rightharpoonup y$. Then it follows from Proposition 5.16(i) that $(\|x_n - x\|_{W_n})_{n\in\mathbb{N}}$ and $(\|x_n - y\|_{W_n})_{n\in\mathbb{N}}$ converge. Moreover, $\|x\|^2_{W_n} = \langle W_n x \mid x\rangle \to \langle Wx \mid x\rangle$ and, likewise, $\|y\|^2_{W_n} \to \langle Wy \mid y\rangle$. Therefore, since

$$(\forall n \in \mathbb{N}) \quad \langle W_n x_n \mid x - y\rangle = \frac{1}{2}\big(\|x_n - y\|^2_{W_n} - \|x_n - x\|^2_{W_n} + \|x\|^2_{W_n} - \|y\|^2_{W_n}\big), \quad (5.37)$$

the sequence $(\langle W_n x_n \mid x-y\rangle)_{n\in\mathbb{N}}$ converges, say $\langle W_n x_n \mid x-y\rangle \to \lambda \in \mathbb{R}$, which implies that

$$\langle x_n \mid W_n(x-y)\rangle \to \lambda \in \mathbb{R}. \tag{5.38}$$

However, since $x_{k_n} \rightharpoonup x$ and $W_{k_n}(x-y) \to W(x-y)$, it follows from (5.38) and [5, Lemma 2.41(iii)] that $\langle x \mid W(x-y)\rangle = \lambda$. Likewise, passing to the limit along the subsequence $(x_{l_n})_{n\in\mathbb{N}}$ in (5.38) yields $\langle y \mid W(x-y)\rangle = \lambda$. Thus,

$$0 = \langle x \mid W(x-y)\rangle - \langle y \mid W(x-y)\rangle = \langle x-y \mid W(x-y)\rangle \geqslant \alpha\|x-y\|^2. \tag{5.39}$$

This shows that $x = y$. Upon invoking Proposition 5.16(ii) and [5, Lemma 2.38], we conclude that $x_n \rightharpoonup x$. ☐

Lemma 5.14 provides instances in which the conditions imposed on $(W_n)_{n\in\mathbb{N}}$ in Theorem 5.17 are satisfied. Next, we present a characterization of strong convergence which can be found in [12, Theorem 3.11] in the special case of quasi-Fejér monotonicity (5.23).

Proposition 5.18 *Let $\alpha \in\,]0,+\infty[$, let $\chi \in [1,+\infty[$, and let $\phi\colon [0,+\infty[\, \to [0,+\infty[$ be an increasing upper semicontinuous function vanishing only at 0 and such that*

$$\big(\forall (\xi_1,\xi_2) \in [0,+\infty[^2\big) \quad \phi(\xi_1+\xi_2) \leqslant \chi\big(\phi(\xi_1)+\phi(\xi_2)\big). \tag{5.40}$$

Let $(W_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{H})$ such that $\mu = \sup_{n\in\mathbb{N}}\|W_n\| < +\infty$, let C be a nonempty closed subset of $\mathcal{H}$, and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that (5.35) is satisfied. Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to a point in C if and only if $\underline{\lim}\, d_C(x_n) = 0$.

Proof. Necessity is clear. For sufficiency, suppose that $\underline{\lim}\, d_C(x_n) = 0$ and set $(\forall n \in \mathbb{N})$ $\xi_n = \inf_{z\in C}\|x_n - z\|_{W_n}$. For every $n \in \mathbb{N}$, let $(z_{n,k})_{k\in\mathbb{N}}$ be a sequence in C such that $\|x_n - z_{n,k}\|_{W_n} \to \xi_n$. Then, since ϕ is increasing, (5.35) yields

$$(\forall n \in \mathbb{N})(\forall k \in \mathbb{N}) \quad \phi(\xi_{n+1}) \leqslant \phi(\|x_{n+1}-z_{n,k}\|_{W_{n+1}}) \leqslant (1+\eta_n)\phi(\|x_n-z_{n,k}\|_{W_n})+\varepsilon_n. \tag{5.41}$$

Hence, it follows from the upper semicontinuity of ϕ that

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \phi(\xi_{n+1}) &\leqslant (1+\eta_n)\varlimsup_{k\to+\infty}\phi(\|x_n - z_{n,k}\|_{W_n}) + \varepsilon_n \\ &\leqslant (1+\eta_n)\phi(\xi_n) + \varepsilon_n. \end{aligned} \tag{5.42}$$

Therefore, by Lemma 5.13,

$$\big(\phi(\xi_n)\big)_{n\in\mathbb{N}} \quad \text{converges.} \tag{5.43}$$

Moreover, since

$$(\forall n \in \mathbb{N})(\forall m \in \mathbb{N})(\forall x \in \mathcal{H}) \quad \alpha\|x_n - x\|^2 \leqslant \|x_n - x\|^2_{W_m} \leqslant \mu\|x_n - x\|^2, \tag{5.44}$$

we have

$$(\forall n \in \mathbb{N}) \quad \sqrt{\alpha} d_C(x_n) \leqslant \xi_n \leqslant \sqrt{\mu} d_C(x_n). \tag{5.45}$$

Consequently, since $\underline{\lim}\, d_C(x_n) = 0$, we derive from (5.45) that $\underline{\lim}\, \xi_n = 0$. Let us extract a subsequence $(\xi_{k_n})_{n\in\mathbb{N}}$ such that $\xi_{k_n} \to 0$. Since ϕ is upper semicontinuous, we have $0 \leqslant \underline{\lim}\, \phi(\xi_{k_n}) \leqslant \overline{\lim}\, \phi(\xi_{k_n}) \leqslant \phi(0) = 0$. In view of (5.43), we therefore obtain $\phi(\xi_n) \to 0$ and, in turn, $\xi_n \to 0$. Hence, we deduce from (5.45) that

$$d_C(x_n) \to 0. \tag{5.46}$$

Next, let N be the smallest integer such that $N > \sqrt{\mu}$, and set $\rho = \chi^{N-1} + \sum_{k=1}^{N-1} \chi^k$ if $N > 1$; $\rho = 1$ if $N = 1$. Moreover, let $x \in C$ and let m and n be strictly positive integers. Using (5.44), the monotonicity of ϕ, and (5.40), we obtain

$$\phi\big(\|x_n - x\|_{W_m}\big) \leqslant \phi\big(\sqrt{\mu}\|x_n - x\|\big) \leqslant \phi\big(N\|x_n - x\|\big) \leqslant \rho\phi\big(\|x_n - x\|\big). \tag{5.47}$$

Now set $\tau = \prod_{k\in\mathbb{N}}(1 + \eta_k)$. Then $\tau < +\infty$ [25, Theorem 3.7.3] and we derive from (5.40), (5.35), and (5.47) that

$$\begin{aligned}
\chi^{-1}\phi\big(\|x_{n+m} - x_n\|_{W_{n+m}}\big) &\leqslant \chi^{-1}\phi\big(\|x_{n+m} - x\|_{W_{n+m}} + \|x_n - x\|_{W_{n+m}}\big) \\
&\leqslant \phi\big(\|x_{n+m} - x\|_{W_{m+n}}\big) + \phi\big(\|x_n - x\|_{W_{m+n}}\big) \\
&\leqslant \tau\left(\phi\big(\|x_n - x\|_{W_n}\big) + \sum_{k=n}^{n+m-1} \varepsilon_k\right) + \phi\big(\|x_n - x\|_{W_{m+n}}\big) \\
&\leqslant \rho(1+\tau)\phi\big(\|x_n - x\|\big) + \tau \sum_{k\geqslant n} \varepsilon_k.
\end{aligned} \tag{5.48}$$

Therefore, upon taking the infimum over $x \in C$, we obtain by upper semicontinuity of ϕ

$$\phi\big(\|x_{n+m} - x_n\|_{W_{n+m}}\big) \leqslant \chi\rho(1+\tau)\phi\big(d_C(x_n)\big) + \chi\tau \sum_{k\geqslant n} \varepsilon_k. \tag{5.49}$$

Hence, appealing to (5.46) and the summability of $(\varepsilon_k)_{k\in\mathbb{N}}$, we deduce from (5.49) that, as $n \to +\infty$, $\phi(\|x_{n+m} - x_n\|_{W_{n+m}}) \to 0$ and, hence, $\alpha\|x_{n+m} - x_n\|^2 \leqslant \|x_{n+m} - x_n\|^2_{W_{n+m}} \to 0$. Thus, $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{H}$ and there exists $\overline{x} \in \mathcal{H}$ such that $x_n \to \overline{x}$. By continuity of d_C and (5.46), we obtain $d_C(\overline{x}) = 0$ and, since C is closed, $\overline{x} \in C$. □

5.5 The quadratic case

In this section, we focus on the important case when $\phi = |\cdot|^2$ in Definition 5.15. Our first result states that variable metric quasi-Fejér monotonicity "spreads" to the convex hull of the target set.

Proposition 5.19 *Let $\alpha \in]0, +\infty[$, let $(\eta_n)_{n\in\mathbb{N}}$ be a sequence in $\ell_+^1(\mathbb{N})$, let $(W_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{H})$ such that*

$$\mu = \sup_{n\in\mathbb{N}} \|W_n\| < +\infty \quad \textit{and} \quad (\forall n \in \mathbb{N}) \quad (1+\eta_n)W_n \succcurlyeq W_{n+1}. \tag{5.50}$$

Let C be a nonempty subset of $\mathcal{H}$ and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that

$$\big(\exists\, (\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})\big)\big(\forall z \in C\big)\big(\exists\, (\varepsilon_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})\big)(\forall n \in \mathbb{N})$$
$$\|x_{n+1} - z\|^2_{W_{n+1}} \leqslant (1+\eta_n)\|x_n - z\|^2_{W_n} + \varepsilon_n. \tag{5.51}$$

Then the following hold.

(i) *$(x_n)_{n\in\mathbb{N}}$ is $|\cdot|^2$-quasi-Fejér monotone with respect to $\operatorname{conv} C$ relative to $(W_n)_{n\in\mathbb{N}}$.*

(ii) *For every $y \in \overline{\operatorname{conv}}\, C$, $(\|x_n - y\|_{W_n})_{n\in\mathbb{N}}$ converges.*

Proof. Let us fix $z \in \operatorname{conv} C$. There exist finite sets $\{z_i\}_{i\in I} \subset C$ and $\{\lambda_i\}_{i\in I} \subset]0,1]$ such that

$$\sum_{i\in I} \lambda_i = 1 \quad \text{and} \quad z = \sum_{i\in I} \lambda_i z_i. \tag{5.52}$$

For every $i \in I$, it follows from (5.51) that there exists a sequence $(\varepsilon_{i,n})_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$ such that

$$(\forall n \in \mathbb{N}) \quad \|x_{n+1} - z_i\|^2_{W_{n+1}} \leqslant (1+\eta_n)\|x_n - z_i\|^2_{W_n} + \varepsilon_{i,n}. \tag{5.53}$$

Now set

$$(\forall n \in \mathbb{N}) \quad \begin{cases} \alpha_n = \dfrac{1}{2}\displaystyle\sum_{i\in I}\sum_{j\in I} \lambda_i\lambda_j \|z_i - z_j\|^2_{W_n} \\ \varepsilon_n = (1+\eta_n)\alpha_n - \alpha_{n+1} + \max\{\varepsilon_{1,n}, \ldots, \varepsilon_{m,n}\}. \end{cases} \tag{5.54}$$

Then $(\max\{\varepsilon_{1,n}, \ldots, \varepsilon_{m,n}\})_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$ and, by (5.50), $(\forall n \in \mathbb{N})$ $(1+\eta_n)\alpha_n \geqslant \alpha_{n+1}$. Hence, Lemma 5.13 asserts that $(\alpha_n)_{n\in\mathbb{N}}$ converges, which implies that $(\varepsilon_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$.

(i) : Using (5.52), [5, Lemma 2.13(ii)], and (5.53), we obtain

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \|x_{n+1} - z\|^2_{W_{n+1}} &= \sum_{i\in I} \lambda_i \|x_{n+1} - z_i\|^2_{W_{n+1}} - \alpha_{n+1} \\ &\leqslant (1+\eta_n)\sum_{i\in I} \lambda_i \|x_n - z_i\|^2_{W_n} - \alpha_{n+1} + \max_{1\leqslant i\leqslant m}\{\varepsilon_{i,n}\} \\ &= (1+\eta_n)\|x_n - z\|^2_{W_n} + (1+\eta_n)\alpha_n - \alpha_{n+1} \\ &\quad + \max_{1\leqslant i\leqslant m}\{\varepsilon_{i,n}\} \\ &= (1+\eta_n)\|x_n - z\|^2_{W_n} + \varepsilon_n. \end{aligned} \tag{5.55}$$

(ii) : It follows from [5, Lemma 2.13(ii)] that

$$(\forall n \in \mathbb{N}) \quad \|x_n - z\|_{W_n}^2 = \sum_{i\in I} \lambda_i \|x_n - z_i\|_{W_n}^2 - \alpha_n. \tag{5.56}$$

However, $(\alpha_n)_{n\in\mathbb{N}}$ converges and, for every $i \in I$, Proposition 5.16(i) asserts that $(\|x_n - z_i\|_{W_n})_{n\in\mathbb{N}}$ converges. Hence, $(\|x_n - z\|_{W_n})_{n\in\mathbb{N}}$ converges. Now let $y \in \overline{\operatorname{conv}}\, C$. Then there exists a sequence $(y_k)_{k\in\mathbb{N}}$ in $\operatorname{conv} C$ such that $y_k \to y$. It follows from (i) and Proposition 5.16(i) that, for every $k \in \mathbb{N}$, $(\|x_n - y_k\|_{W_n})_{n\in\mathbb{N}}$ converges. Moreover, we have

$$\begin{aligned}(\forall k \in \mathbb{N})(\forall n \in \mathbb{N}) \quad -\sqrt{\mu}\|y_k - y\| &\leqslant -\|y_k - y\|_{W_n}\\ &\leqslant \|x_n - y\|_{W_n} - \|x_n - y_k\|_{W_n}\\ &\leqslant \|y_k - y\|_{W_n}\\ &\leqslant \sqrt{\mu}\|y_k - y\|. \end{aligned}\tag{5.57}$$

Consequently,

$$\begin{aligned}(\forall k \in \mathbb{N}) \quad -\sqrt{\mu}\|y_k - y\| &\leqslant \underline{\lim}\, \|x_n - y\|_{W_n} - \lim \|x_n - y_k\|_{W_n}\\ &\leqslant \overline{\lim}\, \|x_n - y\|_{W_n} - \lim \|x_n - y_k\|_{W_n}\\ &\leqslant \sqrt{\mu}\|y_k - y\|. \end{aligned}\tag{5.58}$$

Taking the limit as $k \to +\infty$ yields $\lim_{n\to+\infty} \|x_n - y\|_{W_n} = \lim_{k\to+\infty} \lim_{n\to+\infty} \|x_n - y_k\|_{W_n}$. ☐

Standard Fejér monotone sequences may fail to converge weakly and, even when they converge weakly, strong convergence may fail [12, 23]. However, if the target set C is closed and convex in (5.22), the projected sequence $(P_C x_n)_{n\in\mathbb{N}}$ converges strongly ; see [2, Theorem 2.16(iv)] and [32, Remark 1]. This property, which remains true in the quasi-Fejérian case [12, Proposition 3.6(iv)], is extended below.

Proposition 5.20 *Let $\alpha \in \left]0, +\infty\right[$, let $(\eta_n)_{n\in\mathbb{N}}$ be a sequence in $\ell_+^1(\mathbb{N})$, let $(W_n)_{n\in\mathbb{N}}$ be a uniformly bounded sequence in $\mathcal{P}_\alpha(\mathcal{H})$, let C be a nonempty closed convex subset of $\mathcal{H}$, and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that*

$$\begin{aligned}&\big(\exists\, (\varepsilon_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})\big)\big(\exists\, (\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})\big)(\forall z \in C)(\forall n \in \mathbb{N})\\ &\qquad\qquad \|x_{n+1} - z\|_{W_{n+1}}^2 \leqslant (1 + \eta_n)\|x_n - z\|_{W_n}^2 + \varepsilon_n. \end{aligned}\tag{5.59}$$

Then $(P_C^{W_n} x_n)_{n\in\mathbb{N}}$ converges strongly.

Proof. Set $(\forall n \in \mathbb{N})$ $z_n = P_C^{W_n} x_n$. For every $(m, n) \in \mathbb{N}^2$, since $z_n \in C$ and $z_{m+n} = P_C^{W_{n+m}} x_{n+m}$, the well-known convex projection theorem [5, Theorem 3.14] yields

$$\langle z_n - z_{n+m} \mid x_{n+m} - z_{n+m}\rangle_{W_{n+m}} \leqslant 0, \tag{5.60}$$

which implies that

$$\begin{aligned}\langle z_n - x_{n+m} \mid x_{n+m} - z_{n+m}\rangle_{W_{n+m}} &= \langle z_n - z_{n+m} \mid x_{n+m} - z_{n+m}\rangle_{W_{n+m}} \\ &\quad - \|x_{n+m} - z_{n+m}\|^2_{W_{n+m}} \\ &\leqslant -\|x_{n+m} - z_{n+m}\|^2_{W_{n+m}}. \end{aligned} \tag{5.61}$$

Therefore, for every $(m,n) \in \mathbb{N}^2$,

$$\begin{aligned}\|z_n - z_{n+m}\|^2_{W_{n+m}} &= \|z_n - x_{n+m}\|^2_{W_{n+m}} + 2\langle z_n - x_{n+m} \mid x_{n+m} - z_{n+m}\rangle_{W_{n+m}} \\ &\quad + \|x_{n+m} - z_{n+m}\|^2_{W_{n+m}} \\ &\leqslant \|z_n - x_{n+m}\|^2_{W_{n+m}} - \|x_{n+m} - z_{n+m}\|^2_{W_{n+m}}. \end{aligned} \tag{5.62}$$

Now fix $z \in C$, and set $\mu = \sup_{n\in\mathbb{N}} \|W_n\|$ and $\rho = \sup_{n\in\mathbb{N}} \|x_n - z\|^2_{W_n}$. Then $\mu < +\infty$ and, in view of Proposition 5.16(i), $\rho < +\infty$. It follows from (5.59) that, for every $n \in \mathbb{N}$ and every $m \in \mathbb{N}\smallsetminus\{0\}$, since $P_C^{W_n}$ is nonexpansive with respect to $\|\cdot\|_{W_n}$ [5, Proposition 4.8], we have

$$\begin{aligned}\|x_{n+m} - z_n\|^2_{W_{n+m}} &\leqslant \|x_n - z_n\|^2_{W_n} + \sum_{k=n}^{n+m-1} \left(\eta_k \|x_k - z_n\|^2_{W_k} + \varepsilon_k\right) \\ &\leqslant \|x_n - z_n\|^2_{W_n} + \sum_{k=n}^{n+m-1} \Big(2\eta_k\left(\|x_k - z\|^2_{W_k} + \|z_n - z\|^2_{W_k}\right) + \varepsilon_k\Big) \\ &\leqslant \|x_n - z_n\|^2_{W_n} + \sum_{k=n}^{n+m-1} \Big(2\eta_k\Big(\rho + \frac{\mu}{\alpha}\|P_C^{W_n} x_n - P_C^{W_n} z\|^2_{W_n}\Big) + \varepsilon_k\Big) \\ &\leqslant \|x_n - z_n\|^2_{W_n} + \sum_{k=n}^{n+m-1} \Big(2\eta_k\Big(\rho + \frac{\mu}{\alpha}\|x_n - z\|^2_{W_n}\Big) + \varepsilon_k\Big) \\ &\leqslant \|x_n - z_n\|^2_{W_n} + \sum_{k=n}^{n+m-1} \Big(2\rho\eta_k\Big(1 + \frac{\mu}{\alpha}\Big) + \varepsilon_k\Big). \end{aligned} \tag{5.63}$$

Combining (5.62) and (5.63), we obtain that for every $n \in \mathbb{N}$ and every $m \in \mathbb{N} \smallsetminus \{0\}$,

$$\begin{aligned}\alpha\|z_{n+m} - z_n\|^2 &\leqslant \|z_{n+m} - z_n\|^2_{W_{n+m}} \\ &\leqslant \|x_n - z_n\|^2_{W_n} - \|x_{n+m} - z_{n+m}\|^2_{W_{n+m}} + \sum_{k\geqslant n} \Big(2\rho\eta_k\Big(1 + \frac{\mu}{\alpha}\Big) + \varepsilon_k\Big). \end{aligned} \tag{5.64}$$

On the other hand, (5.59) yields

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \|x_{n+1} - z_{n+1}\|^2_{W_{n+1}} &\leqslant \|x_{n+1} - z_n\|^2_{W_{n+1}} \\ &\leqslant (1 + \eta_n)\|x_n - z_n\|^2_{W_n} + \varepsilon_n, \end{aligned} \tag{5.65}$$

which, by Lemma 5.13, implies that $(\|x_n - z_n\|_{W_n})_{n\in\mathbb{N}}$ converges. Consequently, since $(\eta_k)_{k\in\mathbb{N}}$ and $(\varepsilon_k)_{k\in\mathbb{N}}$ are in $\ell^1_+(\mathbb{N})$, we derive from (5.64) that $(z_n)_{n\in\mathbb{N}}$ is a Cauchy sequence and hence that it converges strongly. □

In the case of classical Fejér monotone sequences, it has been known since [31] that strong convergence is achieved when the interior of the target set is nonempty (see also [12, Proposition 3.10] for the case of quasi-Fejér monotonicity). The following result extends this fact in the context of variable metric quasi-Fejér sequences.

Proposition 5.21 *Let $\alpha \in \left]0,+\infty\right[$, let $(\nu_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N})$, and let $(W_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{H})$ such that*

$$\mu = \sup_{n\in\mathbb{N}} \|W_n\| < +\infty \quad \textit{and} \quad (\forall n \in \mathbb{N}) \quad (1+\nu_n)W_{n+1} \succcurlyeq W_n. \tag{5.66}$$

Furthermore, let C be a subset of $\mathcal{H}$ such that $\operatorname{int} C \neq \varnothing$, let $z \in C$ and $\rho \in \left]0,+\infty\right[$ be such that $B(z;\rho) \subset C$, and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that

$$\begin{aligned}\big(\exists\, (\varepsilon_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N})\big)\big(\exists\, (\eta_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N})\big)(\forall x \in B(z;\rho))(\forall n \in \mathbb{N})&\\ \|x_{n+1} - x\|^2_{W_{n+1}} \leqslant (1+\eta_n)\|x_n - x\|^2_{W_n} + \varepsilon_n.&\end{aligned} \tag{5.67}$$

Then $(x_n)_{n\in\mathbb{N}}$ converges strongly.

Proof. We derive from (5.66) and Proposition 5.16(ii) that

$$\zeta = \sup_{x\in B(z;\rho)} \sup_{n\in\mathbb{N}} \|x_n - x\|^2_{W_n} \leqslant 2\mu\bigg(\sup_{n\in\mathbb{N}} \|x_n - z\|^2 + \sup_{x\in B(z;\rho)} \|x-z\|^2\bigg) < +\infty. \tag{5.68}$$

It follows from (5.67) and (5.68) that

$$(\forall n \in \mathbb{N})(\forall x \in B(z;\rho)) \quad \|x_{n+1}-x\|^2_{W_{n+1}} \leqslant \|x_n - x\|^2_{W_n} + \xi_n, \quad \text{where} \quad \xi_n = \zeta\eta_n + \varepsilon_n. \tag{5.69}$$

Now set

$$(\forall n \in \mathbb{N}) \quad v_n = W_{n+1}(x_{n+1} - z) - W_n(x_n - z), \tag{5.70}$$

and define a sequence $(z_n)_{n\in\mathbb{N}}$ in $B(z;\rho)$ by

$$(\forall n \in \mathbb{N}) \quad z_n = z - \rho u_n, \quad \text{where} \quad u_n = \begin{cases} 0, & \text{if } v_n = 0;\\ v_n/\|v_n\|, & \text{if } v_n \neq 0.\end{cases} \tag{5.71}$$

Then

$$(\forall n \in \mathbb{N}) \quad \begin{cases} \|x_{n+1} - z_n\|^2_{W_{n+1}} & = \|x_{n+1} - z\|^2_{W_{n+1}} + 2\rho\langle W_{n+1}(x_{n+1} - z) \mid u_n\rangle\\ & \quad + \rho^2\|u_n\|^2_{W_{n+1}};\\ \|x_n - z_n\|^2_{W_n} & = \|x_n - z\|^2_{W_n} + 2\rho\langle W_n(x_n - z) \mid u_n\rangle + \rho^2\|u_n\|^2_{W_n}.\end{cases}$$

$$(5.72)$$

On the other hand, (5.69) yields $(\forall n \in \mathbb{N})$ $\|x_{n+1}-z_n\|^2_{W_{n+1}} \leqslant \|x_n-z_n\|^2_{W_n}+\xi_n$. Therefore, it follows from (5.72), (5.70), and (5.66) that

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \|x_{n+1}-z\|^2_{W_{n+1}} &\leqslant \|x_n-z\|^2_{W_n} - 2\rho\|v_n\| + \rho^2\big(\|u_n\|^2_{W_n} - \|u_n\|^2_{W_{n+1}}\big) + \xi_n \\ &\leqslant \|x_n-z\|^2_{W_n} - 2\rho\|v_n\| + \rho^2\mu\nu_n + \xi_n. \end{aligned} \tag{5.73}$$

Since $(\rho^2\mu\nu_n+\xi_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N})$, this implies that

$$\sum_{n\in\mathbb{N}} \|w_{n+1}-w_n\| = \sum_{n\in\mathbb{N}} \|v_n\| < +\infty, \quad \text{where} \quad (\forall n \in \mathbb{N}) \quad w_n = W_n(x_n - z). \tag{5.74}$$

Hence, $(w_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{H}$ and, therefore, there exists $w \in \mathcal{H}$ such that $w_n \to w$. On the other hand, we deduce from (5.66) and Lemma 5.14(ii) that there exists $W \in \mathcal{P}_\alpha(\mathcal{H})$ such that $W_n \to W$. Now set $x = z + W^{-1}w$. Then, since $(W_n)_{n\in\mathbb{N}}$ lies in $\mathcal{P}_\alpha(\mathcal{H})$, it follows from Cauchy-Schwarz that

$$\alpha\|x_n-x\| \leqslant \|W_nx_n-W_nx\| = \|w_n-W_nW^{-1}w\| \leqslant \|w_n-w\|+\|w-W_nW^{-1}w\| \to 0, \tag{5.75}$$

which concludes the proof. □

5.6 Application to convex feasibility

We illustrate our results through an application to the convex feasibility problem, i.e., the generic problem of finding a common point of a family of closed convex sets. As in [4], given $\alpha \in \left]0,+\infty\right[$ and $W \in \mathcal{P}_\alpha(\mathcal{H})$, we say that an operator $T\colon \mathcal{H} \to \mathcal{H}$ with fixed point set $\operatorname{Fix} T$ belongs to $\mathfrak{T}(W)$ if

$$(\forall x \in \mathcal{H})(\forall y \in \operatorname{Fix} T) \quad \langle y - Tx \mid x - Tx\rangle_W \leqslant 0. \tag{5.76}$$

If $T \in \mathfrak{T}(W)$, then [12, Proposition 2.3(ii)] yields

$$\begin{aligned}(\forall x \in \mathcal{H})(\forall y \in \operatorname{Fix} T)(\forall \lambda \in [0,2]) \quad &\|(\operatorname{Id} + \lambda(T - \operatorname{Id}))x - y\|^2_W \\ &\leqslant \|x-y\|^2_W - \lambda(2-\lambda)\|Tx-x\|^2_W. \end{aligned} \tag{5.77}$$

The usefulness of the class $\mathfrak{T}(W)$ stems from the fact that it contains many of the operators commonly encountered in nonlinear analysis : firmly nonexpansive operators (in particular resolvents of maximally monotone operators and proximity operators of proper lower semicontinuous convex functions), subgradient projection operators, projection operators, averaged quasi-nonexpansive operators, and several combinations thereof [4, 6, 12].

Theorem 5.22 *Let $\alpha \in \,]0,+\infty[$, let $(C_i)_{i\in I}$ be a finite or countably infinite family of closed convex subsets of $\mathcal{H}$ such that $C = \bigcap_{i\in I} C_i \neq \varnothing$, let $(a_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that $\sum_{n\in\mathbb{N}} \|a_n\| < +\infty$, let $(\eta_n)_{n\in\mathbb{N}}$ be a sequence in $\ell_+^1(\mathbb{N})$, and let $(W_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{H})$ such that*

$$\mu = \sup_{n\in\mathbb{N}} \|W_n\| < +\infty \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad (1+\eta_n)W_n \succcurlyeq W_{n+1}. \tag{5.78}$$

Let $\mathrm{i}\colon \mathbb{N} \to I$ be such that

$$(\forall j \in I)(\exists\, M_j \in \mathbb{N} \smallsetminus \{0\})(\forall n \in \mathbb{N}) \quad j \in \{\mathrm{i}(n), \ldots, \mathrm{i}(n+M_j-1)\}. \tag{5.79}$$

For every $i \in I$, let $(T_{i,n})_{n\in\mathbb{N}}$ be a sequence of operators such that

$$(\forall n \in \mathbb{N}) \quad T_{i,n} \in \mathfrak{T}(W_n) \quad \text{and} \quad \operatorname{Fix} T_{i,n} = C_i. \tag{5.80}$$

Fix $\varepsilon \in \,]0,1[$ and $x_0 \in \mathcal{H}$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, 2-\varepsilon]$, and set

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n\big(T_{\mathrm{i}(n),n}x_n + a_n - x_n\big). \tag{5.81}$$

Suppose that, for every strictly increasing sequence $(p_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$, every $x \in \mathcal{H}$, and every $j \in I$,

$$\begin{cases} x_{p_n} \rightharpoonup x \\ T_{j,p_n}x_{p_n} - x_{p_n} \to 0 \\ (\forall n \in \mathbb{N}) \;\; j = \mathrm{i}(p_n) \end{cases} \quad \Rightarrow \quad x \in C_j. \tag{5.82}$$

Then the following hold for some $\overline{x} \in C$.

(i) $x_n \rightharpoonup \overline{x}$.

(ii) *Suppose that $\operatorname{int} C \neq \varnothing$ and that there exists $(\nu_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$ such that $(\forall n \in \mathbb{N})$ $(1+\nu_n)W_{n+1} \succcurlyeq W_n$. Then $x_n \to \overline{x}$.*

(iii) *Suppose that $\underline{\lim}\, d_C(x_n) = 0$. Then $x_n \to \overline{x}$.*

(iv) *Suppose that there exists an index $j \in I$ of demicompact regularity : for every strictly increasing sequence $(p_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$,*

$$\begin{cases} \sup_{n\in\mathbb{N}} \|x_{p_n}\| < +\infty \\ T_{j,p_n}x_{p_n} - x_{p_n} \to 0 \\ (\forall n \in \mathbb{N}) \;\; j = \mathrm{i}(p_n) \end{cases} \quad \Rightarrow \quad (x_{p_n})_{n\in\mathbb{N}} \text{ has a strong sequential cluster point.} \tag{5.83}$$

Then $x_n \to \overline{x}$.

Proof. Fix $z \in C$ and set

$$(\forall n \in \mathbb{N}) \quad y_n = x_n + \lambda_n\big(T_{\mathrm{i}(n),n}x_n - x_n\big). \tag{5.84}$$

Appealing to (5.77) and the fact that, by virtue of (5.79), $z \in \bigcap_{i\in I} C_i = \bigcap_{n\in\mathbb{N}} \operatorname{Fix} T_{\mathrm{i}(n),n}$, we obtain,

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \|y_n - z\|_{W_n}^2 &\leqslant \|x_n - z\|_{W_n}^2 - \lambda_n(2-\lambda_n)\|T_{\mathrm{i}(n),n}x_n - x_n\|_{W_n}^2\\ &\leqslant \|x_n - z\|_{W_n}^2 - \varepsilon^2\|T_{\mathrm{i}(n),n}x_n - x_n\|_{W_n}^2. \end{aligned}\tag{5.85}$$

Moreover, it follows from (5.78) that

$$(\forall n \in \mathbb{N}) \quad \|y_n - z\|_{W_{n+1}}^2 \leqslant (1+\eta_n)\|y_n - z\|_{W_n}^2. \tag{5.86}$$

Thus,

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \|y_n - z\|_{W_{n+1}}^2 &\leqslant (1+\eta_n)\|x_n - z\|_{W_n}^2 - \varepsilon^2(1+\eta_n)\|T_{\mathrm{i}(n),n}x_n - x_n\|_{W_n}^2\\ &\leqslant (1+\eta_n)\|x_n - z\|_{W_n}^2 - \varepsilon^2\|T_{\mathrm{i}(n),n}x_n - x_n\|_{W_n}^2 \qquad (5.87)\\ &\leqslant (1+\eta_n)\|x_n - z\|_{W_n}^2. \qquad (5.88)\end{aligned}$$

Using (5.81), (5.84), and (5.88), we get

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \|x_{n+1} - z\|_{W_{n+1}} &\leqslant \|y_n - z\|_{W_{n+1}} + \lambda_n\|a_n\|_{W_{n+1}}\\ &\leqslant \sqrt{1+\eta_n}\|x_n - z\|_{W_n} + \sqrt{\mu}\lambda_n\|a_n\|\\ &\leqslant (1+\eta_n)\|x_n - z\|_{W_n} + 2\sqrt{\mu}\|a_n\|, \end{aligned}\tag{5.89}$$

which shows that

$$(x_n)_{n\in\mathbb{N}} \text{ satisfies (5.35) – and hence (5.34) – with } \phi = |\cdot|. \tag{5.90}$$

It follows from (5.90) and Proposition 5.16(i) that $(\|x_n - z\|_{W_n})_{n\in\mathbb{N}}$ converges, say

$$\|x_n - z\|_{W_n} \to \xi \in \mathbb{R}. \tag{5.91}$$

We therefore derive from (5.89) that $\|y_n - z\|_{W_{n+1}} \to \xi$ and then from (5.87) that

$$\alpha\varepsilon^2\|T_{\mathrm{i}(n),n}x_n - x_n\|^2 \leqslant \varepsilon^2\|T_{\mathrm{i}(n),n}x_n - x_n\|_{W_n}^2 \leqslant (1+\eta_n)\|x_n - z\|_{W_n}^2 - \|y_n - z\|_{W_{n+1}}^2 \to 0. \tag{5.92}$$

(i) : It follows from (5.81) and (5.92) that

$$\begin{aligned}\|x_{n+1} - x_n\| &= \lambda_n\big\|T_{\mathrm{i}(n),n}x_n + a_n - x_n\big\|\\ &\leqslant 2\big(\|T_{\mathrm{i}(n),n}x_n - x_n\| + \|a_n\|\big)\\ &\leqslant 2\big(\|T_{\mathrm{i}(n),n}x_n - x_n\|_{W_n}/\sqrt{\alpha} + \|a_n\|\big)\\ &\to 0. \end{aligned}\tag{5.93}$$

Now, fix $j \in I$ and let x be a weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$. According to (5.79), there exist strictly increasing sequences $(k_n)_{n\in\mathbb{N}}$ and $(p_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that $x_{k_n} \rightharpoonup x$ and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} k_n \leqslant p_n \leqslant k_n + M_j - 1 < k_{n+1} \leqslant p_{n+1}, \\ j = \mathrm{i}(p_n). \end{cases} \tag{5.94}$$

Therefore, we deduce from (5.93) that

$$\begin{aligned} \|x_{p_n} - x_{k_n}\| &\leqslant \sum_{l=k_n}^{k_n+M_j-2} \|x_{l+1} - x_l\| \\ &\leqslant (M_j - 1) \max_{k_n \leqslant l \leqslant k_n+M_j-2} \|x_{l+1} - x_l\| \\ &\to 0, \end{aligned} \tag{5.95}$$

which implies that $x_{p_n} \rightharpoonup x$. We also derive from (5.92) and (5.94) that $T_{j,p_n}x_{p_n} - x_{p_n} = T_{\mathrm{i}(p_n),p_n}x_{p_n} - x_{p_n} \to 0$. Altogether, it follows from (5.82) that $x \in C_j$. Since j was arbitrarily chosen in I, we obtain $x \in C$ and, in view of Lemma 5.14(i) and Theorem 5.17, we conclude that $x_n \rightharpoonup x$.

(ii) : Suppose that $z \in \operatorname{int} C$ and fix $\rho \in \left]0, +\infty\right[$ such that $B(z;\rho) \subset C$. Set $\eta = \sup_{n\in\mathbb{N}} \eta_n$, $\zeta = \sup_{x\in B(z;\rho)} \sup_{n\in\mathbb{N}} \|x_n - x\|_{W_n}$, and

$$(\forall n \in \mathbb{N}) \quad \varepsilon_n = 4\big(\zeta\sqrt{\mu(1+\eta)}\|a_n\| + \mu\|a_n\|^2\big). \tag{5.96}$$

Then $\eta < +\infty$ and, as in (5.68), $\zeta < +\infty$. Therefore $(\varepsilon_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N})$. Furthermore, we derive from (5.81), (5.84), and (5.88) that, for every $x \in B(z;\rho)$ and every $n \in \mathbb{N}$,

$$\begin{aligned} \|x_{n+1} - x\|^2_{W_{n+1}} &\leqslant \|y_n - x\|^2_{W_{n+1}} + 2\lambda_n\|y_n - x\|_{W_{n+1}} \|a_n\|_{W_{n+1}} + \lambda_n^2\|a_n\|^2_{W_{n+1}} \\ &\leqslant (1+\eta_n)\|x_n - x\|^2_{W_n} + 4\sqrt{\mu(1+\eta_n)}\|x_n - x\|_{W_n} \|a_n\| + 4\mu\|a_n\|^2 \\ &\leqslant (1+\eta_n)\|x_n - x\|^2_{W_n} + \varepsilon_n. \end{aligned} \tag{5.97}$$

Altogether, the assertion follows from (i) and Proposition 5.21.

(iii) : This follows from (5.90), Proposition 5.18, and (i).

(iv) : Let $j \in I$ be an index of demicompact regularity and let $(p_n)_{n\in\mathbb{N}}$ be a strictly increasing sequence such that $(\forall n \in \mathbb{N})\ j = \mathrm{i}(p_n)$. Then $(x_{p_n})_{n\in\mathbb{N}}$ is bounded, while (5.92) asserts that $T_{j,p_n}x_{p_n} - x_{p_n} \to 0$. In turn, (5.83) and (i) imply that $x_{p_n} \to \overline{x} \in C$. Therefore $\underline{\lim}\, d_C(x_n) \leqslant \|x_{p_n} - \overline{x}\| \to 0$ and (iii) yields the result. □

Condition (5.79) first appeared in [9, Definition 5]. Property (5.82) was introduced in [2, Definition 3.7] and property (5.83) in [12, Definition 6.5]. Examples of sequences of operators that satisfy (5.82) can be found in [2, 6, 12]. Here is a simple application of Theorem 5.22 to a variable metric periodic projection method.

Corollary 5.23 *Let $\alpha \in]0,+\infty[$, let m be a strictly positive integer, let $I = \{1,\ldots,m\}$, let $(C_i)_{i\in I}$ be family of closed convex subsets of $\mathcal{H}$ such that $C = \bigcap_{i\in I} C_i \neq \varnothing$, let $(a_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that $\sum_{n\in\mathbb{N}} \|a_n\| < +\infty$, let $(\eta_n)_{n\in\mathbb{N}}$ be a sequence in $\ell_+^1(\mathbb{N})$, and let $(W_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{H})$ such that $\sup_{n\in\mathbb{N}} \|W_n\| < +\infty$ and $(\forall n \in \mathbb{N})\ (1+\eta_n)W_n \succcurlyeq W_{n+1}$. Fix $\varepsilon \in]0,1[$ and $x_0 \in \mathcal{H}$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, 2-\varepsilon]$, and set*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n\Big(P^{W_n}_{C_{1+\mathrm{rem}(n,m)}} x_n + a_n - x_n\Big), \tag{5.98}$$

where $\mathrm{rem}(\cdot, m)$ is the remainder function of the division by m. Then the following hold for some $\overline{x} \in C$.

(i) $x_n \rightharpoonup \overline{x}$.

(ii) *Suppose that $\operatorname{int} C \neq \varnothing$ and that there exists $(\nu_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$ such that $(\forall n \in \mathbb{N})$ $(1+\nu_n)W_{n+1} \succcurlyeq W_n$. Then $x_n \to \overline{x}$.*

(iii) *Suppose that there exists $j \in I$ such that C_j is boundedly compact, i.e., its intersection with every closed ball of $\mathcal{H}$ is compact. Then $x_n \to \overline{x}$.*

Proof. The function $\mathrm{i}\colon \mathbb{N} \to I \colon n \mapsto 1 + \mathrm{rem}(n,m)$ satisfies (5.79) with $(\forall j \in I)\ M_j = m$. Now, set $(\forall i \in I)(\forall n \in \mathbb{N})\ T_{i,n} = P^{W_n}_{C_i}$. Then $(\forall i \in I)(\forall n \in \mathbb{N})\ T_{i,n} \in \mathfrak{T}(W_n)$ and $\operatorname{Fix} T_{i,n} = C_i$. Hence, (5.98) is a special case of (5.81).

(i)–(ii) : Fix $j \in I$ and let $(x_{p_n})_{n\in\mathbb{N}}$ be a weakly convergent subsequence of $(x_n)_{n\in\mathbb{N}}$, say $x_{p_n} \rightharpoonup x$, such that $T_{j,p_n}x_{p_n} - x_{p_n} \to 0$ and $(\forall n \in \mathbb{N})\ j = \mathrm{i}(p_n)$. Then $C_j \ni P^{W_{p_n}}_{C_j} x_{p_n} = T_{j,p_n}x_{p_n} \rightharpoonup x$ and, since C_j is weakly closed [5, Theorem 3.32], we have $x \in C_j$. This shows that (5.82) holds. Altogether, the claims follow from Theorem 5.22(i)–(ii).

(iii) : Let $(p_n)_{n\in\mathbb{N}}$ be a strictly increasing sequence in $\mathbb{N}$ such that $P^{W_{p_n}}_{C_j} x_{p_n} - x_{p_n} = T_{j,p_n}x_{p_n} - x_{p_n} \to 0$ and $(\forall n \in \mathbb{N})\ j = \mathrm{i}(p_n)$. Then

$$\|P_{C_j} x_{p_n} - x_{p_n}\| \leqslant \|P^{W_{p_n}}_{C_j} x_{p_n} - x_{p_n}\| \to 0. \tag{5.99}$$

On the other hand, since $(x_{p_n})_{n\in\mathbb{N}}$ is bounded and P_{C_j} is nonexpansive, $(P_{C_j}x_{p_n})_{n\in\mathbb{N}}$ is a bounded sequence in the boundedly compact set C_j. Hence, $(P_{C_j}x_{p_n})_{n\in\mathbb{N}}$ admits a strong sequential cluster point and so does $(x_{p_n})_{n\in\mathbb{N}}$ since $P_{C_j}x_{p_n} - x_{p_n} \to 0$. Thus, $j \in I$ is an index of demicompact regularity and the claim therefore follows from Theorem 5.22(iv). ☐

Remark 5.24 In the special case when, for every $n \in \mathbb{N}$, $W_n = \mathrm{Id}$ and $\eta_n = 0$, Corollary 5.23(i) was established in [8] (with $(\forall n \in \mathbb{N})\ \lambda_n = 1$), and Corollary 5.23(ii) in [22].

Next is an application of Corollary 5.23 to the problem of solving linear inequalities. In Euclidean spaces, the use of periodic projection methods to solve this problem goes back to [27].

Example 5.25 Let $\alpha \in \left]0,+\infty\right[$, let m be a strictly positive integer, let $I = \{1,\ldots,m\}$, let $(\eta_i)_{i\in I}$ be real numbers, and suppose that $(u_i)_{i\in I}$ are nonzero vectors in $\mathcal{H}$ such that

$$C = \big\{x \in \mathcal{H} \;\big|\; (\forall i \in I)\;\; \langle x \mid u_i\rangle \leqslant \eta_i\big\} \neq \varnothing. \tag{5.100}$$

Let $(\eta_n)_{n\in\mathbb{N}}$ be a sequence in $\ell_+^1(\mathbb{N})$, and let $(W_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{H})$ such that $\sup_{n\in\mathbb{N}}\|W_n\| < +\infty$ and $(\forall n \in \mathbb{N})\ (1+\eta_n)W_n \succcurlyeq W_{n+1}$. Fix $\varepsilon \in \left]0,1\right[$ and $x_0 \in \mathcal{H}$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, 2-\varepsilon]$, and set

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} \mathrm{i}(n) = 1 + \mathrm{rem}(n,m) \\ \text{if } \big\langle x_n \mid u_{\mathrm{i}(n)}\big\rangle \leqslant \eta_{\mathrm{i}(n)} \\ \lfloor\; y_n = x_n \\ \text{if } \big\langle x_n \mid u_{\mathrm{i}(n)}\big\rangle > \eta_{\mathrm{i}(n)} \\ \left\lfloor\; y_n = x_n + \dfrac{\eta_{\mathrm{i}(n)} - \big\langle x_n \mid u_{\mathrm{i}(n)}\big\rangle}{\big\langle u_{\mathrm{i}(n)} \mid W_n^{-1}u_{\mathrm{i}(n)}\big\rangle} W_n^{-1}u_{\mathrm{i}(n)} \right. \\ x_{n+1} = x_n + \lambda_n(y_n - x_n). \end{array}\right. \tag{5.101}$$

Then there exists $\overline{x} \in C$ such that $x_n \rightharpoonup \overline{x}$.

Proof. Set $(\forall i \in I)\ C_i = \big\{x \in \mathcal{H} \;\big|\; \langle x \mid u_i\rangle \leqslant \eta_i\big\}$. Then it follows from [5, Example 28.16(iii)] that (5.101) can be rewritten as

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n\Big(P^{W_n}_{C_{1+\mathrm{rem}(n,m)}} x_n - x_n\Big). \tag{5.102}$$

The claim is therefore a consequence of Corollary 5.23(i). □

We now turn our attention to the problem of finding a zero of a maximally monotone operator $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ (see [5] for background) via a variable metric proximal point algorithm. Let $\alpha \in \left]0,+\infty\right[$, let $\gamma \in \left]0,+\infty\right[$, let $W \in \mathcal{P}_\alpha(\mathcal{H})$, and let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone with graph $\mathrm{gra}A$. It follows from [3, Corollary 3.14(ii)] (applied with $f\colon x \mapsto \langle Wx \mid x\rangle/2$) that

$$J^W_{\gamma A}\colon \mathcal{H} \to \mathcal{H}\colon x \mapsto (W + \gamma A)^{-1}(Wx) \tag{5.103}$$

is well-defined, and that

$$J^W_{\gamma A} \in \mathfrak{T}(W) \quad \text{and} \quad \operatorname{Fix} J^W_{\gamma A} = \big\{z \in \mathcal{H} \;\big|\; 0 \in Az\big\}. \tag{5.104}$$

We write $J^{\mathrm{Id}}_{\gamma A} = J_{\gamma A}$.

Corollary 5.26 *Let $\alpha \in \left]0,+\infty\right[$, let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator such that $C = \big\{z \in \mathcal{H} \;\big|\; 0 \in Az\big\} \neq \varnothing$, let $(a_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that $\sum_{n\in\mathbb{N}}\|a_n\| < +\infty$, let $(\eta_n)_{n\in\mathbb{N}}$ be a sequence in $\ell_+^1(\mathbb{N})$, and let $(W_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{H})$ such that $\mu = \sup_{n\in\mathbb{N}}\|W_n\| < +\infty$ and $(\forall n \in \mathbb{N})\ (1+\eta_n)W_n \succcurlyeq W_{n+1}$. Fix $\varepsilon \in \left]0,1\right[$ and $x_0 \in \mathcal{H}$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, 2-\varepsilon]$, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $\left[\varepsilon,+\infty\right[$, and set*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n\Big(J^{W_n}_{\gamma_n A} x_n + a_n - x_n\Big). \tag{5.105}$$

Then the following hold for some $\overline{x} \in C$.

(i) $x_n \rightharpoonup \overline{x}$.

(ii) *Suppose that* $\operatorname{int} C \neq \varnothing$ *and that there exists* $(\nu_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$ *such that* $(\forall n \in \mathbb{N})$ $(1+\nu_n)W_{n+1} \succcurlyeq W_n$. *Then* $x_n \to \overline{x}$.

(iii) *Suppose that* A *is pointwise uniformly monotone on* C, *i.e., for every* $x \in C$ *there exists an increasing function* $\phi\colon [0,+\infty[\to [0,+\infty]$ *vanishing only at* 0 *such that*

$$(\forall u \in Ax)(\forall (y,v) \in \operatorname{gra} A)\ \ \langle x-y \mid u-v\rangle \geqslant \phi(\|x-y\|). \tag{5.106}$$

Then $x_n \to \overline{x}$.

Proof. In view of (5.104), (5.105) is a special case of (5.81) with $I = \{1\}$ and $(\forall n \in \mathbb{N})$ $T_{1,n} = J_{\gamma_n A}^{W_n}$. Hence, using Theorem 5.22(i)–(ii), to show (i)–(ii), it suffices to prove that (5.82) holds. To this end, let $(x_{p_n})_{n\in\mathbb{N}}$ be a weakly convergent subsequence of $(x_n)_{n\in\mathbb{N}}$, say $x_{p_n} \rightharpoonup x$, such that $J_{\gamma_{p_n} A}^{W_{p_n}} x_{p_n} - x_{p_n} \to 0$. To show that $0 \in Ax$, let us set

$$(\forall n \in \mathbb{N}) \quad y_n = J_{\gamma_n A}^{W_n} x_n \quad \text{and} \quad v_n = \frac{1}{\gamma_n} W_n(x_n - y_n). \tag{5.107}$$

Then (5.103) yields $(\forall n \in \mathbb{N})$ $v_n \in Ay_n$. On the other hand, since $y_{p_n} - x_{p_n} \to 0$, we have

$$\|v_{p_n}\| = \frac{\|W_{p_n}(x_{p_n} - y_{p_n})\|}{\gamma_{p_n}} \leqslant \frac{\mu}{\varepsilon}\|x_{p_n} - y_{p_n}\| \to 0. \tag{5.108}$$

Thus, $y_{p_n} \rightharpoonup x$ and $Ay_{p_n} \ni v_{p_n} \to 0$. Since $\operatorname{gra} A$ is sequentially closed in $\mathcal{H}^{\text{weak}} \times \mathcal{H}^{\text{strong}}$ [5, Proposition 20.33(ii)], we conclude that $0 \in Ax$. Let us now show (iii). We have $0 \in A\overline{x}$ and $(\forall n \in \mathbb{N})$ $v_{p_n} \in Ay_{p_n}$. Hence, it follows from (5.106) that there exists an increasing function $\phi\colon [0,+\infty[\to [0,+\infty]$ vanishing only at 0 such that

$$(\forall n \in \mathbb{N}) \quad \langle y_{p_n} - \overline{x} \mid v_{p_n}\rangle \geqslant \phi(\|y_{p_n} - \overline{x}\|). \tag{5.109}$$

Since $v_{p_n} \to 0$, we get $\phi(\|y_{p_n} - \overline{x}\|) \to 0$ and, in turn, $\|y_{p_n} - \overline{x}\| \to 0$. It follows that $\|x_{p_n} - \overline{x}\| \to 0$ and hence that $\underline{\lim}\, d_C(x_n) = 0$. In view of Theorem 5.22(iii), we conclude that $x_n \to \overline{x}$. ☐

Remark 5.27 Corollary 5.26(i) reduces to the classical result of [34, Theorem 1] when $(\forall n \in \mathbb{N})$ $W_n = \operatorname{Id}$, $\eta_n = 0$, and $\lambda_n = 1$. In this context, Corollary 5.26(ii) appears in [28, Section 6]. In a finite-dimensional setting, an alternative variable metric proximal point algorithm is proposed in [29], which also uses the above conditions on $(W_n)_{n\in\mathbb{N}}$ but alternative error terms and relaxation parameters.

5.7 Application to inverse problems

In this section, we consider an application to a structured variational inverse problem. Henceforth, $\Gamma_0(\mathcal{H})$ denotes the class of proper lower semicontinuous convex functions from $\mathcal{H}$ to $]-\infty,+\infty]$.

Problem 5.28 Let $f \in \Gamma_0(\mathcal{H})$ and let I be a nonempty finite index set. For every $i \in I$, let $(\mathcal{G}_i, \|\cdot\|_i)$ be a real Hilbert space, let $L_i \colon \mathcal{H} \to \mathcal{G}_i$ be a nonzero bounded linear operator, let $r_i \in \mathcal{G}_i$, and let $\mu_i \in \left]0, +\infty\right[$. The problem is to

$$\underset{x \in \mathcal{H}}{\text{minimize}} \;\; f(x) + \frac{1}{2}\sum_{i \in I} \mu_i \|L_i x - r_i\|_i^2. \tag{5.110}$$

This formulation covers many inverse problems (see [17, Section 5] and the references therein) and it can be interpreted as follows : an ideal object $\widetilde{x} \in \mathcal{H}$ is to be recovered from noisy linear measurements $r_i = L_i\widetilde{x} + w_i \in \mathcal{G}_i$, where w_i represents noise ($i \in I$), and the function f penalizes the violation of prior information on $\widetilde{x}$. Thus, (5.110) attempts to strike a balance between the observation model, represented by the data fitting term $x \mapsto (1/2)\sum_{i \in I} \mu_i \|L_i x - r_i\|_i^2$, and a priori knowledge, represented by f. To solve this problem within our framework, we require the following facts.

Let $\alpha \in \left]0, +\infty\right[$, let $W \in \mathcal{P}_\alpha(\mathcal{H})$, and let $\varphi \in \Gamma_0(\mathcal{H})$. The proximity operator of φ relative to the metric induced by W is

$$\operatorname{prox}_\varphi^W \colon \mathcal{H} \to \mathcal{H} \colon x \mapsto \underset{y \in \mathcal{H}}{\operatorname{argmin}} \left(\varphi(y) + \frac{1}{2}\|x - y\|_W^2 \right). \tag{5.111}$$

Now, let $\partial\varphi$ be the subdifferential of φ [5, Chapter 16]. Then, in connection with (5.103), $\partial\varphi$ is maximally monotone and we have [16, Section 3.3]

$$(\forall \gamma \in \left]0, +\infty\right[) \quad \operatorname{prox}_{\gamma\varphi}^W = J_{\gamma\partial\varphi}^W = (W + \gamma\partial\varphi)^{-1} \circ W. \tag{5.112}$$

We write $\operatorname{prox}_{\gamma\varphi}^{\mathrm{Id}} = \operatorname{prox}_{\gamma\varphi}$.

Lemma 5.29 *Let $A \colon \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone, let U be a nonzero operator in $\mathcal{P}_0(\mathcal{H})$, let $\gamma \in \left]0, 1/\|U\|\right[$, let $u \in \mathcal{H}$, set $W = \mathrm{Id} - \gamma U$, and set $B = A + U + \{u\}$. Then*

$$(\forall x \in \mathcal{H}) \quad J_{\gamma B}^W x = J_{\gamma A}\big(Wx - \gamma u\big). \tag{5.113}$$

Proof. Since $U \in \mathcal{P}_0(\mathcal{H})$, U is maximally monotone [5, Example 20.29]. In turn, it follows from [5, Corollary 24.4(i)] that B is maximally monotone. Moreover, $W \in \mathcal{P}_\alpha(\mathcal{H})$, where $\alpha = 1 - \gamma\|U\|$. Now, let x and p be in $\mathcal{H}$. Then it follows from (5.103) that

$$p = J_{\gamma B}^W x \Leftrightarrow Wx \in Wp + \gamma Bp \Leftrightarrow Wx - \gamma u \in p + \gamma Ap \Leftrightarrow p = J_{\gamma A}\big(Wx - \gamma u\big), \tag{5.114}$$

which completes the proof. □

Proposition 5.30 *Let $\varepsilon \in \left]0, 1/(1 + \sum_{i \in I} \mu_i \|L_i\|^2)\right[$, let $(a_n)_{n \in \mathbb{N}}$ be a sequence in $\mathcal{H}$ such that $\sum_{n \in \mathbb{N}} \|a_n\| < +\infty$, let $(\eta_n)_{n \in \mathbb{N}}$ be a sequence in $\ell_+^1(\mathbb{N})$, and let $(\gamma_n)_{n \in \mathbb{N}}$ be a sequence in $\mathbb{R}$ such that*

$$(\forall n \in \mathbb{N}) \quad \varepsilon \leqslant \gamma_n \leqslant \frac{1 - \varepsilon}{\sum_{i \in I} \mu_i \|L_i\|^2} \quad \text{and} \quad (1 + \eta_n)\gamma_n - \gamma_{n+1} \leqslant \frac{\eta_n}{\sum_{i \in I} \mu_i \|L_i\|^2}. \tag{5.115}$$

Furthermore, let C be the set of solutions to Problem 5.28, let $x_0 \in \mathcal{H}$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, 2-\varepsilon]$, and set

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n\Big(\operatorname{prox}_{\gamma_n f}\Big(x_n + \gamma_n \sum_{i\in I} \mu_i L_i^*\big(r_i - L_i x_n\big)\Big) + a_n - x_n\Big). \tag{5.116}$$

Then the following hold for some $\overline{x} \in C$.

(i) *Suppose that*

$$\lim_{\|x\|\to+\infty} \quad f(x) + \frac{1}{2}\sum_{i\in I} \mu_i \|L_i x - r_i\|_i^2 = +\infty. \tag{5.117}$$

Then $x_n \rightharpoonup \overline{x}$.

(ii) *Suppose that there exists $j \in I$ such that L_j is bounded below, say,*

$$(\exists\, \beta \in\,]0,+\infty[)(\forall x \in \mathcal{H}) \quad \|L_j x\|_j \geqslant \beta\|x\|. \tag{5.118}$$

Then $C = \{\overline{x}\}$ and $x_n \to \overline{x}$.

Proof. Set $U = \sum_{i\in I} \mu_i L_i^* L_i$ and $u = -\sum_{i\in I} \mu_i L_i^* r_i$. Then

$$\|U\| \leqslant \sum_{i\in I} \mu_i \|L_i\|^2, \tag{5.119}$$

and the assumptions imply that $0 \neq U \in \mathcal{P}_0(\mathcal{H})$ and that $(\forall n \in \mathbb{N})\ \varepsilon \leqslant \gamma_n \leqslant (1-\varepsilon)/\|U\|$. Now set

$$g\colon \mathcal{H} \to\,]-\infty,+\infty] : x \mapsto f(x) + \frac{1}{2}\langle Ux \mid x\rangle + \langle x \mid u\rangle \tag{5.120}$$

and

$$(\forall n \in \mathbb{N}) \quad W_n = \operatorname{Id} - \gamma_n U. \tag{5.121}$$

Then (5.110) is equivalent to minimizing g. Furthermore, it follows from (5.115) that $(W_n)_{n\in\mathbb{N}}$ lies in $\mathcal{P}_\varepsilon(\mathcal{H})$ and that $\sup_{n\in\mathbb{N}} \|W_n\| \leqslant 2-\varepsilon$. In addition, we have

$$(\forall n \in \mathbb{N}) \quad \eta_n \geqslant \big((1+\eta_n)\gamma_n - \gamma_{n+1}\big)\|U\|. \tag{5.122}$$

Indeed if, for some $n \in \mathbb{N}$, $(1+\eta_n)\gamma_n \leqslant \gamma_{n+1}$ then $\eta_n \geqslant 0 \geqslant ((1+\eta_n)\gamma_n - \gamma_{n+1})\|U\|$; otherwise we deduce from (5.115) and (5.119) that $\eta_n \geqslant ((1+\eta_n)\gamma_n - \gamma_{n+1}) \sum_{i\in I} \mu_i \|L_i\|^2 \geqslant ((1+\eta_n)\gamma_n - \gamma_{n+1})\|U\|$. Thus, since $U \in \mathcal{P}_0(\mathcal{H})$, we have $\|U\| = \sup_{\|x\|\leqslant 1} \langle Ux \mid x\rangle$ and therefore

$$\begin{aligned}
(5.122) \quad &\Rightarrow \quad (\forall n \in \mathbb{N})(\forall x \in \mathcal{H}) \quad \eta_n\|x\|^2 \geqslant \big((1+\eta_n)\gamma_n - \gamma_{n+1}\big)\langle Ux \mid x\rangle \\
&\Rightarrow \quad (\forall n \in \mathbb{N})(\forall x \in \mathcal{H}) \quad (1+\eta_n)(\|x\|^2 - \gamma_n\langle Ux \mid x\rangle) \geqslant \|x\|^2 - \gamma_{n+1}\langle Ux \mid x\rangle \\
&\Rightarrow \quad (\forall n \in \mathbb{N}) \quad (1+\eta_n)W_n \succcurlyeq W_{n+1}.
\end{aligned} \tag{5.123}$$

Now set $A=\partial f$ and $B=A+U+\{u\}$. Then we derive from [5, Corollary 16.38(iii)] that $B=\partial g$. Hence, using (5.112), (5.121), and Lemma 5.29, (5.116) can be rewritten as

$$(\forall n\in\mathbb{N})\quad x_{n+1}=x_n+\lambda_n\Big(\mathrm{prox}_{\gamma_n f}\big(x_n-\gamma_n(Ux_n+u)\big)+a_n-x_n\Big)$$
$$=x_n+\lambda_n\Big(J_{\gamma_n A}\big(W_nx_n-\gamma_n u\big)+a_n-x_n\Big)$$
$$=x_n+\lambda_n\Big(J^{W_n}_{\gamma_n B}x_n+a_n-x_n\Big). \tag{5.124}$$

On the other hand, it follows from Fermat's rule [5, Theorem 16.2] that

$$\big\{z\in\mathcal{H}\;\big|\;0\in Bz\big\}=\operatorname{Argmin} g=C. \tag{5.125}$$

(i) : Since $f\in\Gamma_0(\mathcal{H})$ and $U\in\mathcal{P}_0(\mathcal{H})$, it follows from [5, Proposition 11.14(i)] that Problem 5.28 admits at least one solution. Altogether, the result follows from Corollary 5.26(i).

(ii) : It follows from (5.118) that $L_j^*L_j\in\mathcal{P}_{\beta^2}(\mathcal{H})$. Therefore, $U\in\mathcal{P}_{\mu_j\beta^2}(\mathcal{H})$ and, since $f\in\Gamma_0(\mathcal{H})$, we derive from (5.120) that $g\in\Gamma_0(\mathcal{H})$ is strongly convex. Hence, [5, Corollary 11.16] asserts that (5.110) possesses a unique solution, while [5, Example 22.3(iv)] asserts that B is strongly – hence uniformly – monotone. Altogether, the claim follows from Corollary 5.26(iii). □

Remark 5.31 In Problem 5.28 suppose that $I=\{1\}$, $\mu_1=1$, $L_1=L$, and $r_1=r$, and that $\lim_{\|x\|\to+\infty} f(x)+\|Lx-r\|_1^2/2=+\infty$. Then (5.116) reduces to the proximal Landweber method

$$(\forall n\in\mathbb{N})\quad x_{n+1}=x_n+\lambda_n\Big(\mathrm{prox}_{\gamma_n f}\big(x_n+\gamma_nL^*(r-Lx_n)\big)+a_n-x_n\Big), \tag{5.126}$$

and we derive from Proposition 5.30(i) that $(x_n)_{n\in\mathbb{N}}$ converges weakly to a minimizer of $x\mapsto f(x)+\|Lx-r\|_1^2/2$ if

$$(\forall n\in\mathbb{N})\quad \begin{cases}\varepsilon\leqslant\gamma_n\leqslant(1-\varepsilon)/\|L\|^2\\ (1+\eta_n)\gamma_n\leqslant\gamma_{n+1}+\eta_n/\|L\|^2\\ \varepsilon\leqslant\lambda_n\leqslant 2-\varepsilon.\end{cases} \tag{5.127}$$

This result complements [17, Theorem 5.5(i)], which establishes weak convergence under alternative conditions on the parameters $(\gamma_n)_{n\in\mathbb{N}}$ and $(\lambda_n)_{n\in\mathbb{N}}$, namely

$$(\forall n\in\mathbb{N})\quad \begin{cases}\varepsilon\leqslant\gamma_n\leqslant(2-\varepsilon)/\|L\|^2\\ \varepsilon\leqslant\lambda_n\leqslant 1.\end{cases} \tag{5.128}$$

In particular, suppose that $\mathcal{H}$ is separable, let $(e_k)_{k\in\mathbb{N}}$ be an orthonormal basis of $\mathcal{H}$, and set $f\colon x\mapsto\sum_{k\in\mathbb{N}}\phi_k(\langle x\mid e_k\rangle)$, where $(\forall k\in\mathbb{N})\ \Gamma_0(\mathbb{R})\ni\phi_k\geqslant\phi_k(0)=0$. Moreover, for every $n \in \mathbb{N}$, let $(\alpha_{n,k})_{k\in\mathbb{N}}$ be a sequence in $\ell^2(\mathbb{N})$ and suppose that $\sum_{n\in\mathbb{N}} \sqrt{\sum_{k\in\mathbb{N}} |\alpha_{n,k}|^2} < +\infty$. Now set $(\forall n \in \mathbb{N})\ a_n = \sum_{k\in\mathbb{N}} \alpha_{n,k} e_k$. Then, arguing as in [17, Section 5.4], (5.126) becomes

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n \bigg(\sum_{k\in\mathbb{N}} \big(\alpha_{n,k} + \mathrm{prox}_{\gamma_n \phi_k} \langle x_n + \gamma_n L^*(r - Lx_n) \mid e_k\rangle\big) e_k - x_n \bigg), \tag{5.129}$$

and we obtain convergence under the new condition (5.127) (see also [15] for potential signal and image processing applications of this result).

5.8 Bibliographie

[1] Ya. I. Alber, A. N. Iusem, and M. V. Solodov, On the projected subgradient method for nonsmooth convex optimization in a Hilbert space, *Math. Programming A,* vol. 81, pp, 23–35, 1998.

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

[3] H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Bregman monotone optimization algorithms, *SIAM J. Control Optim.,* vol. 42, pp. 596–636, 2003.

[4] H. H. Bauschke and P. L. Combettes, A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces, *Math. Oper. Res.*, vol. 26, pp. 248–264, 2001.

[5] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York, 2011.

[6] H. H. Bauschke, P. L. Combettes, and S. G. Kruk, Extrapolation algorithm for affine-convex feasibility problems, *Numer. Algorithms,* vol. 41, pp. 239–274, 2006.

[7] J. F. Bonnans, J. Ch. Gilbert, C. Lemaréchal, and C. A. Sagastizábal, A family of variable metric proximal methods, *Math. Programming*, vol. 68, pp. 15–47, 1995.

[8] L. M. Bregman, The method of successive projection for finding a common point of convex sets, *Soviet Math. Dokl.*, vol. 6, pp. 688–692, 1965.

[9] F. E. Browder, Convergence theorems for sequences of nonlinear operators in Banach spaces, *Math. Z.,* vol. 100, pp. 201–225, 1967.

[10] J. V. Burke and M. Qian, A variable metric proximal point algorithm for monotone operators, *SIAM J. Control Optim.,* vol. 37, pp. 353–375, 1999.

[11] G. H.-G. Chen and R. T. Rockafellar, Convergence rates in forward-backward splitting, *SIAM J. Optim.*, vol. 7, pp. 421–444, 1997.

[12] P. L. Combettes, Quasi-Fejérian analysis of some optimization algorithms, in : *Inherently Parallel Algorithms for Feasibility and Optimization*, (D. Butnariu, Y. Censor, and S. Reich, eds.), pp. 115–152. Elsevier, New York, 2001.

[13] P. L. Combettes, Solving monotone inclusions via compositions of nonexpansive averaged operators, *Optimization,* vol. 53, pp. 475–504, 2004.

[14] P. L. Combettes, Fejér monotonicity in convex optimization, in : *Encyclopedia of Optimization,* 2nd ed. (C. A. Floudas and P. M. Pardalos, eds.), pp. 1016–1024. Springer, New York, 2009.

[15] P. L. Combettes and J.-C. Pesquet, Proximal thresholding algorithm for minimization over orthonormal bases, *SIAM J. Optim.,* vol. 18, pp. 1351–1376, 2007.

[16] P. L. Combettes and B. C. Vũ, Variable metric forward-backward splitting with applications to monotone inclusions in duality, *Optimization,* to appear, 2013.

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

[18] I. I. Eremin, Fejér mappings and convex programming, *Siberian Math. J.,* vol. 10, pp. 762–772, 1969.

[19] I. I. Eremin and L. D. Popov, Fejér processes in theory and practice : Recent results, *Russian Math. (Iz. VUZ),* vol. 53, pp. 36–55, 2009.

[20] Yu. M. Ermol'ev and A. D. Tuniev, Random Fejér and quasi-Fejér sequences, *Theory of Optimal Solutions– Akademiya Nauk Ukrainskoĭ SSR Kiev,* vol. 2, pp. 76–83, 1968 ; translated in : *American Mathematical Society Selected Translations in Mathematical Statistics and Probability,* vol. 13, pp. 143–148, 1973.

[21] L. Fejér, Über die Lage der Nullstellen von Polynomen, die aus Minimumforderungen gewisser Art entspringen, *Math. Ann.*, vol. 85, pp. 41–48, 1922.

[22] L. G. Gubin, B. T. Polyak, and E. V. Raik, The method of projections for finding the common point of convex sets, *USSR Comput. Math. and Math. Phys.,* vol. 7, pp. 1–24, 1967.

[23] H. S. Hundal, An alternating projection that does not converge in norm, *Nonlinear Anal.*, vol. 57, pp. 35–61, 2004.

[24] T. Kato, *Perturbation Theory for Linear Operators,* 2nd ed. Springer-Verlag, New York, 1980.

[25] K. Knopp, *Infinite Sequences and Series.* Dover, New York, 1956.

[26] P. A. Lotito, L. A. Parente, and M. V. Solodov, A class of variable metric decomposition methods for monotone variational inclusions, *J. Convex Anal.*, vol. 16, pp. 857–880, 2009.

[27] T. S. Motzkin and I. J. Schoenberg, The relaxation method for linear inequalities, *Canadian J. Math,* vol. 6, pp. 393–404, 1954.

[28] O. Nevanlinna and S. Reich, Strong convergence of contraction semigroups and of iterative methods for accretive operators in Banach spaces, *Israel J. Math.,* vol. 32, pp. 44–58, 1979.

[29] L. A. Parente, P. A. Lotito, and M. V. Solodov, A class of inexact variable metric proximal point algorithms, *SIAM J. Optim.*, vol. 19, pp. 240–260, 2008.

[30] B. T. Polyak, *Introduction to Optimization*. Optimization Software Inc., New York, 1987.

[31] È. Raĭk, A class of iterative methods with Fejér-monotone sequences, *Eesti NSV Tead. Akad. Toimetised Füüs.-Mat.,* vol. 18, pp. 22–26, 1969.

[32] S. Reich, Weak convergence theorems for nonexpansive mappings in Banach spaces, *J. Math. Anal. Appl.,* vol. 67, pp. 274–276, 1979.

[33] F. Riesz and B. Sz.-Nagy, *Leçons d'Analyse Fonctionnelle,* 5th ed. Gauthier-Villars, Paris, 1968.

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

[35] D. Schott, A general iterative scheme with applications to convex optimization and related fields, *Optimization,* vol. 22, pp. 885–902, 1991.

Chapitre 6

Méthode explicite-implicite à métrique variable

Nous proposons une méthode explicite-implicite à métrique variable pour résoudre des inclusions monotones et montrons sa convergence dans des espaces hilbertiens réels. Nous l'appliquons aux problèmes d'inclusions fortement monotones en dualité et aux inclusions monotones impliquant des opérateurs cocoercifs.

6.1 Description et résultats principaux

Le résultat principal de ce chapitre est le suivant.

Théorème 6.1 *Soit $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ un opérateur maximalement monotone, soit $\alpha \in \left]0,+\infty\right[$, soit $\beta \in \left]0,+\infty\right[$, soit $B\colon \mathcal{H} \to \mathcal{H}$ un opérateur β-cocoercif, soit $(\eta_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N})$, et soit $(U_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{P}_\alpha(\mathcal{H})$ telle que*

$$\mu = \sup_{n\in\mathbb{N}} \|U_n\| < +\infty \quad \textit{et} \quad (\forall n \in \mathbb{N}) \quad (1+\eta_n)U_{n+1} \succcurlyeq U_n. \tag{6.1}$$

Soit $\varepsilon \in \left]0, \min\{1, 2\beta/(\mu+1)\}\right]$, soit $(\lambda_n)_{n\in\mathbb{N}}$ une suite dans $[\varepsilon, 1]$, soit $(\gamma_n)_{n\in\mathbb{N}}$ une suite dans $[\varepsilon, (2\beta-\varepsilon)/\mu]$, soit $x_0 \in \mathcal{H}$, et soient $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ deux suites absolument sommables dans $\mathcal{H}$. Supposons que

$$Z = \operatorname{zer}(A+B) \neq \varnothing, \tag{6.2}$$

et posons

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} y_n = x_n - \gamma_n U_n(Bx_n + b_n) \\ x_{n+1} = x_n + \lambda_n\big(J_{\gamma_n U_n A}(y_n) + a_n - x_n\big). \end{array} \right. \tag{6.3}$$

Alors, on a les résultats suivants pour un point $\overline{x} \in Z$.

(i) $x_n \rightharpoonup \overline{x}$.

(ii) $\sum_{n\in\mathbb{N}} \|Bx_n - B\overline{x}\|^2 < +\infty$.

(iii) *Supposons que l'une des conditions suivantes soit satisfaite.*

  (a) $\underline{\lim}\, d_Z(x_n) = 0$.

  (b) *En tout point dans Z, A ou B est demirégulier (voir Definition 6.22).*

  (c) $\operatorname{int} Z \neq \varnothing$ *et il existe* $(\nu_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$ *telle que* $(\forall n \in \mathbb{N})\ (1+\nu_n)U_n \succcurlyeq U_{n+1}$.

Alors $x_n \to \overline{x}$.

Remarque 6.2

(i) Supposons que $(\forall n \in \mathbb{N})\ U_n = \operatorname{Id}$. Alors l'algorithme (6.3) se réduit à la méthode explicite-implicite (1.4) étudiée dans [1, 12] où on trouve des cas particuliers tels que [27, 29, 40]. Le Théorème 6.1 étend les résultats de convergence de ces articles.

(ii) Comme on a vu dans [18, Remark 5.12], la convergence de la méthode explicite-implicite vers une solution peut être seulement faible et pas forte, d'où la nécessité d'ajouter des conditions dans le Théorème 6.1(iii).

(iii) Dans des espaces euclidiens, la condition (6.1) a été utilisée dans [32] avec l'algorithme proximal à métrique variable et ensuite dans [28] dans un cadre plus général.

Nous présentons ci-dessous des applications aux inclusions monotones. Nous considérons tout d'abord des inclusions fortement monotones qui comportent des sommes parallèles.

Problème 6.3 Soit $z \in \mathcal{H}$, soit $\rho \in \,]0,+\infty[$, soit $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ un opérateur maximalement monotone, et soit m un entier strictement positif. Pour tout $i \in \{1,\ldots,m\}$, soit $r_i \in \mathcal{G}_i$, soit $B_i\colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ un opérateur maximalement monotone, soit $\nu_i \in \,]0,+\infty[$, soit $D_i\colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ un opérateur maximalement monotone et ν_i-fortement monotone, et supposons que $0 \neq L_i \in \mathcal{B}(\mathcal{H},\mathcal{G}_i)$. De plus, supposons que

$$z \in \operatorname{ran}\left(A + \sum_{i=1}^m L_i^*\big((B_i \,\square\, D_i)(L_i \cdot - r_i)\big) + \rho \operatorname{Id}\right). \tag{6.4}$$

Le problème est de résoudre l'inclusion primale

$$\text{trouver } \overline{x} \in \mathcal{H} \text{ tel que } z \in A\overline{x} + \sum_{i=1}^m L_i^*\big((B_i \,\square\, D_i)(L_i\overline{x} - r_i)\big) + \rho\overline{x}, \tag{6.5}$$

et l'inclusion duale

$$\text{trouver } \overline{v_1} \in \mathcal{G}_1, \ldots, \overline{v_m} \in \mathcal{G}_m \text{ tels que}$$
$$(\forall i \in \{1,\ldots,m\}) \quad r_i \in L_i\left(J_{\rho^{-1}A}\left(\rho^{-1}\left(z - \sum_{j=1}^m L_j^*\overline{v_j}\right)\right)\right) - B_i^{-1}\overline{v_i} - D_i^{-1}\overline{v_i}. \tag{6.6}$$

Voici quelques propriétés préliminaires.

Proposition 6.4 *Dans le Problème 6.3, posons*

$$\overline{x} = J_{\rho^{-1}M}\big(\rho^{-1}z\big), \quad \text{où} \quad M = A + \sum_{i=1}^{m} L_i^* \circ (B_i \,\square\, D_i) \circ (L_i \cdot - r_i). \tag{6.7}$$

Alors nous avons les résultats suivants.

(i) $\overline{x}$ *est la solution unique de l'inclusion primale* (6.5).

(ii) *L'inclusion duale* (6.6) *admet au moins une solution.*

(iii) *Soit* $(\overline{v_1}, \ldots, \overline{v_m})$ *une solution de* (6.6). *Alors* $\overline{x} = J_{\rho^{-1}A}\big(\rho^{-1}\big(z - \sum_{i=1}^m L_i^*\overline{v_i}\big)\big)$.

(iv) *La condition* (6.4) *est vérifiée pour tout* z *dans* $\mathcal{H}$ *si et seulement si* M *est maximalement monotone. C'est le cas lorsque l'une des conditions suivantes est vérifiée.*

(a) *L'enveloppe conique de*

$$E = \left\{ \big(L_i x - r_i - v_i\big)_{1\leqslant i\leqslant m} \;\middle|\; x \in \operatorname{dom} A \text{ et } (v_i)_{1\leqslant i\leqslant m} \in \bigtimes_{i=1}^{m} \operatorname{ran}\big(B_i^{-1} + D_i^{-1}\big) \right\} \tag{6.8}$$

est un sous-espace vectoriel fermé.

(b) $A = \partial f$ *avec* $f \in \Gamma_0(\mathcal{H})$, *et pour tout* $i \in \{1, \ldots, m\}$, $B_i = \partial g_i$ *avec* $g_i \in \Gamma_0(\mathcal{G}_i)$ *et* $D_i = \partial \ell_i$ *où* $\ell_i \in \Gamma_0(\mathcal{G}_i)$ *est une fonction fortement convexe, et l'une des conditions suivantes est vérifiée.*

1/ $(r_1, \ldots, r_m) \in \operatorname{sri}\big\{(L_i x - y_i)_{1\leqslant i\leqslant m} \;\big|\; x \in \operatorname{dom} f$ *et*

$$(\forall i \in \{1, \ldots, m\}) \;\; y_i \in \operatorname{dom} g_i + \operatorname{dom} \ell_i\big\}.$$

2/ *Pour tout* $i \in \{1, \ldots, m\}$, g_i *ou* ℓ_i *est une fonction à valeurs réelles.*

3/ $\mathcal{H}$ *et* $(\mathcal{G}_i)_{1\leqslant i\leqslant m}$ *sont de dimensions finies, et il existe* $x \in \operatorname{ri}\operatorname{dom} f$ *tel que*

$$(\forall i \in \{1, \ldots, m\}) \quad L_i x - r_i \in \operatorname{ri}\operatorname{dom} g_i + \operatorname{ri}\operatorname{dom} \ell_i. \tag{6.9}$$

En appliquant la méthode explicite-implicite à métrique variable (6.3) au problème dual (6.6), nous obtenons l'algorithme primal-dual suivant.

Corollaire 6.5 *Dans le Problème 6.3, posons*

$$\beta = \frac{1}{\displaystyle\max_{1\leqslant i\leqslant m} \frac{1}{\nu_i} + \frac{1}{\rho} \sum_{1\leqslant i\leqslant m} \|L_i\|^2}. \tag{6.10}$$

Soit $(a_n)_{n\in\mathbb{N}}$ *une suite absolument sommable dans* $\mathcal{H}$, *soit* $\alpha \in \,]0,+\infty[$, *et soit* $(\eta_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N})$. *Pour tout* $i \in \{1,\ldots,m\}$, *soit* $v_{i,0} \in \mathcal{G}_i$, *soient* $(b_{i,n})_{n\in\mathbb{N}}$ *et* $(d_{i,n})_{n\in\mathbb{N}}$ *des suites absolument sommables dans* $\mathcal{G}_i$, *et soit* $(U_{i,n})_{n\in\mathbb{N}}$ *une suite dans* $\mathcal{P}_\alpha(\mathcal{G}_i)$. *Supposons que*

$$\mu = \max_{1\leqslant i\leqslant m}\sup_{n\in\mathbb{N}}\|U_{i,n}\| < +\infty \quad \textit{et} \quad (\forall i \in \{1,\ldots,m\})(\forall n\in\mathbb{N}) \quad (1+\eta_n)U_{i,n+1} \succcurlyeq U_{i,n}. \tag{6.11}$$

Soit $\varepsilon \in \,]0,\min\{1,2\beta/(\mu+1)\}]$, *soit* $(\lambda_n)_{n\in\mathbb{N}}$ *une suite dans* $[\varepsilon,1]$, *et soit* $(\gamma_n)_{n\in\mathbb{N}}$ *une suite dans* $[\varepsilon,(2\beta-\varepsilon)/\mu]$. *Posons*

$$(\forall n\in\mathbb{N}) \quad \left\lfloor \begin{array}{l} s_n = z - \sum_{i=1}^m L_i^* v_{i,n} \\ x_n = J_{\rho^{-1}A}(\rho^{-1}s_n) + a_n \\ \text{Pour } i = 1,\ldots,m \\ \left\lfloor \begin{array}{l} w_{i,n} = v_{i,n} + \gamma_n U_{i,n}\big(L_i x_n - r_i - D_i^{-1}v_{i,n} - d_{i,n}\big) \\ v_{i,n+1} = v_{i,n} + \lambda_n\Big(J_{\gamma_n U_{i,n} B_i^{-1}}(w_{i,n}) + b_{i,n} - v_{i,n}\Big). \end{array} \right. \end{array} \right. \tag{6.12}$$

Alors, nous avons les résultats suivants pour la solution $\overline{x}$ *du problème* (6.5) *et pour une solution* $(\overline{v_1},\ldots,\overline{v_m})$ *du problème* (6.6).

(i) $(\forall i\in\{1,\ldots,m\})$ $v_{i,n} \rightharpoonup \overline{v_i}$. *De plus,* $\overline{x} = J_{\rho^{-1}A}\big(\rho^{-1}\big(z - \sum_{i=1}^m L_i^*\overline{v_i}\big)\big)$.

(ii) $x_n \to \overline{x}$.

On voit que l'algorithme (2.6) et l'algorithme (3.2) sont deux cas particuliers de (6.12). Des applications de l'algorithme (6.12) aux problèmes variationnels et problèmes de meilleure approximation sont présentées dans les Exemples 6.33 et 6.34, respectivement.

Corollaire 6.6 *Dans le Problème* 4.3, *mettons* $(\forall i\in\{1,\ldots,m\})$ $\omega_i = 1$, *et supposons que*

$$z \in \operatorname{ran}\Big(A + \sum_{i=1}^m L_i^*\big((B_i \,\square\, D_i)(L_i\cdot - r_i)\big) + C\Big), \tag{6.13}$$

et posons

$$\beta = \min\{\mu,\nu_1,\ldots,\nu_m\}. \tag{6.14}$$

Soit $\varepsilon \in \,]0,\min\{1,\beta\}[$, *soit* $\alpha\in\,]0,+\infty[$, *soit* $(\lambda_n)_{n\in\mathbb{N}}$ *une suite dans* $[\varepsilon,1]$, *soit* $x_0\in\mathcal{H}$, *soient* $(a_n)_{n\in\mathbb{N}}$ *et* $(c_n)_{n\in\mathbb{N}}$ *des suites absolument sommables dans* $\mathcal{H}$, *et soit* $(U_n)_{n\in\mathbb{N}}$ *une suite dans* $\mathcal{P}_\alpha(\mathcal{H})$ *telle que* $(\forall n\in\mathbb{N})$ $U_{n+1}\succcurlyeq U_n$. *Pour tout* $i\in\{1,\ldots,m\}$, *soit* $v_{i,0}\in\mathcal{G}_i$, *et soient* $(b_{i,n})_{n\in\mathbb{N}}$ *et* $(d_{i,n})_{n\in\mathbb{N}}$ *des suites absolument sommables dans* $\mathcal{G}_i$, *et soit* $(U_{i,n})_{n\in\mathbb{N}}$ *une suite dans* $\mathcal{P}_\alpha(\mathcal{G}_i)$ *telle que* $(\forall n\in\mathbb{N})$ $U_{i,n+1}\succcurlyeq U_{i,n}$. *Pour tout* $n\in\mathbb{N}$, *posons*

$$\delta_n = \left(\sqrt{\sum_{i=1}^m \|\sqrt{U_{i,n}}L_i\sqrt{U_n}\|^2}\right)^{-1} - 1, \tag{6.15}$$

et supposons que

$$\zeta_n = \frac{\delta_n}{(1+\delta_n)\max\{\|U_n\|, \|U_{1,n}\|, \ldots, \|U_{m,n}\|\}} \geqslant \frac{1}{2\beta - \varepsilon}. \tag{6.16}$$

Posons

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} p_n = J_{U_n A}\Big(x_n - U_n\big(\sum_{i=1}^m L_i^* v_{i,n} + Cx_n + c_n - z\big)\Big) + a_n \\ y_n = 2p_n - x_n \\ x_{n+1} = x_n + \lambda_n(p_n - x_n) \\ \text{Pour } i = 1, \ldots, m \\ \left\lfloor \begin{array}{l} q_{i,n} = J_{U_{i,n} B_i^{-1}}\Big(v_{i,n} + U_{i,n}\big(L_i y_n - D_i^{-1} v_{i,n} - d_{i,n} - r_i\big)\Big) + b_{i,n} \\ v_{i,n+1} = v_{i,n} + \lambda_n(q_{i,n} - v_{i,n}). \end{array} \right. \end{array} \right. \tag{6.17}$$

Alors, on a les résultats suivants pour un point $\overline{x} \in \mathcal{P}$ et pour un point $(\overline{v_1}, \ldots, \overline{v_m}) \in \mathcal{D}$.

(i) $x_n \rightharpoonup \overline{x}$.

(ii) $(\forall i \in \{1, \ldots, m\})\ v_{i,n} \rightharpoonup \overline{v_i}$.

(iii) *Supposons que C soit demirégulier en $\overline{x}$. Alors $x_n \to \overline{x}$.*

(iv) *Supposons que, pour quelque $j \in \{1, \ldots, m\}$, D_j^{-1} soit demirégulier en $\overline{v_j}$. Alors $v_{j,n} \to \overline{v_j}$.*

Le petit exemple suivant illustre la convergence de la méthode explicite-implicite à métrique variable en comparaison avec le cas de la métrique constante.

Exemple 6.7 Considérons le système d'équations dans $\mathbb{R}^2$ de trouver $(\xi_1, \xi_2) \in \mathbb{R}^2$ tel que

$$\begin{cases} \xi_1 = 0, \\ -5\xi_1 + \xi_2 = 0. \end{cases} \tag{6.18}$$

Notons $H_1 = \big\{(\xi_1, \xi_2) \ \big|\ \xi_1 = 0\big\} \subset \mathbb{R}^2$ et $H_2 = \big\{(\xi_1, \xi_2) \ \big|\ -5\xi_1 + \xi_2 = 0\big\} \subset \mathbb{R}^2$. Le problème (6.18) est équivalent au problème de trouver un point dans $H_1 \cap H_2$. Nous utilisons la méthode de projection alternées à métrique variable (Var-POCS),

$$x_0 = (15, 15) \in \mathbb{R}^2 \quad \text{et} \quad (\forall n \in \mathbb{N}) \quad x_{n+1} = P_{H_1}^{U_n^{-1}} P_{H_2}^{U_n^{-1}} x_n, \tag{6.19}$$

où

$$U_0 = \begin{bmatrix} 5 & -1 \\ -1 & 1.5 \end{bmatrix} \quad \text{et} \quad (\forall n \in \mathbb{N}) \quad U_{n+1} = U_0 + \sum_{k=0}^{n} \mathrm{Id}\,/k^2. \tag{6.20}$$

Cette méthode est un cas particulier de la méthode explicite-implicite [18] à métrique variable avec les paramètres suivants,

$$(\forall n \in \mathbb{N}) \quad \gamma_n = 1, \quad \lambda_n = 1, \quad a_n = 0, \quad b_n = 0. \tag{6.21}$$

Nous comparons cette méthode avec la méthode de projection alternées à métrique constante (POCS),

$$y_0 = x_0 = (15, 15) \in \mathbb{R}^2 \quad \text{et} \quad (\forall n \in \mathbb{N}) \quad y_{n+1} = P_{H_1} P_{H_2} y_n. \tag{6.22}$$

Nous obtenons les résultats dans les figures 6.1 et 6.2.

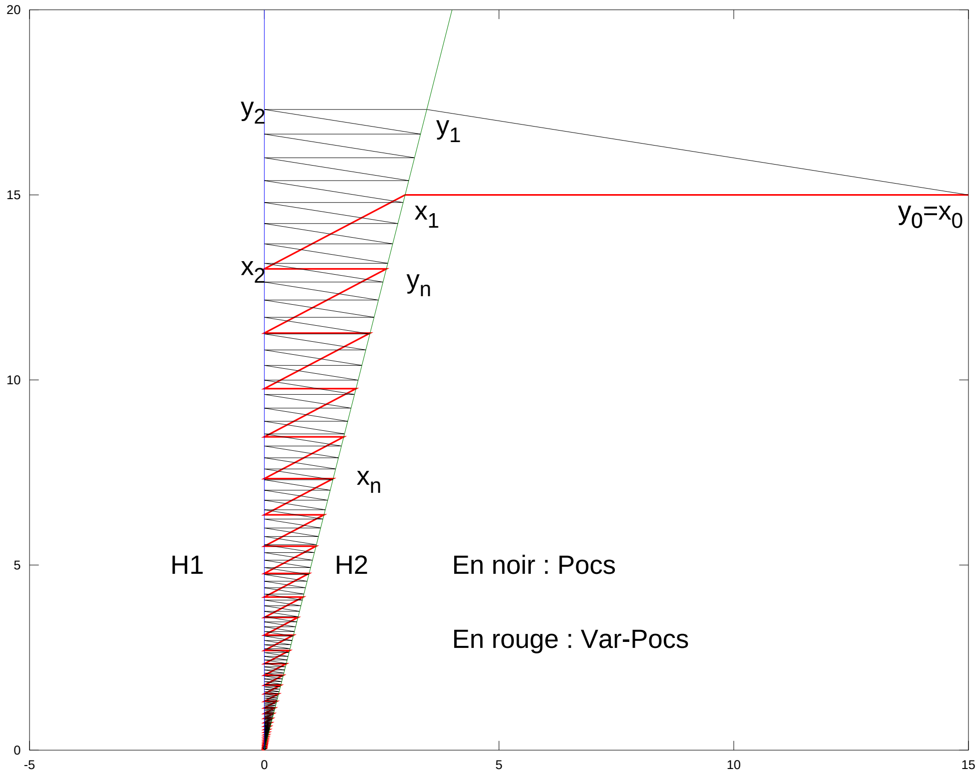

FIGURE 6.1 – Les suites $(x_n)_{n\in\mathbb{N}}$ produite par Var-POCS et $(y_n)_{n\in\mathbb{N}}$ produite par POCS.

On voit que la méthode Var-POCS est plus rapide que la méthode POCS dans cet exemple.

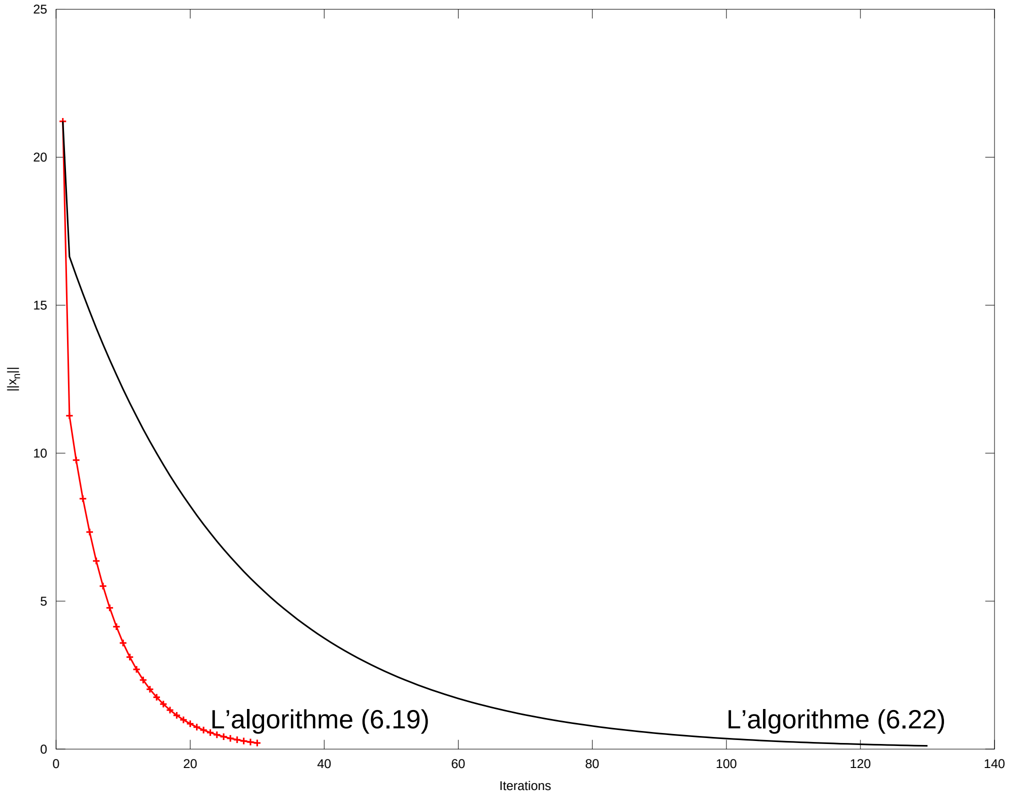

FIGURE 6.2 – La convergence de Var-POCS et POCS.

6.2 Article en anglais

VARIABLE METRIC FORWARD-BACKWARD SPLITTING WITH APPLICATIONS TO MONOTONE INCLUSIONS IN DUALITY [1]

Abstract : We propose a variable metric forward-backward splitting algorithm and prove its convergence in real Hilbert spaces. We then use this framework to derive primal-dual splitting algorithms for solving various classes of monotone inclusions in duality. Some of these algorithms are new even when specialized to the fixed metric case. Various applications are discussed.

1. P. L. Combettes, B. C. Vũ, Variable metric forward-backward splitting with applications to monotone inclusions in duality, *Optimization*, to appear
http://www.tandfonline.com/doi/full/10.1080/02331934.2012.733883

6.2.1 Introduction

The forward-backward algorithm has a long history going back to the projected gradient method (see [1, 12] for historical background). It addresses the problem of finding a zero of the sum of two operators acting on a real Hilbert space $\mathcal{H}$, namely,

$$\text{find} \quad x \in \mathcal{H} \quad \text{such that} \quad 0 \in Ax + Bx, \tag{6.23}$$

under the assumption that $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ is maximally monotone and that $B\colon \mathcal{H} \to \mathcal{H}$ is β-cocoercive for some $\beta \in \left]0, +\infty\right[$, i.e. [4],

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \langle x - y \mid Bx - By \rangle \geqslant \beta \|Bx - By\|^2. \tag{6.24}$$

This framework is quite central due to the large class of problems it encompasses in areas such as partial differential equations, mechanics, evolution inclusions, signal and image processing, best approximation, convex optimization, learning theory, inverse problems, statistics, game theory, and variational inequalities [1, 4, 7, 10, 12, 15, 18, 20, 21, 23, 24, 29, 30, 39, 40, 42]. The forward-backward algorithm operates according to the routine

$$x_0 \in \mathcal{H} \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad x_{n+1} = (\mathrm{Id} + \gamma_n A)^{-1}(x_n - \gamma_n B x_n), \quad \text{where} \quad 0 < \gamma_n < 2\beta. \tag{6.25}$$

In classical optimization methods, the benefits of changing the underlying metric over the course of the iterations to improve convergence profiles has long been recognized [19, 33]. In proximal methods, variable metrics have been investigated mostly when $B = 0$ in (6.23). In such instances (6.25) reduces to the proximal point algorithm

$$x_0 \in \mathcal{H} \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad x_{n+1} = (\mathrm{Id} + \gamma_n A)^{-1} x_n, \quad \text{where} \quad \gamma_n > 0. \tag{6.26}$$

In the case when A is the subdifferential of a real-valued convex function in a finite dimensional setting, variable metric versions of (6.26) have been proposed in [5, 11, 27, 35]. These methods draw heavily on the fact that the proximal point algorithm for minimizing a function corresponds to the gradient descent method applied to its Moreau envelope. In the same spirit, variable metric proximal point algorithms for a general maximally monotone operator A were considered in [8, 36]. In [8], superlinear convergence rates were shown to be achievable under suitable hypotheses (see also [9] for further developments). The finite dimensional variable metric proximal point algorithm proposed in [32] allows for errors in the proximal steps and features a flexible class of exogenous metrics to implement the algorithm. The first variable metric forward-backward algorithm appears to be that introduced in [10, Section 5]. It focuses on linear convergence results in the case when $A + B$ is strongly monotone and $\mathcal{H}$ is finite-dimensional. The variable metric splitting algorithm of [28] provides a framework which can be used to solve (6.23) in instances when $\mathcal{H}$ is finite-dimensional and B is merely Lipschitzian. However, it does not exploit the cocoercivity property (6.24) and it is more cumbersome to implement than the forward-backward iteration. Let us add that, in the important case when B is the gradient of a convex function, the Baillon-Haddad theorem asserts that the notions of cocoercivity and Lipschitz-continuity coincide [4, Corollary 18.16].

The goal of this paper is two-fold. First, we propose a general purpose variable metric forward-backward algorithm to solve (6.23)–(6.24) in Hilbert spaces and analyze its asymptotic behavior, both in terms of weak and strong convergence. Second, we show that this algorithm can be used to solve a broad class of composite monotone inclusion problems in duality by formulating them as instances of (6.23)–(6.24) in alternate Hilbert spaces. Even when restricted to the constant metric case, some of these results are new.

The paper is organized as follows. Section 6.2.2 is devoted to notation and background. In Section 6.2.3, we provide preliminary results. The variable metric forward-backward algorithm is introduced and analyzed in Section 6.2.4. In Section 6.2.5, we present a new variable metric primal-dual splitting algorithm for strongly monotone composite inclusions. This algorithm is obtained by applying the forward-backward algorithm of Section 6.2.4 to the dual inclusion. In Section 6.2.6, we consider a more general class of composite inclusions in duality and show that they can be solved by applying the forward-backward algorithm of Section 6.2.4 to a certain inclusion problem posed in the primal-dual product space. Applications to minimization problems, variational inequalities, and best approximation are discussed.

6.2.2 Notation and background

We recall some notation and background from convex analysis and monotone operator theory (see [4] for a detailed account).

Throughout, $\mathcal{H}$, $\mathcal{G}$, and $(\mathcal{G}_i)_{1\leqslant i\leqslant m}$ are real Hilbert spaces. We denote the scalar product of a Hilbert space by $\langle\cdot\mid\cdot\rangle$ and the associated norm by $\|\cdot\|$. The symbols $\rightharpoonup$ and $\to$ denote respectively weak and strong convergence, and Id denotes the identity operator. We denote by $\mathcal{B}(\mathcal{H},\mathcal{G})$ the space of bounded linear operators from $\mathcal{H}$ to $\mathcal{G}$, we set $\mathcal{B}(\mathcal{H})=\mathcal{B}(\mathcal{H},\mathcal{H})$ and $\mathcal{S}(\mathcal{H})=\big\{L\in\mathcal{B}(\mathcal{H})\;\big|\;L=L^*\big\}$, where L^* denotes the adjoint of L. The Loewner partial ordering on $\mathcal{S}(\mathcal{H})$ is defined by

$$(\forall U\in\mathcal{S}(\mathcal{H}))(\forall V\in\mathcal{S}(\mathcal{H}))\quad U\succcurlyeq V\quad\Leftrightarrow\quad(\forall x\in\mathcal{H})\quad\langle Ux\mid x\rangle\geqslant\langle Vx\mid x\rangle. \tag{6.27}$$

Now let $\alpha\in\left[0,+\infty\right[$. We set

$$\mathcal{P}_\alpha(\mathcal{H})=\big\{U\in\mathcal{S}(\mathcal{H})\;\big|\;U\succcurlyeq\alpha\,\mathrm{Id}\big\}, \tag{6.28}$$

and we denote by $\sqrt{U}$ the square root of $U\in\mathcal{P}_\alpha(\mathcal{H})$. Moreover, for every $U\in\mathcal{P}_\alpha(\mathcal{H})$, we define a semi-scalar product and a semi-norm (a scalar product and a norm if $\alpha>0$)

by

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \langle x \mid y\rangle_U = \langle Ux \mid y\rangle \quad \text{and} \quad \|x\|_U = \sqrt{\langle Ux \mid x\rangle}. \tag{6.29}$$

Notation 6.8 *We denote by $\boldsymbol{\mathcal{G}} = \mathcal{G}_1 \oplus \cdots \oplus \mathcal{G}_m$ the Hilbert direct sum of the Hilbert spaces $(\mathcal{G}_i)_{1\leqslant i\leqslant m}$, i.e., their product space equipped with the scalar product and the associated norm respectively defined by*

$$\langle\langle \cdot \mid \cdot\rangle\rangle \colon (\boldsymbol{x}, \boldsymbol{y}) \mapsto \sum_{i=1}^{m} \langle x_i \mid y_i\rangle \quad \textit{and} \quad |||\cdot||| \colon \boldsymbol{x} \mapsto \sqrt{\sum_{i=1}^{m} \|x_i\|^2}, \tag{6.30}$$

where $\boldsymbol{x} = (x_i)_{1\leqslant i\leqslant m}$ and $\boldsymbol{y} = (y_i)_{1\leqslant i\leqslant m}$ denote generic elements in $\boldsymbol{\mathcal{G}}$.

Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a set-valued operator. The domain and the graph of A are respectively defined by $\operatorname{dom} A = \big\{x \in \mathcal{H} \;\big|\; Ax \neq \varnothing\big\}$ and $\operatorname{gra} A = \big\{(x,u) \in \mathcal{H} \times \mathcal{H} \;\big|\; u \in Ax\big\}$. We denote by $\operatorname{zer} A = \big\{x \in \mathcal{H} \;\big|\; 0 \in Ax\big\}$ the set of zeros of A and by $\operatorname{ran} A = \big\{u \in \mathcal{H} \;\big|\; (\exists\, x \in \mathcal{H})\; u \in Ax\big\}$ the range of A. The inverse of A is $A^{-1}\colon \mathcal{H} \mapsto 2^{\mathcal{H}}\colon u \mapsto \big\{x \in \mathcal{H} \;\big|\; u \in Ax\big\}$, and the resolvent of A is

$$J_A = (\operatorname{Id} + A)^{-1}. \tag{6.31}$$

Moreover, A is monotone if

$$(\forall (x,y) \in \mathcal{H} \times \mathcal{H})(\forall (u,v) \in Ax \times Ay) \quad \langle x - y \mid u - v\rangle \geqslant 0, \tag{6.32}$$

and maximally monotone if it is monotone and there exists no monotone operator $B\colon \mathcal{H} \to 2^{\mathcal{H}}$ such that $\operatorname{gra} A \subset \operatorname{gra} B$ and $A \neq B$. The parallel sum of A and $B\colon \mathcal{H} \to 2^{\mathcal{H}}$ is

$$A \,\square\, B = (A^{-1} + B^{-1})^{-1}. \tag{6.33}$$

The conjugate of $f\colon \mathcal{H} \to \left]-\infty, +\infty\right]$ is

$$f^*\colon \mathcal{H} \to [-\infty, +\infty] : u \mapsto \sup_{x\in\mathcal{H}} \big(\langle x \mid u\rangle - f(x)\big), \tag{6.34}$$

and the infimal convolution of f with $g\colon \mathcal{H} \to \left]-\infty, +\infty\right]$ is

$$f \,\square\, g\colon \mathcal{H} \to [-\infty, +\infty] : x \mapsto \inf_{y\in\mathcal{H}} \big(f(y) + g(x-y)\big). \tag{6.35}$$

The class of lower semicontinuous convex functions $f\colon \mathcal{H} \to \left]-\infty, +\infty\right]$ such that $\operatorname{dom} f = \big\{x \in \mathcal{H} \;\big|\; f(x) < +\infty\big\} \neq \varnothing$ is denoted by $\Gamma_0(\mathcal{H})$. If $f \in \Gamma_0(\mathcal{H})$, then $f^* \in \Gamma_0(\mathcal{H})$ and the subdifferential of f is the maximally monotone operator

$$\partial f\colon \mathcal{H} \to 2^{\mathcal{H}}\colon x \mapsto \big\{u \in \mathcal{H} \;\big|\; (\forall y \in \mathcal{H})\; \langle y - x \mid u\rangle + f(x) \leqslant f(y)\big\} \tag{6.36}$$

with inverse $(\partial f)^{-1} = \partial f^*$. Let C be a nonempty subset of $\mathcal{H}$. The indicator function and the distance function of C are defined on $\mathcal{H}$ as

$$\iota_C \colon x \mapsto \begin{cases} 0, & \text{if } x \in C; \\ +\infty, & \text{if } x \notin C \end{cases} \qquad \text{and} \quad d_C = \iota_C \,\square\, \|\cdot\| \colon x \mapsto \inf_{y \in C} \|x - y\|. \tag{6.37}$$

respectively. The interior of C is $\operatorname{int} C$ and the support function of C is $\sigma_C = \iota_C^*$. Now suppose that C is convex. The normal cone operator of C is defined as

$$N_C = \partial \iota_C \colon \mathcal{H} \to 2^{\mathcal{H}} \colon x \mapsto \begin{cases} \big\{u \in \mathcal{H} \mid (\forall y \in C) \ \ \langle y - x \mid u \rangle \leqslant 0\big\}, & \text{if } x \in C; \\ \varnothing, & \text{otherwise.} \end{cases} \tag{6.38}$$

The strong relative interior of C, i.e., the set of points $x \in C$ such that the conical hull of $-x + C$ is a closed vector subspace of $\mathcal{H}$, is denoted by $\operatorname{sri} C$; if $\mathcal{H}$ is finite-dimensional, $\operatorname{sri} C$ coincides with the relative interior of C, denoted by $\operatorname{ri} C$. If C is also closed, its projector is denoted by P_C, i.e., $P_C \colon \mathcal{H} \to C \colon x \mapsto \operatorname{argmin}_{y \in C} \|x - y\|$.

Finally, $\ell_+^1(\mathbb{N})$ denotes the set of summable sequences in $[0, +\infty[$.

6.2.3 Preliminary results

6.2.3.1 Technical results

The following properties can be found in [26, Section VI.2.6] (see [17, Lemma 2.1] for an alternate short proof).

Lemma 6.9 *Let $\alpha \in\,]0, +\infty[$ and $\mu \in\,]0, +\infty[$, and assume that A and B are operators in $\mathcal{S}(\mathcal{H})$ such that $\mu \operatorname{Id} \succcurlyeq A \succcurlyeq B \succcurlyeq \alpha \operatorname{Id}$. Then the following hold.*

(i) $\alpha^{-1} \operatorname{Id} \succcurlyeq B^{-1} \succcurlyeq A^{-1} \succcurlyeq \mu^{-1} \operatorname{Id}$.

(ii) $(\forall x \in \mathcal{H})\ \langle A^{-1}x \mid x \rangle \geqslant \|A\|^{-1}\|x\|^2$.

(iii) $\|A^{-1}\| \leqslant \alpha^{-1}$.

The next fact concerns sums of composite cocoercive operators.

Proposition 6.10 *Let I be a finite index set. For every $i \in I$, let $0 \neq L_i \in \mathcal{B}(\mathcal{H}, \mathcal{G}_i)$, let $\beta_i \in\,]0, +\infty[$, and let $T_i \colon \mathcal{G}_i \to \mathcal{G}_i$ be β_i-cocoercive. Set $T = \sum_{i \in I} L_i^* T_i L_i$ and $\beta = 1/\big(\sum_{i \in I} \|L_i\|^2/\beta_i\big)$. Then T is β-cocoercive.*

Proof. Set $(\forall i \in I)\ \alpha_i = \beta\|L_i\|^2/\beta_i$. Then $\sum_{i\in I} \alpha_i = 1$ and, using the convexity of $\|\cdot\|^2$ and (6.24), we have

$$
\begin{aligned}
(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \langle x - y \mid Tx - Ty\rangle &= \sum_{i\in I} \langle x - y \mid L_i^* T_i L_i x - L_i^* T_i L_i y\rangle \\
&= \sum_{i\in I} \langle L_i x - L_i y \mid T_i L_i x - T_i L_i y\rangle \\
&\geqslant \sum_{i\in I} \beta_i \|T_i L_i x - T_i L_i y\|^2 \\
&\geqslant \sum_{i\in I} \frac{\beta_i}{\|L_i\|^2} \|L_i^* T_i L_i x - L_i^* T_i L_i y\|^2 \\
&= \beta \sum_{i\in I} \alpha_i \left\| \frac{1}{\alpha_i}(L_i^* T_i L_i x - L_i^* T_i L_i y)\right\|^2 \\
&\geqslant \beta \left\| \sum_{i\in I} (L_i^* T_i L_i x - L_i^* T_i L_i y)\right\|^2 \\
&= \beta\|Tx - Ty\|^2, \qquad (6.39)
\end{aligned}
$$

which concludes the proof. □

6.2.3.2 Variable metric quasi-Fejér sequences

The following results are from [17].

Proposition 6.11 *Let $\alpha \in \left]0, +\infty\right[$, let $(W_n)_{n\in\mathbb{N}}$ be in $\mathcal{P}_\alpha(\mathcal{H})$, let C be a nonempty subset of $\mathcal{H}$, and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that*

$$
\begin{aligned}
&\big(\exists\, (\eta_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N})\big)\big(\forall z \in C\big)\big(\exists\, (\varepsilon_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N})\big)(\forall n \in \mathbb{N}) \\
&\qquad \|x_{n+1} - z\|_{W_{n+1}} \leqslant (1 + \eta_n)\|x_n - z\|_{W_n} + \varepsilon_n. \qquad (6.40)
\end{aligned}
$$

Then $(x_n)_{n\in\mathbb{N}}$ is bounded and, for every $z \in C$, $(\|x_n - z\|_{W_n})_{n\in\mathbb{N}}$ converges.

Proposition 6.12 *Let $\alpha \in \left]0, +\infty\right[$, and let $(W_n)_{n\in\mathbb{N}}$ and W be operators in $\mathcal{P}_\alpha(\mathcal{H})$ such that $W_n \to W$ pointwise as $n \to +\infty$, as is the case when*

$$
\sup_{n\in\mathbb{N}} \|W_n\| < +\infty \quad \text{and} \quad (\exists\, (\eta_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N}))(\forall n \in \mathbb{N}) \quad (1 + \eta_n)W_n \succcurlyeq W_{n+1}. \qquad (6.41)
$$

Let C be a nonempty subset of $\mathcal{H}$, and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that (6.40) is satisfied. Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in C if and only if every weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ is in C.

Proposition 6.13 *Let $\alpha \in \left]0,+\infty\right[$, let $(W_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{H})$ such that $\sup_{n\in\mathbb{N}}\|W_n\| < +\infty$, let C be a nonempty closed subset of $\mathcal{H}$, and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that*

$$\big(\exists\,(\varepsilon_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})\big)\big(\exists\,(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})\big)(\forall z \in C)(\forall n \in \mathbb{N}) \\ \|x_{n+1} - z\|_{W_{n+1}} \leqslant (1+\eta_n)\|x_n - z\|_{W_n} + \varepsilon_n. \quad (6.42)$$

Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to a point in C if and only if $\underline{\lim}\, d_C(x_n) = 0$.

Proposition 6.14 *Let $\alpha \in \left]0,+\infty\right[$, let $(\nu_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$, and let $(W_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{H})$ such that $\sup_{n\in\mathbb{N}}\|W_n\| < +\infty$ and $(\forall n \in \mathbb{N})$ $(1+\nu_n)W_{n+1} \succcurlyeq W_n$. Furthermore, let C be a subset of $\mathcal{H}$ such that $\operatorname{int} C \neq \varnothing$, let $z \in C$ and $\rho \in \left]0,+\infty\right[$ be such that $B(z;\rho) \subset C$, and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that*

$$\big(\exists\,(\varepsilon_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})\big)\big(\exists\,(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})\big)(\forall x \in B(z;\rho))(\forall n \in \mathbb{N}) \\ \|x_{n+1} - x\|^2_{W_{n+1}} \leqslant (1+\eta_n)\|x_n - x\|^2_{W_n} + \varepsilon_n. \quad (6.43)$$

Then $(x_n)_{n\in\mathbb{N}}$ converges strongly.

6.2.3.3 Monotone operators

We establish some results on monotone operators in a variable metric environment.

Lemma 6.15 *Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone, let $\alpha \in \left]0,+\infty\right[$, let $U \in \mathcal{P}_\alpha(\mathcal{H})$, and let $\mathcal{G}$ be the real Hilbert space obtained by endowing $\mathcal{H}$ with the scalar product $(x,y) \mapsto \langle x \mid y\rangle_{U^{-1}} = \langle x \mid U^{-1}y\rangle$. Then the following hold.*

(i) *$UA\colon \mathcal{G} \to 2^{\mathcal{G}}$ is maximally monotone.*

(ii) *$J_{UA}\colon \mathcal{G} \to \mathcal{G}$ is 1-cocoercive, i.e., firmly nonexpansive, hence nonexpansive.*

(iii) *$J_{UA} = (U^{-1} + A)^{-1} \circ U^{-1}$.*

Proof. (i) : Set $B = UA$ and $V = U^{-1}$. For every $(x,u) \in \operatorname{gra} B$ and every $(y,v) \in \operatorname{gra} B$, $Vu \in VBx = Ax$ and $Vv \in VBy = Ay$, so that

$$\langle x - y \mid u - v\rangle_V = \langle x - y \mid Vu - Vv\rangle \geqslant 0 \quad (6.44)$$

by monotonicity of A on $\mathcal{H}$. This shows that B is monotone on $\mathcal{G}$. Now let $(y,v) \in \mathcal{H}^2$ be such that

$$(\forall (x,u) \in \operatorname{gra} B) \quad \langle x - y \mid u - v\rangle_V \geqslant 0. \quad (6.45)$$

Then, for every $(x,u) \in \operatorname{gra} A$, $(x,Uu) \in \operatorname{gra} B$ and we derive from (6.45) that

$$\langle x - y \mid u - Vv\rangle = \langle x - y \mid Uu - v\rangle_V \geqslant 0. \quad (6.46)$$

Since A is maximally monotone on $\mathcal{H}$, (6.46) gives $(y, Vv) \in \operatorname{gra} A$, which implies that $(y, v) \in \operatorname{gra} B$. Hence, B is maximally monotone on $\mathcal{G}$.

(ii) : This follows from (i) and [4, Corollary 23.8].

(iii) : Let x and p be in $\mathcal{G}$. Then $p = J_{UA}x \Leftrightarrow x \in p + UAp \Leftrightarrow U^{-1}x \in (U^{-1} + A)p \Leftrightarrow p = (U^{-1} + A)^{-1}(U^{-1}x)$. □

Remark 6.16 let $\alpha \in \left]0, +\infty\right[$, let $U \in \mathcal{P}_\alpha(\mathcal{H})$, set $f\colon \mathcal{H} \to \mathbb{R}\colon x \mapsto \langle U^{-1}x \mid x\rangle/2$, and let $D\colon (x, y) \mapsto f(x) - f(y) - \langle x - y \mid \nabla f(y)\rangle$ be the associated Bregman distance. Then Lemma 6.15(iii) asserts that $J_{UA} = (\nabla f + A)^{-1} \circ \nabla f$. In other words, J_{UA} is the D-resolvent of A introduced in [3, Definition 3.7].

Let $U \in \mathcal{P}_\alpha(\mathcal{H})$ for some $\alpha \in \left]0, +\infty\right[$. The proximity operator of $f \in \Gamma_0(\mathcal{H})$ relative to the metric induced by U is [25, Section XV.4]

$$\operatorname{prox}^U_f\colon \mathcal{H} \to \mathcal{H}\colon x \mapsto \underset{y\in\mathcal{H}}{\operatorname{argmin}} \left(f(y) + \frac{1}{2}\|x - y\|^2_U\right), \tag{6.47}$$

and the projector onto a nonempty closed convex subset C of $\mathcal{H}$ relative to the norm $\|\cdot\|_U$ is denoted by P^U_C. We have

$$\operatorname{prox}^U_f = J_{U^{-1}\partial f} \quad \text{and} \quad P^U_C = \operatorname{prox}^U_{\iota_C}, \tag{6.48}$$

and we write $\operatorname{prox}^{\mathrm{Id}}_f = \operatorname{prox}_f$.

In the case when $U = \mathrm{Id}$ in Lemma 6.15, examples of closed form expressions for J_{UA} and basic resolvent calculus rules can be found in [4, 15, 18]. A few examples illustrating the case when $U \neq \mathrm{Id}$ are provided below. The first result is an extension of the well-known resolvent identity $J_A + J_{A^{-1}} = \mathrm{Id}$.

Example 6.17 Let $\alpha \in \left]0, +\infty\right[$, let $\gamma \in \left]0, +\infty\right[$, and let $U \in \mathcal{P}_\alpha(\mathcal{H})$. Then the following hold.

(i) Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone. Then

$$J_{\gamma UA} = \sqrt{U} J_{\gamma\sqrt{U}A\sqrt{U}} \sqrt{U}^{-1} = \mathrm{Id} - \gamma U J_{\gamma^{-1}U^{-1}A^{-1}}(\gamma^{-1}U^{-1}). \tag{6.49}$$

(ii) Let $f \in \Gamma_0(\mathcal{H})$. Then

$$\operatorname{prox}^U_{\gamma f} = \sqrt{U}^{-1} \operatorname{prox}_{\gamma f\circ\sqrt{U}^{-1}} \sqrt{U} = \mathrm{Id} - \gamma U^{-1} \operatorname{prox}^{U^{-1}}_{\gamma^{-1}f^*}(\gamma^{-1}U).$$

(iii) Let C be a nonempty closed convex subset of $\mathcal{H}$. Then

$$\operatorname{prox}^U_{\gamma\sigma_C} = \sqrt{U}^{-1} \operatorname{prox}_{\gamma\sigma_C\circ\sqrt{U}^{-1}} \sqrt{U} = \mathrm{Id} - \gamma U^{-1} P^{U^{-1}}_C(\gamma^{-1}U).$$

Proof. (i) : Let x and p be in $\mathcal{H}$. Then

$$\begin{aligned} p = J_{\gamma UA}x \quad &\Leftrightarrow \quad x - p \in \gamma UAp \\ &\Leftrightarrow \quad \sqrt{U}^{-1}x - \sqrt{U}^{-1}p \in \gamma\sqrt{U}A\sqrt{U}\sqrt{U}^{-1}p \\ &\Leftrightarrow \quad \sqrt{U}^{-1}p = J_{\gamma\sqrt{U}A\sqrt{U}}\big(\sqrt{U}^{-1}x\big) \\ &\Leftrightarrow \quad p = \sqrt{U}J_{\gamma\sqrt{U}A\sqrt{U}}\big(\sqrt{U}^{-1}x\big). \end{aligned} \tag{6.50}$$

Furthermore, by [4, Proposition 23.23(ii)], $J_{\sqrt{U}(\gamma A)\sqrt{U}} = \mathrm{Id} - \sqrt{U}\big(U + (\gamma A)^{-1}\big)^{-1}\sqrt{U}$. Hence, (6.50) yields

$$J_{\gamma UA} = \mathrm{Id} - U\big(U + (\gamma A)^{-1}\big)^{-1}. \tag{6.51}$$

However

$$\begin{aligned} p = \big(U + (\gamma A)^{-1}\big)^{-1}x \quad &\Leftrightarrow \quad x \in Up + (\gamma A)^{-1}p \\ &\Leftrightarrow \quad \gamma^{-1}p \in A(x - Up) \\ &\Leftrightarrow \quad x - Up \in A^{-1}(\gamma^{-1}p) \\ &\Leftrightarrow \quad \gamma^{-1}U^{-1}x \in \big(\mathrm{Id} + \gamma^{-1}U^{-1}A^{-1}\big)(\gamma^{-1}p) \\ &\Leftrightarrow \quad \gamma^{-1}p = J_{\gamma^{-1}U^{-1}A^{-1}}(\gamma^{-1}U^{-1}x). \end{aligned} \tag{6.52}$$

Hence, $(U + (\gamma A)^{-1})^{-1} = \gamma J_{\gamma^{-1}U^{-1}A^{-1}}(\gamma^{-1}U^{-1})$ and, using (6.51), we obtain the rightmost identity in (i).

(ii) : Apply (i) to $A = \partial f$, and use (6.48) and the fact that $\partial(f \circ \sqrt{U}^{-1}) = (\sqrt{U}^{-1})^* \circ (\partial f) \circ \sqrt{U}^{-1} = \sqrt{U}^{-1} \circ (\partial f) \circ \sqrt{U}^{-1}$ [4, Corollary 16.42(i)].

(iii) : Apply (ii) to $f = \sigma_C$, and use (6.48). ☐

Example 6.18 Define $\boldsymbol{\mathcal{G}}$ as in Notation 6.8, let $\alpha \in \mathbb{R}$, and, for every $i \in \{1,\ldots,m\}$, let $A_i\colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ be maximally monotone and let $U_i \in \mathcal{P}_\alpha(\mathcal{G}_i)$. Set $\boldsymbol{A}\colon \boldsymbol{\mathcal{G}} \to 2^{\boldsymbol{\mathcal{G}}}\colon (x_i)_{1\leqslant i\leqslant m} \mapsto \times_{i=1}^m A_i x_i$ and $\boldsymbol{U}\colon \boldsymbol{\mathcal{G}} \to \boldsymbol{\mathcal{G}}\colon (x_i)_{1\leqslant i\leqslant m} \mapsto (U_i x_i)_{1\leqslant i\leqslant m}$. Then $\boldsymbol{UA}$ is maximally monotone and

$$(\forall (x_i)_{1\leqslant i\leqslant m} \in \boldsymbol{\mathcal{G}}) \quad J_{\boldsymbol{UA}}(x_i)_{1\leqslant i\leqslant m} = (J_{U_iA_i}x_i)_{1\leqslant i\leqslant m}. \tag{6.53}$$

Proof. This follows from Lemma 6.15(i) and [4, Proposition 23.16]. ☐

Example 6.19 Let $\alpha \in \left]0,+\infty\right[$, let $\xi \in \mathbb{R}$, let $U \in \mathcal{P}_\alpha(\mathcal{H})$, let $\phi \in \Gamma_0(\mathbb{R})$, suppose that $0 \neq u \in \mathcal{H}$, and set $H = \big\{x \in \mathcal{H} \;\big|\; \langle x \mid u\rangle \leqslant \xi\big\}$ and $g = \phi(\langle\cdot \mid u\rangle)$. Then $g \in \Gamma_0(\mathcal{H})$ and

$$(\forall x \in \mathcal{H}) \quad \mathrm{prox}_g^U x = x + \frac{\mathrm{prox}_{\|\sqrt{U^{-1}}u\|^2\phi}\langle x \mid u\rangle - \langle x \mid u\rangle}{\|\sqrt{U^{-1}}u\|^2}U^{-1}u \tag{6.54}$$

and

$$P_H^U x = \begin{cases} x, & \text{if } \langle x \mid u\rangle \leqslant \xi; \\ x + \dfrac{\xi - \langle x \mid u\rangle}{\langle u \mid U^{-1}u\rangle} U^{-1}u, & \text{if } \langle x \mid u\rangle > \xi. \end{cases} \tag{6.55}$$

Proof. It follows from Example 6.17(ii) that

$$(\forall x \in \mathcal{H}) \quad \text{prox}_g^U x = \sqrt{U^{-1}}\text{prox}_{g\circ\sqrt{U^{-1}}}\sqrt{U}x. \tag{6.56}$$

Moreover, $g \circ \sqrt{U^{-1}} = \phi(\langle\cdot \mid \sqrt{U^{-1}}u\rangle)$. Hence, using (6.56) and [4, Corollary 23.33], we obtain

$$\begin{aligned}(\forall x \in \mathcal{H}) \quad \text{prox}_g^U x &= \sqrt{U^{-1}}\text{prox}_{\phi(\langle\cdot\mid\sqrt{U^{-1}}u\rangle)}\sqrt{U}x \\ &= x + \frac{\text{prox}_{\|\sqrt{U^{-1}}u\|^2\phi}\langle x \mid u\rangle - \langle x \mid u\rangle}{\|\sqrt{U^{-1}}u\|^2}U^{-1}u. \end{aligned}\tag{6.57}$$

Finally, upon setting $\phi = \iota_{]-\infty,\xi]}$, we obtain (6.55) from (6.54). □

Example 6.20 Let $\alpha \in \left]0,+\infty\right[$, let $\gamma \in \mathbb{R}$, let $A \in \mathcal{P}_0(\mathcal{H})$, let $u \in \mathcal{H}$, let $U \in \mathcal{P}_\alpha(\mathcal{H})$, and set $\varphi\colon \mathcal{H} \to \mathbb{R}\colon x \mapsto \langle Ax \mid x\rangle/2 + \langle x \mid u\rangle + \gamma$. Then $\varphi \in \Gamma_0(\mathcal{H})$ and

$$(\forall x \in \mathcal{H}) \quad \text{prox}_\varphi^U x = (\text{Id} + U^{-1}A)^{-1}(x - U^{-1}u). \tag{6.58}$$

Proof. Let $x \in \mathcal{H}$. Then $p = \text{prox}_\varphi^U x \Leftrightarrow x - p = U^{-1}\nabla\varphi(p) \Leftrightarrow x - p = U^{-1}(Ap + u) \Leftrightarrow x - U^{-1}u = (\text{Id} + U^{-1}A)p \Leftrightarrow p = (\text{Id} + U^{-1}A)^{-1}(x - U^{-1}u)$. □

Example 6.21 Let $\alpha \in \left]0,+\infty\right[$ and let $U \in \mathcal{P}_\alpha(\mathcal{H})$. For every $i \in \{1,\ldots,m\}$, let $r_i \in \mathcal{G}_i$, let $\omega_i \in \left]0,+\infty\right[$, and let $L_i \in \mathcal{B}(\mathcal{H},\mathcal{G}_i)$. Set $\varphi\colon x \mapsto (1/2)\sum_{i=1}^m \omega_i\|L_i x - r_i\|^2$. Then $\varphi \in \Gamma_0(\mathcal{H})$ and

$$(\forall x \in \mathcal{H}) \quad \text{prox}_\varphi^U x = \left(\text{Id} + U^{-1}\sum_{i=1}^m \omega_i L_i^* L_i\right)^{-1}\left(x + U^{-1}\sum_{i=1}^m \omega_i L_i^* r_i\right). \tag{6.59}$$

Proof. We have $\varphi\colon x \mapsto \langle Ax \mid x\rangle/2 + \langle x \mid u\rangle + \gamma$, where $A = \sum_{i=1}^m \omega_i L_i^* L_i$, $u = -\sum_{i=1}^m \omega_i L_i^* r_i$, and $\gamma = \sum_{i=1}^m \omega_i\|r_i\|^2/2$. Hence, (6.59) follows from (6.58). □

6.2.3.4 Demiregularity

Definition 6.22 [1, Definition 2.3] An operator $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ is *demiregular* at $x \in \text{dom}\,A$ if, for every sequence $((x_n,u_n))_{n\in\mathbb{N}}$ in $\text{gra}A$ and every $u \in Ax$ such that $x_n \rightharpoonup x$ and $u_n \to u$ as $n \to +\infty$, we have $x_n \to x$ as $n \to +\infty$.

Lemma 6.23 [1, Proposition 2.4] *Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be monotone and suppose that $x \in \operatorname{dom} A$. Then A is demiregular at x in each of the following cases.*

(i) *A is uniformly monotone at x, i.e., there exists an increasing function $\phi\colon [0,+\infty[\to [0,+\infty]$ that vanishes only at 0 such that $(\forall u \in Ax)(\forall (y,v) \in \operatorname{gra} A)$ $\langle x-y \mid u-v\rangle \geqslant \phi(\|x-y\|)$.*

(ii) *A is strongly monotone, i.e., there exists $\alpha \in]0,+\infty[$ such that $A-\alpha\operatorname{Id}$ is monotone.*

(iii) *J_A is compact, i.e., for every bounded set $C \subset \mathcal{H}$, the closure of $J_A(C)$ is compact. In particular, $\operatorname{dom} A$ is boundedly relatively compact, i.e., the intersection of its closure with every closed ball is compact.*

(iv) *$A\colon \mathcal{H} \to \mathcal{H}$ is single-valued with a single-valued continuous inverse.*

(v) *A is single-valued on $\operatorname{dom} A$ and $\operatorname{Id} - A$ is demicompact, i.e., for every bounded sequence $(x_n)_{n\in\mathbb{N}}$ in $\operatorname{dom} A$ such that $(Ax_n)_{n\in\mathbb{N}}$ converges strongly, $(x_n)_{n\in\mathbb{N}}$ admits a strong cluster point.*

(vi) *$A = \partial f$, where $f \in \Gamma_0(\mathcal{H})$ is uniformly convex at x, i.e., there exists an increasing function $\phi\colon [0,+\infty[\to [0,+\infty]$ that vanishes only at 0 such that $(\forall \alpha \in]0,1[)(\forall y \in \operatorname{dom} f)$ $f\big(\alpha x + (1-\alpha)y\big) + \alpha(1-\alpha)\phi(\|x-y\|) \leqslant \alpha f(x) + (1-\alpha)f(y)$.*

(vii) *$A = \partial f$, where $f \in \Gamma_0(\mathcal{H})$ and, for every $\xi \in \mathbb{R}$, $\big\{x \in \mathcal{H} \;\big|\; f(x) \leqslant \xi\big\}$ is boundedly compact.*

6.2.4 Algorithm and convergence

Our main result is stated in the following theorem.

Theorem 6.24 *Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone, let $\alpha \in]0,+\infty[$, let $\beta \in]0,+\infty[$, let $B\colon \mathcal{H} \to \mathcal{H}$ be β-cocoercive, let $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$, and let $(U_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{H})$ such that*

$$\mu = \sup_{n\in\mathbb{N}} \|U_n\| < +\infty \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad (1+\eta_n)U_{n+1} \succcurlyeq U_n. \tag{6.60}$$

Let $\varepsilon \in]0, \min\{1, 2\beta/(\mu+1)\}[$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon,1]$, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, (2\beta-\varepsilon)/\mu]$, let $x_0 \in \mathcal{H}$, and let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{H}$. Suppose that

$$Z = \operatorname{zer}(A+B) \neq \varnothing, \tag{6.61}$$

and set

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} y_n = x_n - \gamma_n U_n(Bx_n + b_n) \\ x_{n+1} = x_n + \lambda_n\big(J_{\gamma_n U_n A}(y_n) + a_n - x_n\big). \end{array} \right. \tag{6.62}$$

Then the following hold for some $\overline{x} \in Z$.

(i) $x_n \rightharpoonup \overline{x}$ *as* $n \to +\infty$.

(ii) $\sum_{n\in\mathbb{N}} \|Bx_n - B\overline{x}\|^2 < +\infty$.

(iii) *Suppose that one of the following holds.*

(a) $\underline{\lim}\, d_Z(x_n) = 0$.

(b) *At every point in* Z, A *or* B *is demiregular (see Lemma 6.23 for special cases).*

(c) $\operatorname{int} Z \neq \varnothing$ *and there exists* $(\nu_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N})$ *such that* $(\forall n \in \mathbb{N})$ $(1+\nu_n)U_n \succcurlyeq U_{n+1}$.

Then $x_n \to \overline{x}$ *as* $n \to +\infty$.

Proof. Set

$$(\forall n \in \mathbb{N}) \quad \begin{cases} A_n = \gamma_n U_n A \\ B_n = \gamma_n U_n B \end{cases} \quad \text{and} \quad \begin{cases} p_n = J_{A_n} y_n \\ q_n = J_{A_n}(x_n - B_n x_n) \\ s_n = x_n + \lambda_n(q_n - x_n). \end{cases} \tag{6.63}$$

Then (6.62) can be written as

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n(p_n + a_n - x_n). \tag{6.64}$$

On the other hand, (6.60) and Lemma 6.9(i)&(iii) yield

$$(\forall n \in \mathbb{N}) \quad \|U_n^{-1}\| \leqslant \frac{1}{\alpha}, \quad U_n^{-1} \in \mathcal{P}_{1/\mu}(\mathcal{H}), \quad \text{and} \quad (1+\eta_n)U_n^{-1} \succcurlyeq U_{n+1}^{-1} \tag{6.65}$$

and, therefore,

$$(\forall n \in \mathbb{N})(\forall x \in \mathcal{H}) \quad (1+\eta_n)\|x\|^2_{U_n^{-1}} \geqslant \|x\|^2_{U_{n+1}^{-1}}. \tag{6.66}$$

Hence, we derive from (6.64), (6.63), Lemma 6.15(ii), (6.65) and (6.60) that

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \|x_{n+1} - s_n\|_{U_n^{-1}} &\leqslant \lambda_n\Big(\|a_n\|_{U_n^{-1}} + \|p_n - q_n\|_{U_n^{-1}}\Big) \\ &\leqslant \|a_n\|_{U_n^{-1}} + \|y_n - x_n + B_n x_n\|_{U_n^{-1}} \\ &\leqslant \|a_n\|_{U_n^{-1}} + \gamma_n\|U_n b_n\|_{U_n^{-1}} \\ &\leqslant \sqrt{\|U_n^{-1}\|}\,\|a_n\| + \gamma_n\sqrt{\|U_n\|}\,\|b_n\| \\ &\leqslant \frac{1}{\sqrt{\alpha}}\|a_n\| + \frac{2\beta-\varepsilon}{\sqrt{\mu}}\|b_n\|. \end{aligned} \tag{6.67}$$

Now let $z \in Z$. Since B is β-cocoercive,

$$(\forall n \in \mathbb{N}) \quad \langle x_n - z \mid Bx_n - Bz\rangle \geqslant \beta\|Bx_n - Bz\|^2. \tag{6.68}$$

On the other hand, it follows from (6.60) that

$$(\forall n \in \mathbb{N}) \quad \|B_n x_n - B_n z\|^2_{U_n^{-1}} \leqslant \gamma_n^2 \|U_n\| \, \|Bx_n - Bz\|^2 \leqslant \gamma_n^2 \mu \|Bx_n - Bz\|^2. \tag{6.69}$$

We also note that, since $-Bz \in Az$, (6.63) yields

$$(\forall n \in \mathbb{N}) \quad z = J_{A_n}(z - B_n z). \tag{6.70}$$

Altogether, it follows from (6.63), (6.70), Lemma 6.15(ii), (6.68), and (6.69) that

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \|q_n - z\|^2_{U_n^{-1}} &\leqslant \|(x_n - z) - (B_n x_n - B_n z)\|^2_{U_n^{-1}} \\ &\quad - \|(x_n - q_n) - (B_n x_n - B_n z)\|^2_{U_n^{-1}} \\ &= \|x_n - z\|^2_{U_n^{-1}} - 2\langle x_n - z \mid B_n x_n - B_n z\rangle_{U_n^{-1}} \\ &\quad + \|B_n x_n - B_n z\|^2_{U_n^{-1}} - \|(x_n - q_n) - (B_n x_n - B_n z)\|^2_{U_n^{-1}} \\ &= \|x_n - z\|^2_{U_n^{-1}} - 2\gamma_n \langle x_n - z \mid Bx_n - Bz\rangle \\ &\quad + \|B_n x_n - B_n z\|^2_{U_n^{-1}} - \|(x_n - q_n) - (B_n x_n - B_n z)\|^2_{U_n^{-1}} \\ &\leqslant \|x_n - z\|^2_{U_n^{-1}} - \gamma_n(2\beta - \mu\gamma_n)\|Bx_n - Bz\|^2 \\ &\quad - \|(x_n - q_n) - (B_n x_n - B_n z)\|^2_{U_n^{-1}} \\ &\leqslant \|x_n - z\|^2_{U_n^{-1}} - \varepsilon^2 \|Bx_n - Bz\|^2 \\ &\quad - \|(x_n - q_n) - (B_n x_n - B_n z)\|^2_{U_n^{-1}}. \end{aligned} \tag{6.71}$$

In turn, we derive from (6.66) and (6.63) that

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad (1+\eta_n)^{-1}\|s_n - z\|^2_{U_{n+1}^{-1}} &\leqslant \|s_n - z\|^2_{U_n^{-1}} \\ &\leqslant (1-\lambda_n)\|x_n - z\|^2_{U_n^{-1}} + \lambda_n \|q_n - z\|^2_{U_n^{-1}} \\ &\leqslant \|x_n - z\|^2_{U_n^{-1}} - \varepsilon^3 \|Bx_n - Bz\|^2 \\ &\quad - \varepsilon \|(x_n - q_n) - (B_n x_n - B_n z)\|^2_{U_n^{-1}}, \end{aligned} \tag{6.72}$$

which implies that

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \|s_n - z\|^2_{U_{n+1}^{-1}} &\leqslant (1+\eta_n)\|x_n - z\|^2_{U_n^{-1}} - \varepsilon^3 \|Bx_n - Bz\|^2 \\ &\quad - \varepsilon \|(x_n - q_n) - (B_n x_n - B_n z)\|^2_{U_n^{-1}} \qquad (6.73) \\ &\leqslant \delta^2 \|x_n - z\|^2_{U_n^{-1}}, \qquad (6.74) \end{aligned}$$

where

$$\delta = \sup_{n \in \mathbb{N}} \sqrt{1+\eta_n}. \tag{6.75}$$

Next, we set

$$(\forall n \in \mathbb{N}) \quad \varepsilon_n = \delta\left(\frac{1}{\sqrt{\alpha}}\|a_n\| + \frac{2\beta - \varepsilon}{\sqrt{\mu}}\|b_n\|\right). \tag{6.76}$$

Then our assumptions yield

$$\sum_{n\in\mathbb{N}} \varepsilon_n < +\infty. \tag{6.77}$$

Moreover, using (6.66), (6.73), and (6.67), we obtain

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \|x_{n+1} - z\|_{U_{n+1}^{-1}} &\leqslant \|x_{n+1} - s_n\|_{U_{n+1}^{-1}} + \|s_n - z\|_{U_{n+1}^{-1}} \\ &\leqslant \sqrt{1+\eta_n}\|x_{n+1} - s_n\|_{U_n^{-1}} + \sqrt{1+\eta_n}\|x_n - z\|_{U_n^{-1}} \\ &\leqslant \delta\|x_{n+1} - s_n\|_{U_n^{-1}} + \sqrt{1+\eta_n}\|x_n - z\|_{U_n^{-1}} \\ &\leqslant \sqrt{1+\eta_n}\|x_n - z\|_{U_n^{-1}} + \varepsilon_n \\ &\leqslant (1+\eta_n)\|x_n - z\|_{U_n^{-1}} + \varepsilon_n. \end{aligned} \tag{6.78}$$

In view of (6.65), (6.77), and (6.78), we can apply Proposition 6.11 to assert that $(\|x_n - z\|_{U_n^{-1}})_{n\in\mathbb{N}}$ converges and, therefore, that

$$\zeta = \sup_{n\in\mathbb{N}} \|x_n - z\|_{U_n^{-1}} < +\infty. \tag{6.79}$$

On the other hand, (6.66), (6.67), and (6.76) yield

$$(\forall n \in \mathbb{N}) \quad \|x_{n+1} - s_n\|^2_{U_{n+1}^{-1}} \leqslant (1+\eta_n)\|x_{n+1} - s_n\|^2_{U_n^{-1}} \leqslant \varepsilon_n^2. \tag{6.80}$$

Hence, using (6.73), (6.74), (6.75), and (6.79), we get

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \|x_{n+1} - z\|^2_{U_{n+1}^{-1}} &\leqslant \|s_n - z\|^2_{U_{n+1}^{-1}} + 2\|s_n - z\|_{U_{n+1}^{-1}}\|x_{n+1} - s_n\|_{U_{n+1}^{-1}} \\ &\quad + \|x_{n+1} - s_n\|^2_{U_{n+1}^{-1}} \\ &\leqslant (1+\eta_n)\|x_n - z\|^2_{U_n^{-1}} - \varepsilon^3\|Bx_n - Bz\|^2 \\ &\quad - \varepsilon\|x_n - q_n - B_nx_n + B_nz\|^2_{U_n^{-1}} + 2\delta\zeta\varepsilon_n + \varepsilon_n^2 \\ &\leqslant \|x_n - z\|^2_{U_n^{-1}} - \varepsilon^3\|Bx_n - Bz\|^2 \\ &\quad - \varepsilon\|x_n - q_n - B_nx_n + B_nz\|^2_{U_n^{-1}} + \zeta^2\eta_n + 2\delta\zeta\varepsilon_n + \varepsilon_n^2. \end{aligned} \tag{6.81}$$

Consequently, for every $N \in \mathbb{N}$,

$$\begin{aligned}\varepsilon^3 \sum_{n=0}^{N} \|Bx_n - Bz\|^2 &\leqslant \|x_0 - z\|^2_{U_0^{-1}} - \|x_{N+1} - z\|^2_{U_{N+1}^{-1}} + \sum_{n=0}^{N}\big(\zeta^2\eta_n + 2\delta\zeta\varepsilon_n + \varepsilon_n^2\big) \\ &\leqslant \zeta^2 + \sum_{n=0}^{N}\big(\zeta^2\eta_n + 2\delta\zeta\varepsilon_n + \varepsilon_n^2\big). \end{aligned} \tag{6.82}$$

Appealing to (6.77) and the summability of $(\eta_n)_{n\in\mathbb{N}}$, taking the limit as $N \to +\infty$, yields

$$\sum_{n\in\mathbb{N}} \|Bx_n - Bz\|^2 \leqslant \frac{1}{\varepsilon^3}\bigg(\zeta^2 + \sum_{n\in\mathbb{N}} \big(\zeta^2\eta_n + 2\delta\zeta\varepsilon_n + \varepsilon_n^2\big)\bigg) < +\infty. \tag{6.83}$$

We likewise derive from (6.81) that

$$\sum_{n\in\mathbb{N}} \|x_n - q_n - B_n x_n + B_n z\|^2_{U_n^{-1}} < +\infty. \tag{6.84}$$

(i) : Let x be a weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$, say $x_{k_n} \rightharpoonup x$ as $n \to +\infty$. In view of (6.78), (6.65), and Proposition 6.12, it is enough to show that $x \in Z$. On the one hand, (6.83) yields $Bx_{k_n} \to Bz$ as $n \to +\infty$. On the other hand, since B is cocoercive, it is maximally monotone [4, Example 20.28] and its graph is therefore sequentially closed in $\mathcal{H}^{\text{weak}} \times \mathcal{H}^{\text{strong}}$ [4, Proposition 20.33(ii)]. This implies that $Bx = Bz$ and hence that $Bx_{k_n} \to Bx$ as $n \to +\infty$. Thus, in view of (6.83),

$$\sum_{n\in\mathbb{N}} \|Bx_n - Bx\|^2 < +\infty. \tag{6.85}$$

Now set

$$(\forall n \in \mathbb{N}) \quad u_n = \frac{1}{\gamma_n} U_n^{-1}(x_n - q_n) - Bx_n. \tag{6.86}$$

Then it follows from (6.63) that

$$(\forall n \in \mathbb{N}) \quad u_n \in Aq_n. \tag{6.87}$$

In addition, (6.63), (6.65), and (6.84) yield

$$\begin{aligned}
\|u_n + Bx\| &= \frac{1}{\gamma_n}\|U_n^{-1}(x_n - q_n - B_n x_n + B_n x)\| \\
&\leqslant \frac{1}{\varepsilon\alpha}\|x_n - q_n - B_n x_n + B_n x\| \\
&\leqslant \frac{\sqrt{\mu}}{\varepsilon\alpha}\|x_n - q_n - B_n x_n + B_n x\|_{U_n^{-1}} \\
&\to 0 \quad \text{as} \quad n \to +\infty.
\end{aligned} \tag{6.88}$$

Moreover, it follows from (6.63), (6.60), and (6.85) that

$$\begin{aligned}
\|x_n - q_n\| &\leqslant \|x_n - q_n - B_n x_n + B_n x\| + \|B_n x_n - B_n x\| \\
&\leqslant \|x_n - q_n - B_n x_n + B_n x\| + \gamma_n\|U_n\|\,\|Bx_n - Bx\| \\
&\leqslant \|x_n - q_n - B_n x_n + B_n x\| + (2\beta - \varepsilon)\|Bx_n - Bx\| \\
&\to 0 \quad \text{as} \quad n \to +\infty.
\end{aligned} \tag{6.89}$$

and, therefore, since $x_{k_n} \rightharpoonup x$ as $n \to +\infty$, that $q_{k_n} \rightharpoonup x$ as $n \to +\infty$. To sum up,

$$\begin{cases} q_{k_n} \rightharpoonup x \quad \text{and } u_{k_n} \to -Bx \ \text{ as } n \to +\infty, \\ (\forall n \in \mathbb{N})\ (q_{k_n}, u_{k_n}) \in \operatorname{gra} A. \end{cases} \tag{6.90}$$

Hence, using the sequential closedness of gra$A$ in $\mathcal{H}^{\text{weak}} \times \mathcal{H}^{\text{strong}}$ [4, Proposition 20.33(ii)], we conclude that $-Bx \in Ax$, i.e., $x \in Z$.

(ii) : Since $\overline{x} \in Z$, the claim follows from (6.83).

(iii) : We now prove strong convergence.

(iii)(a) : Since A and B are maximally monotone and dom $B = \mathcal{H}$, $A + B$ is maximally monotone [4, Corollary 24.4(i)] and Z is therefore closed [4, Proposition 23.39]. Hence, the claim follows from (i), (6.78), and Proposition 6.13.

(iii)(b) : It follows from (i) and (6.89) that $q_n \rightharpoonup \overline{x} \in Z$ as $n \to +\infty$ and from (6.88) that $u_n \to -B\overline{x} \in A\overline{x}$ as $n \to +\infty$. Hence, if A is demiregular at $\overline{x}$, (6.87) yields $q_n \to \overline{x}$ as $n \to +\infty$. In view of (6.89), we conclude that $x_n \to \overline{x}$ as $n \to +\infty$. Now suppose that B is demiregular at $\overline{x}$. Then since $x_n \rightharpoonup \overline{x} \in Z$ as $n \to +\infty$ by (i) and $Bx_n \to B\overline{x}$ as $n \to +\infty$ by (ii), we conclude that $x_n \to \overline{x}$ as $n \to +\infty$.

(iii)(c) : Suppose that $z \in \operatorname{int} Z$ and fix $\rho \in \left]0, +\infty\right[$ such that $B(z;\rho) \subset Z$. It follows from (6.79) that $\theta = \sup_{x \in B(z;\rho)} \sup_{n\in\mathbb{N}} \|x_n - x\|_{U_n^{-1}} \leqslant (1/\sqrt{\alpha})(\sup_{n\in\mathbb{N}} \|x_n - z\| + \sup_{x\in B(z;\rho)} \|x - z\|) < +\infty$ and from (6.81) that

$$(\forall n \in \mathbb{N})(\forall x \in B(z;\rho)) \quad \|x_{n+1} - x\|^2_{U_{n+1}^{-1}} \leqslant \|x_n - x\|^2_{U_n^{-1}} + \theta^2\eta_n + 2\delta\theta\varepsilon_n + \varepsilon_n^2. \tag{6.91}$$

Hence, the claim follows from (i), Lemma 6.9, and Proposition 6.14. □

Remark 6.25 Here are some observations on Theorem 6.24.

(i) Suppose that $(\forall n \in \mathbb{N})\ U_n = \operatorname{Id}$. Then (6.62) relapses to the forward-backward algorithm studied in [1, 12], which itself captures those of [27, 29, 40]. Theorem 6.24 extends the convergence results of these papers.

(ii) As shown in [18, Remark 5.12], the convergence of the forward-backward iterates to a solution may be only weak and not strong, hence the necessity of the additional conditions in Theorem 6.24(iii).

(iii) In Euclidean spaces, condition (6.60) was used in [32] in a variable metric proximal point algorithm and then in [28] in a more general splitting algorithm.

Next, we describe direct applications of Theorem 6.24, which yield new variable metric splitting schemes. We start with minimization problems, an area in which the forward-backward algorithm has found numerous applications, e.g., [15, 18, 21, 39, 40].

Example 6.26 Let $f \in \Gamma_0(\mathcal{H})$, let $\alpha \in]0,+\infty[$, let $\beta \in]0,+\infty[$, let $g\colon \mathcal{H} \to \mathbb{R}$ be convex and differentiable with a $1/\beta$-Lipschitzian gradient, let $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$, and let $(U_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{H})$ such that (6.60) holds. Furthermore, let $\varepsilon \in]0, \min\{1, 2\beta/(\mu+1)\}[$ where μ is given by (6.60), let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, 1]$, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, (2\beta-\varepsilon)/\mu]$, let $x_0 \in \mathcal{H}$, and let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{H}$. Suppose that $\operatorname{Argmin}(f+g) \neq \varnothing$ and set

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} y_n = x_n - \gamma_n U_n(\nabla g(x_n) + b_n) \\ x_{n+1} = x_n + \lambda_n\Big(\operatorname{prox}_{\gamma_n f}^{U_n^{-1}} y_n + a_n - x_n\Big). \end{array} \right. \tag{6.92}$$

Then the following hold for some $\overline{x} \in \operatorname{Argmin}(f+g)$.

(i) $x_n \rightharpoonup \overline{x}$ as $n \to +\infty$.

(ii) $\sum_{n\in\mathbb{N}} \|\nabla g(x_n) - \nabla g(\overline{x})\|^2 < +\infty$.

(iii) Suppose that one of the following holds.

(a) $\underline{\lim}\, d_{\operatorname{Argmin}(f+g)}(x_n) = 0$.

(b) At every point in $\operatorname{Argmin}(f+g)$, f or g is uniformly convex (see Lemma 6.23(vi)).

(c) $\operatorname{int}\operatorname{Argmin}(f+g) \neq \varnothing$ and there exists $(\nu_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$ such that $(\forall n \in \mathbb{N})$ $(1+\nu_n)U_n \succcurlyeq U_{n+1}$.

Then $x_n \to \overline{x}$ as $n \to +\infty$.

Proof. An application of Theorem 6.24 with $A = \partial f$ and $B = \nabla g$, since the Baillon-Haddad theorem [4, Corollary 18.16] ensures that ∇g is β-cocoercive and since, by [4, Corollary 26.3], $\operatorname{Argmin}(f+g) = \operatorname{zer}(A+B)$. □

The next example addresses variational inequalities, another area of application of forward-backward splitting [4, 23, 39, 40].

Example 6.27 Let $f \in \Gamma_0(\mathcal{H})$, let $\alpha \in]0,+\infty[$, let $\beta \in]0,+\infty[$, let $B\colon \mathcal{H} \to \mathcal{H}$ be β-cocoercive, let $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$, and let $(U_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{H})$ that satisfies (6.60). Furthermore, let $\varepsilon \in]0, \min\{1, 2\beta/(\mu+1)\}[$ where μ is given by (6.60), let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, 1]$, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, (2\beta-\varepsilon)/\mu]$, let $x_0 \in \mathcal{H}$, and let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{H}$. Suppose that the variational inequality

$$\text{find} \quad x \in \mathcal{H} \quad \text{such that} \quad (\forall y \in \mathcal{H}) \quad \langle x - y \mid Bx \rangle + f(x) \leqslant f(y) \tag{6.93}$$

admits at least one solution and set

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} y_n = x_n - \gamma_n U_n(Bx_n + b_n) \\ x_{n+1} = x_n + \lambda_n\big(\operatorname{prox}_{\gamma_n f}^{U_n^{-1}} y_n + a_n - x_n\big). \end{array} \right. \tag{6.94}$$

Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a solution $\overline{x}$ to (6.93).

Proof. Set $A = \partial f$ in Theorem 6.24(i). □

6.2.5 Strongly monotone inclusions in duality

In [13], strongly convex composite minimization problems of the form

$$\underset{x\in\mathcal{H}}{\text{minimize}}\;\; f(x)+g(Lx-r)+\frac{1}{2}\|x-z\|^2, \tag{6.95}$$

where $z\in\mathcal{H}$, $r\in\mathcal{G}$, $f\in\Gamma_0(\mathcal{H})$, $g\in\Gamma_0(\mathcal{G})$, and $L\in\mathcal{B}(\mathcal{H},\mathcal{G})$, were solved by applying the forward-backward algorithm to the Fenchel-Rockafellar dual problem

$$\underset{v\in\mathcal{G}}{\text{minimize}}\;\; \widetilde{f^*}(z-L^*v)+g^*(v)+\langle v\mid r\rangle, \tag{6.96}$$

where $\widetilde{f^*}=f^*\,\square\,(\|\cdot\|^2/2)$ denotes the Moreau envelope of f^*. This framework was shown to capture and extend various formulations in areas such as sparse signal recovery, best approximation theory, and inverse problems. In this section, we use the results of Section 6.2.4 to generalize this framework in several directions simultaneously. First, we consider general monotone inclusions, not just minimization problems. Second, we incorporate parallel sum components (see (6.33)) in the model. Third, our algorithm allows for a variable metric. The following problem is formulated using the duality framework of [16], which itself extends those of [2, 22, 31, 34, 37, 38].

Problem 6.28 Let $z\in\mathcal{H}$, let $\rho\in\left]0,+\infty\right[$, let $A\colon\mathcal{H}\to 2^{\mathcal{H}}$ be maximally monotone, and let m be a strictly positive integer. For every $i\in\{1,\ldots,m\}$, let $r_i\in\mathcal{G}_i$, let $B_i\colon\mathcal{G}_i\to 2^{\mathcal{G}_i}$ be maximally monotone, let $\nu_i\in\left]0,+\infty\right[$, let $D_i\colon\mathcal{G}_i\to 2^{\mathcal{G}_i}$ be maximally monotone and ν_i-strongly monotone, and suppose that $0\neq L_i\in\mathcal{B}(\mathcal{H},\mathcal{G}_i)$. Furthermore, suppose that

$$z\in\operatorname{ran}\Bigg(A+\sum_{i=1}^m L_i^*\big((B_i\,\square\,D_i)(L_i\cdot-r_i)\big)+\rho\,\mathrm{Id}\Bigg). \tag{6.97}$$

The problem is to solve the primal inclusion

$$\text{find}\;\; \overline{x}\in\mathcal{H}\;\;\text{such that}\;\; z\in A\overline{x}+\sum_{i=1}^m L_i^*\big((B_i\,\square\,D_i)(L_i\overline{x}-r_i)\big)+\rho\overline{x}, \tag{6.98}$$

together with the dual inclusion

$$\begin{aligned}&\text{find}\;\; \overline{v_1}\in\mathcal{G}_1,\;\ldots,\;\overline{v_m}\in\mathcal{G}_m\;\;\text{such that}\\ &\quad(\forall i\in\{1,\ldots,m\})\quad r_i\in L_i\Bigg(J_{\rho^{-1}A}\Bigg(\rho^{-1}\Bigg(z-\sum_{j=1}^m L_j^*\overline{v_j}\Bigg)\Bigg)\Bigg)-B_i^{-1}\overline{v_i}-D_i^{-1}\overline{v_i}.\end{aligned} \tag{6.99}$$

Let us start with some properties of Problem 6.28.

Proposition 6.29 *In Problem 6.28, set*

$$\overline{x} = J_{\rho^{-1}M}\big(\rho^{-1}z\big), \quad \text{where} \quad M = A + \sum_{i=1}^{m} L_i^* \circ (B_i \,\square\, D_i) \circ (L_i \cdot - r_i). \tag{6.100}$$

Then the following hold.

(i) $\overline{x}$ *is the unique solution to the primal problem* (6.98).

(ii) *The dual problem* (6.99) *admits at least one solution.*

(iii) *Let* $(\overline{v_1}, \dots, \overline{v_m})$ *be a solution to* (6.99). *Then* $\overline{x} = J_{\rho^{-1}A}\big(\rho^{-1}\big(z - \sum_{i=1}^m L_i^*\overline{v_i}\big)\big)$.

(iv) *Condition* (6.97) *is satisfied for every* z *in* $\mathcal{H}$ *if and only if* M *is maximally monotone. This is true when one of the following holds.*

(a) *The conical hull of*

$$E = \Big\{ \big(L_i x - r_i - v_i\big)_{1 \leqslant i \leqslant m} \;\Big|\; x \in \operatorname{dom} A \text{ and } (v_i)_{1 \leqslant i \leqslant m} \in \bigtimes_{i=1}^{m} \operatorname{ran}\big(B_i^{-1} + D_i^{-1}\big) \Big\} \tag{6.101}$$

is a closed vector subspace.

(b) $A = \partial f$ *for some* $f \in \Gamma_0(\mathcal{H})$, *for every* $i \in \{1, \dots, m\}$, $B_i = \partial g_i$ *for some* $g_i \in \Gamma_0(\mathcal{G}_i)$ *and* $D_i = \partial \ell_i$ *for some strongly convex function* $\ell_i \in \Gamma_0(\mathcal{G}_i)$, *and one of the following holds.*

1/ $(r_1, \dots, r_m) \in \operatorname{sri}\big\{(L_i x - y_i)_{1 \leqslant i \leqslant m} \mid x \in \operatorname{dom} f$ *and* $(\forall i \in \{1, \dots, m\}) \;\; y_i \in \operatorname{dom} g_i + \operatorname{dom} \ell_i\big\}$.

2/ *For every* $i \in \{1, \dots, m\}$, g_i *or* ℓ_i *is real-valued.*

3/ $\mathcal{H}$ *and* $(\mathcal{G}_i)_{1 \leqslant i \leqslant m}$ *are finite-dimensional, and there exists* $x \in \operatorname{ri} \operatorname{dom} f$ *such that*

$$(\forall i \in \{1, \dots, m\}) \quad L_i x - r_i \in \operatorname{ri} \operatorname{dom} g_i + \operatorname{ri} \operatorname{dom} \ell_i. \tag{6.102}$$

Proof. (i) : It follows from our assumptions and [4, Proposition 20.10] that $\rho^{-1}M$ is a monotone operator. Hence, $J_{\rho^{-1}M}$ is a single-valued operator with domain $\operatorname{ran}(\mathrm{Id} + \rho^{-1}M)$ [4, Proposition 23.9(ii)]. Moreover, (6.97) $\Leftrightarrow$ $\rho^{-1}z \in \operatorname{ran}(\mathrm{Id} + \rho^{-1}M) = \operatorname{dom} J_{\rho^{-1}M}$, and, in view of (6.31), the inclusion in (6.98) is equivalent to $\overline{x} = J_{\rho^{-1}M}(\rho^{-1}z)$.

(ii)&(iii) : It follows from (6.31) and (6.33) that

$$\begin{aligned}
(i) \;\Leftrightarrow\; & (\exists\, \overline{v_1} \in \mathcal{G}_1) \cdots (\exists\, \overline{v_m} \in \mathcal{G}_m) \quad \begin{cases} (\forall i \in \{1, \dots, m\}) \quad \overline{v_i} \in (B_i \,\square\, D_i)(L_i\overline{x} - r_i) \\ z - \sum_{i=1}^m L_i^*\overline{v_i} \in A\overline{x} + \rho\overline{x} \end{cases} \\
\Leftrightarrow\; & (\exists\, \overline{v_1} \in \mathcal{G}_1) \cdots (\exists\, \overline{v_m} \in \mathcal{G}_m) \quad \begin{cases} (\forall i \in \{1, \dots, m\}) \quad r_i \in L_i\overline{x} - B_i^{-1}\overline{v_i} - D_i^{-1}\overline{v_i} \\ \overline{x} = J_{\rho^{-1}A}\big(\rho^{-1}\big(z - \sum_{j=1}^m L_j^*\overline{v_j}\big)\big) \end{cases} \\
\Leftrightarrow\; & \begin{cases} (\overline{v_1}, \dots, \overline{v_m}) \text{ solves } (6.99) \\ \overline{x} = J_{\rho^{-1}A}\big(\rho^{-1}\big(z - \sum_{j=1}^m L_j^*\overline{v_j}\big)\big). \end{cases}
\end{aligned} \tag{6.103}$$

(iv) : It follows from Minty's theorem [4, Theorem 21.1], that $M+\rho\,\mathrm{Id}$ is surjective if and only if M is maximally monotone.

(iv)(a) : Using Notation 6.8, let us set

$$\boldsymbol{L}\colon \mathcal{H}\to\boldsymbol{\mathcal{G}}\colon x\mapsto \big(L_ix\big)_{1\leqslant i\leqslant m}\quad\text{and}\quad \boldsymbol{B}\colon\boldsymbol{\mathcal{G}}\to 2^{\boldsymbol{\mathcal{G}}}\colon \boldsymbol{y}\mapsto \big((B_i\,\Box\,D_i)(y_i-r_i)\big)_{1\leqslant i\leqslant m}. \tag{6.104}$$

Then it follows from (6.100) that $M = A+\boldsymbol{L}^*\circ\boldsymbol{B}\circ\boldsymbol{L}$ and from (6.101) that $E = \boldsymbol{L}(\operatorname{dom}A)-\operatorname{dom}\boldsymbol{B}$. Hence, since $\operatorname{cone}(E) = \overline{\operatorname{span}}\,(E)$, in view of [6, Section 24], to conclude that M is maximally monotone, it is enough to show that $\boldsymbol{B}$ is. For every $i\in\{1,\ldots,m\}$, since D_i is maximally monotone and strongly monotone, $\operatorname{dom}D_i^{-1} = \operatorname{ran}D_i = \mathcal{G}_i$ [4, Proposition 22.8(ii)] and it follows from [4, Proposition 20.22 and Corollary 24.4(i)] that $B_i\,\Box\,D_i$ is maximally monotone. This shows that $\boldsymbol{B}$ is maximally monotone.

(iv)(b) : This follows from [16, Proposition 4.3]. □

Remark 6.30 In connection with Proposition 6.29(iv), let us note that even in the simple setting of normal cone operators in finite dimension, some constraint qualification is required to ensure the existence of a primal solution for every $z\in\mathcal{H}$. To see this, suppose that, in Problem 6.28, $\mathcal{H}$ is the Euclidean plane, $m=1$, $\rho=1$, $\mathcal{G}_1=\mathcal{H}$, $L_1=\mathrm{Id}$, $z=(\zeta_1,\zeta_2)$, $r_1=0$, $D_1=\{0\}^{-1}$, $A=N_C$, and $B_1=N_K$, where $C=\big\{(\xi_1,\xi_2)\in\mathbb{R}^2\;\big|\;(\xi_1-1)^2+\xi_2^2\leqslant 1\big\}$ and $K=\big\{(\xi_1,\xi_2)\in\mathbb{R}^2\;\big|\;\xi_1\leqslant 0\big\}$. Then $\operatorname{dom}(A+B_1+\mathrm{Id}) = \operatorname{dom}A\cap\operatorname{dom}B_1 = C\cap K = \{0\}$ and the primal inclusion $z\in A\overline{x}+B_1\overline{x}+\overline{x}$ reduces to $(\zeta_1,\zeta_2)\in N_C0+N_K0 = \left]-\infty,0\right]\times\{0\}+\left[0,+\infty\right[\times\{0\} = \mathbb{R}\times\{0\}$, which has no solution if $\zeta_2\neq 0$. Here $\operatorname{cone}(\operatorname{dom}A-\operatorname{dom}B_1) = \operatorname{cone}(C-K) = -K$ is not a vector subspace.

In the following result we derive from Theorem 6.24 a parallel primal-dual algorithm for solving Problem 6.28.

Corollary 6.31 *In Problem 6.28, set*

$$\beta = \frac{1}{\displaystyle\max_{1\leqslant i\leqslant m}\frac{1}{\nu_i}+\frac{1}{\rho}\sum_{1\leqslant i\leqslant m}\|L_i\|^2}. \tag{6.105}$$

Let $(a_n)_{n\in\mathbb{N}}$ be an absolutely summable sequence in $\mathcal{H}$, let $\alpha\in\left]0,+\infty\right[$, and let $(\eta_n)_{n\in\mathbb{N}}\in\ell_+^1(\mathbb{N})$. For every $i\in\{1,\ldots,m\}$, let $v_{i,0}\in\mathcal{G}_i$, let $(b_{i,n})_{n\in\mathbb{N}}$ and $(d_{i,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{G}_i$, and let $(U_{i,n})_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{G}_i)$. Suppose that

$$\mu = \max_{1\leqslant i\leqslant m}\sup_{n\in\mathbb{N}}\|U_{i,n}\|<+\infty\quad\textit{and}\quad(\forall i\in\{1,\ldots,m\})(\forall n\in\mathbb{N})\quad(1+\eta_n)U_{i,n+1}\succcurlyeq U_{i,n}. \tag{6.106}$$

Let $\varepsilon \in \left]0, \min\{1, 2\beta/(\mu+1)\}\right[$*, let* $(\lambda_n)_{n\in\mathbb{N}}$ *be a sequence in* $[\varepsilon, 1]$*, and let* $(\gamma_n)_{n\in\mathbb{N}}$ *be a sequence in* $[\varepsilon, (2\beta-\varepsilon)/\mu]$*. Set*

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} s_n = z - \sum_{i=1}^m L_i^* v_{i,n} \\ x_n = J_{\rho^{-1}A}(\rho^{-1}s_n) + a_n \\ \text{For } i = 1, \ldots, m \\ \left\lfloor \begin{array}{l} w_{i,n} = v_{i,n} + \gamma_n U_{i,n}\big(L_i x_n - r_i - D_i^{-1} v_{i,n} - d_{i,n}\big) \\ v_{i,n+1} = v_{i,n} + \lambda_n\Big(J_{\gamma_n U_{i,n} B_i^{-1}}(w_{i,n}) + b_{i,n} - v_{i,n}\Big). \end{array} \right. \end{array} \right. \tag{6.107}$$

Then the following hold for the solution $\overline{x}$ *to* (6.98) *and for some solution* $(\overline{v_1}, \ldots, \overline{v_m})$ *to* (6.99)*.*

(i) $(\forall i \in \{1, \ldots, m\})$ $v_{i,n} \rightharpoonup \overline{v_i}$ *as* $n \to +\infty$*. In addition,* $\overline{x} = J_{\rho^{-1}A}\big(\rho^{-1}\big(z - \sum_{i=1}^m L_i^* \overline{v_i}\big)\big)$*.*

(ii) $x_n \to \overline{x}$ *as* $n \to +\infty$*.*

Proof. For every $i \in \{1, \ldots, m\}$, since D_i is maximally monotone and ν_i-strongly monotone, D_i^{-1} is ν_i-cocoercive with $\operatorname{dom} D_i^{-1} = \operatorname{ran} D_i = \mathcal{G}_i$ [4, Proposition 22.8(ii)]. Let us define $\boldsymbol{\mathcal{G}}$ as in Notation 6.8, and let us introduce the operators

$$\begin{cases} T\colon \mathcal{H} \to \mathcal{H}\colon x \mapsto J_{\rho^{-1}A}\big(\rho^{-1}(z - x)\big) \\ \boldsymbol{A}\colon \boldsymbol{\mathcal{G}} \to 2^{\boldsymbol{\mathcal{G}}}\colon \boldsymbol{v} \mapsto \big(B_i^{-1} v_i\big)_{1\leqslant i\leqslant m} \\ \boldsymbol{D}\colon \boldsymbol{\mathcal{G}} \to \boldsymbol{\mathcal{G}}\colon \boldsymbol{v} \mapsto \big(r_i + D_i^{-1} v_i\big)_{1\leqslant i\leqslant m} \\ \boldsymbol{L}\colon \mathcal{H} \to \boldsymbol{\mathcal{G}}\colon x \mapsto \big(L_i x\big)_{1\leqslant i\leqslant m} \end{cases} \tag{6.108}$$

and

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{U}_n\colon \boldsymbol{\mathcal{G}} \to \boldsymbol{\mathcal{G}}\colon \boldsymbol{v} \mapsto \big(U_{i,n} v_i\big)_{1\leqslant i\leqslant m}. \tag{6.109}$$

(i) : In view of (6.30) and (6.108),

$$\boldsymbol{A} \text{ is maximally monotone,} \tag{6.110}$$

$\boldsymbol{D}$ is $(\min_{1\leqslant i\leqslant m} \nu_i)$-cocoercive, Lemma 6.15(ii) implies that

$$-T \text{ is } \rho\text{-cocoercive,} \tag{6.111}$$

while $\|\boldsymbol{L}\|^2 \leqslant \sum_{i=1}^m \|L_i\|^2$. Hence, we derive from (6.105) and Proposition 6.10 that

$$\boldsymbol{B} = \boldsymbol{D} - \boldsymbol{L}T\boldsymbol{L}^* \text{ is } \beta\text{-cocoercive.} \tag{6.112}$$

Moreover, it follows from (6.106), (6.109), and (6.30) that

$$\sup_{n\in\mathbb{N}} \|\boldsymbol{U}_n\| = \mu \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad (1+\eta_n)\boldsymbol{U}_{n+1} \succcurlyeq \boldsymbol{U}_n \in \mathcal{P}_\alpha(\boldsymbol{\mathcal{G}}). \tag{6.113}$$

Now set

$$(\forall n \in \mathbb{N}) \quad \begin{cases} \boldsymbol{a}_n = \big(b_{i,n}\big)_{1 \leqslant i \leqslant m} \\ \boldsymbol{b}_n = \big(d_{i,n} - L_i a_n\big)_{1 \leqslant i \leqslant m} \\ \boldsymbol{v}_n = \big(v_{i,n}\big)_{1 \leqslant i \leqslant m} \\ \boldsymbol{w}_n = \big(w_{i,n}\big)_{1 \leqslant i \leqslant m}. \end{cases} \tag{6.114}$$

Then $\sum_{n\in\mathbb{N}} |||\boldsymbol{a}_n||| < +\infty$, $\sum_{n\in\mathbb{N}} |||\boldsymbol{b}_n||| < +\infty$, and (6.107) can be rewritten as

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} \boldsymbol{w}_n = \boldsymbol{v}_n - \gamma_n \boldsymbol{U}_n(\boldsymbol{B}\boldsymbol{v}_n + \boldsymbol{b}_n) \\ \boldsymbol{v}_{n+1} = \boldsymbol{v}_n + \lambda_n\big(J_{\gamma_n \boldsymbol{U}_n \boldsymbol{A}}(\boldsymbol{w}_n) + \boldsymbol{a}_n - \boldsymbol{v}_n\big). \end{array} \right. \tag{6.115}$$

Furthermore, the dual problem (6.99) is equivalent to

$$\text{find} \quad \overline{\boldsymbol{v}} \in \boldsymbol{\mathcal{G}} \quad \text{such that} \quad \boldsymbol{0} \in \boldsymbol{A}\overline{\boldsymbol{v}} + \boldsymbol{B}\overline{\boldsymbol{v}} \tag{6.116}$$

which, in view of (6.110), (6.112), and Proposition 6.29(ii), can be solved using (6.115). Altogether, the claims follow from Theorem 6.24(i) and Proposition 6.29(iii).

(ii) : Set $(\forall n \in \mathbb{N})\ z_n = x_n - a_n$. It follows from (i), (6.107) and (6.108) that

$$\overline{x} = T(\boldsymbol{L}^*\overline{\boldsymbol{v}}) \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad z_n = T(\boldsymbol{L}^*\boldsymbol{v}_n). \tag{6.117}$$

In turn, we deduce from (6.111), (i), (6.112), and the monotonicity of $\boldsymbol{D}$ that

$$\begin{aligned} \rho\|z_n - \overline{x}\|^2 &= \rho\|T(\boldsymbol{L}^*\boldsymbol{v}_n) - T(\boldsymbol{L}^*\overline{\boldsymbol{v}})\|^2 \\ &\leqslant \langle \boldsymbol{L}^*(\boldsymbol{v}_n - \overline{\boldsymbol{v}}) \mid T(\boldsymbol{L}^*\overline{\boldsymbol{v}}) - T(\boldsymbol{L}^*\boldsymbol{v}_n)\rangle \\ &\leqslant \langle\langle \boldsymbol{v}_n - \overline{\boldsymbol{v}} \mid \boldsymbol{L}T(\boldsymbol{L}^*\overline{\boldsymbol{v}}) - \boldsymbol{L}T(\boldsymbol{L}^*\boldsymbol{v}_n)\rangle\rangle \\ &\leqslant \langle\langle \boldsymbol{v}_n - \overline{\boldsymbol{v}} \mid \boldsymbol{D}\boldsymbol{v}_n - \boldsymbol{D}\overline{\boldsymbol{v}}\rangle\rangle - \langle\langle \boldsymbol{v}_n - \overline{\boldsymbol{v}} \mid \boldsymbol{L}T(\boldsymbol{L}^*\boldsymbol{v}_n) - \boldsymbol{L}T(\boldsymbol{L}^*\overline{\boldsymbol{v}})\rangle\rangle \\ &= \langle\langle \boldsymbol{v}_n - \overline{\boldsymbol{v}} \mid \boldsymbol{B}\boldsymbol{v}_n - \boldsymbol{B}\overline{\boldsymbol{v}}\rangle\rangle \\ &\leqslant \delta|||\boldsymbol{B}\boldsymbol{v}_n - \boldsymbol{B}\overline{\boldsymbol{v}}|||, \end{aligned} \tag{6.118}$$

where $\delta = \sup_{n\in\mathbb{N}} |||\boldsymbol{v}_n - \overline{\boldsymbol{v}}||| < +\infty$ by (i). Therefore, it follows from (6.115) and Theorem 6.24(ii) that $\|z_n - \overline{x}\| \to 0$. Since $a_n \to 0$ as $n \to +\infty$, we conclude that $x_n \to \overline{x}$ as $n \to +\infty$. □

Remark 6.32 Here are some observations on Corollary 6.31.

(i) At iteration n, the vectors a_n, $b_{i,n}$, and $d_{i,n}$ model errors in the implementation of the nonlinear operators. Note also that, thanks to Example 6.17(i), the computation of $v_{i,n+1}$ in (6.107) can be implemented using $J_{\gamma_n^{-1}U_{i,n}^{-1}B_i}$ rather than $J_{\gamma_n U_{i,n} B_i^{-1}}$.

(ii) Corollary 6.31 provides a general algorithm for solving strongly monotone composite inclusions which is new even in the fixed standard metric case, i.e., $(\forall i \in \{1,\ldots,m\})(\forall n \in \mathbb{N})\ U_{i,n} = \mathrm{Id}$.

The following example describes an application of Corollary 6.31 to strongly convex minimization problems which extends the primal-dual formulation (6.95)–(6.96) of [13] and solves it with a variable metric scheme. It also extends the framework of [14], where $f = 0$ and $(\forall i \in \{1,\ldots,m\})\ \ell_i = \iota_{\{0\}}$ and $(\forall n \in \mathbb{N})\ U_{i,n} = \mathrm{Id}$.

Example 6.33 Let $z \in \mathcal{H}$, let $f \in \Gamma_0(\mathcal{H})$, let $\alpha \in \left]0,+\infty\right[$, let $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$, let $(a_n)_{n\in\mathbb{N}}$ be an absolutely summable sequence in $\mathcal{H}$, and let m be a strictly positive integer. For every $i \in \{1,\ldots,m\}$, let $r_i \in \mathcal{G}_i$, let $g_i \in \Gamma_0(\mathcal{G}_i)$, let $\nu_i \in \left]0,+\infty\right[$, let $\ell_i \in \Gamma_0(\mathcal{G}_i)$ be ν_i-strongly convex, let $v_{i,0} \in \mathcal{G}_i$, let $(b_{i,n})_{n\in\mathbb{N}}$ and $(d_{i,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{G}_i$, let $(U_{i,n})_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{G}_i)$, and suppose that $0 \neq L_i \in \mathcal{B}(\mathcal{H},\mathcal{G}_i)$. Furthermore, suppose that (see Proposition 6.29(iv)(b) for special cases)

$$z \in \mathrm{ran}\left(\partial f + \sum_{i=1}^m L_i^*(\partial g_i \,\square\, \partial \ell_i)(L_i\cdot - r_i) + \mathrm{Id}\right). \tag{6.119}$$

The primal problem is

$$\underset{x\in\mathcal{H}}{\text{minimize}}\ \ f(x) + \sum_{i=1}^m (g_i \,\square\, \ell_i)(L_i x - r_i) + \frac{1}{2}\|x-z\|^2, \tag{6.120}$$

and the dual problem is

$$\underset{v_1\in\mathcal{G}_1,\ldots,v_m\in\mathcal{G}_m}{\text{minimize}}\ \ \widetilde{f}^*\left(z - \sum_{i=1}^m L_i^* v_i\right) + \sum_{i=1}^m \big(g_i^*(v_i) + \ell_i^*(v_i) + \langle v_i \mid r_i\rangle\big). \tag{6.121}$$

Suppose that (6.106) holds, let $\varepsilon \in \left]0, \min\{1, 2\beta/(\mu+1)\}\right[$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon,1]$, and let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon,(2\beta-\varepsilon)/\mu]$, where β is defined in (6.105) and μ in (6.106). Set

$$(\forall n \in \mathbb{N})\quad \left\lfloor\begin{array}{l} s_n = z - \sum_{i=1}^m L_i^* v_{i,n} \\ x_n = \mathrm{prox}_f s_n + a_n \\ \text{For } i = 1,\ldots,m \\ \left\lfloor\begin{array}{l} w_{i,n} = v_{i,n} + \gamma_n U_{i,n}\big(L_i x_n - r_i - \nabla \ell_i^*(v_{i,n}) - d_{i,n}\big) \\ v_{i,n+1} = v_{i,n} + \lambda_n\Big(\mathrm{prox}_{\gamma_n g_i^*}^{U_{i,n}^{-1}} w_{i,n} + b_{i,n} - v_{i,n}\Big). \end{array}\right. \end{array}\right. \tag{6.122}$$

Then (6.120) admits a unique solution $\overline{x}$ and the following hold for some solution $(\overline{v_1},\ldots,\overline{v_m})$ to (6.121).

(i) $(\forall i \in \{1,\ldots,m\})\ v_{i,n} \rightharpoonup \overline{v_i}$ as $n \to +\infty$. In addition, $\overline{x} = \mathrm{prox}_f(z - \sum_{i=1}^m L_i^*\overline{v_i})$.

(ii) $x_n \to \overline{x}$ as $n \to +\infty$.

Proof. Set

$$\rho = 1, \quad A = \partial f, \quad \text{and} \quad (\forall i \in \{1,\ldots,m\})\quad B_i = \partial g_i \quad \text{and} \quad D_i = \partial \ell_i. \tag{6.123}$$

It follows from [4, Theorem 20.40] that the operators A, $(B_i)_{1\leqslant i\leqslant m}$, and $(D_i)_{1\leqslant i\leqslant m}$ are maximally monotone. We also observe that (6.119) implies that (6.97) is satisfied. Moreover, for every $i \in \{1,\ldots,m\}$, D_i is ν_i-strongly monotone [4, Example 22.3(iv)], ℓ_i^* is Fréchet differentiable on $\mathcal{G}_i$ [4, Corollary 13.33 and Theorem 18.15], and $D_i^{-1} = (\partial\ell_i)^{-1} = \partial\ell_i^* = \{\nabla\ell_i^*\}$ [4, Corollary 16.24 and Proposition 17.26(i)]. Since, for every $i \in \{1,\ldots,m\}$, $\operatorname{dom}\ell_i^* = \mathcal{G}_i$, [4, Proposition 24.27] yields

$$(\forall i \in \{1,\ldots,m\}) \quad B_i \,\square\, D_i = \partial g_i \,\square\, \partial\ell_i = \partial(g_i \,\square\, \ell_i), \tag{6.124}$$

while [4, Corollaries 16.24 and 16.38(iii)] yield

$$(\forall i \in \{1,\ldots,m\}) \quad B_i^{-1} + D_i^{-1} = \partial g_i^* + \{\nabla\ell_i^*\} = \partial\big(g_i^* + \ell_i^*\big). \tag{6.125}$$

Moreover, (6.48) implies that (6.122) is a special case of (6.107). Hence, in view of Corollary 6.31, it remains to show that (6.98) and (6.99) yield (6.120) and (6.121), respectively. Let us set $q = \|\cdot\|^2/2$. We derive from [4, Example 16.33] that

$$\partial\big(f + q(\cdot - z)\big) = \partial f + \mathrm{Id} - z. \tag{6.126}$$

On the other hand, it follows from (6.119) and [4, Proposition 16.5(ii)] that

$$\partial\big(f+q(\cdot-z)\big)+\sum_{i=1}^m L_i^*\big(\partial(g_i\,\square\,\ell_i)\big)(L_i\cdot-r_i) \subset \partial\bigg(f+q(\cdot-z)+\sum_{i=1}^m(g_i\,\square\,\ell_i)\circ(L_i\cdot-r_i)\bigg) \tag{6.127}$$

and that $x \mapsto f(x) + \sum_{i=1}^m (g_i\,\square\,\ell_i)(L_ix - r_i) + \|x-z\|^2/2$ is a strongly convex function in $\Gamma_0(\mathcal{H})$. Therefore [4, Corollary 11.16] asserts that (6.120) possesses a unique solution $\overline{x}$. Next, we deduce from (6.126), (6.123), (6.124), and Fermat's rule [4, Theorem 16.2] that, for every $x \in \mathcal{H}$,

$$\begin{aligned}
x \text{ solves (6.98)} \quad &\Leftrightarrow \quad z \in \partial f(x) + \sum_{i=1}^m L_i^*\big((\partial g_i\,\square\,\partial\ell_i)(L_ix - r_i)\big) + x \\
&\Leftrightarrow \quad 0 \in \partial\big(f + q(\cdot - z)\big)(x) + \bigg(\sum_{i=1}^m L_i^* \circ \partial(g_i\,\square\,\ell_i) \circ (L_i\cdot - r_i)\bigg)(x) \\
&\Rightarrow \quad 0 \in \partial\bigg(f + q(\cdot - z) + \sum_{i=1}^m (g_i\,\square\,\ell_i) \circ (L_i\cdot - r_i)\bigg)(x) \\
&\Leftrightarrow \quad x \text{ solves (6.120)}.
\end{aligned} \tag{6.128}$$

Finally, set $\boldsymbol{L}\colon \mathcal{H} \to \boldsymbol{\mathcal{G}}\colon x \mapsto (L_ix)_{1\leqslant i\leqslant m}$ and $\boldsymbol{h}\colon \boldsymbol{\mathcal{G}} \to \left]-\infty,+\infty\right]\colon \boldsymbol{v} \mapsto \sum_{i=1}^m(g_i^*(v_i) + \ell_i^*(v_i) + \langle v_i \mid r_i\rangle)$. We recall that $\widetilde{f^*} = f^*\,\square\,q$ is Fréchet differentiable on $\mathcal{H}$ with $\nabla\widetilde{f^*} = \operatorname{prox}_f$ [4, Remark 14.4]. Hence, it follows from (6.123), (6.125), [4, Proposition 16.8 and Theorem 16.37(i)], and Fermat's rule [4, Theorem 16.2] that, for every

$\boldsymbol{v} = (v_i)_{1\leqslant i\leqslant m} \in \boldsymbol{\mathcal{G}}$,

$$\begin{aligned}
\boldsymbol{v} \text{ solves } (6.99) &\Leftrightarrow (\forall i \in \{1,\ldots,m\})\ r_i \in L_i\Big(J_A\Big(z - \sum_{j=1}^m L_j^* v_j\Big)\Big) \\
&\qquad\qquad - B_i^{-1}v_i - D_i^{-1}v_i \\
&\Leftrightarrow (\forall i \in \{1,\ldots,m\})\ r_i \in L_i\Big(\operatorname{prox}_f\Big(z - \sum_{j=1}^m L_j^* v_j\Big)\Big) - \partial(g_i^* + \ell_i^*)(v_i) \\
&\Leftrightarrow (0,\ldots,0) \in -\boldsymbol{L}\Big(\nabla \widetilde{f}^*\big(z - \boldsymbol{L}^*\boldsymbol{v}\big)\Big) + \bigtimes_{i=1}^m \partial\big(g_i^* + \ell_i^* + \langle\cdot \mid r_i\rangle\big)(v_i) \\
&\qquad = \big(-\boldsymbol{L}^*\big)^*\Big(\nabla \widetilde{f}^*\big(z - \boldsymbol{L}^*\boldsymbol{v}\big)\Big) + \partial\boldsymbol{h}(\boldsymbol{v}) \\
&\qquad = \partial\Big(\widetilde{f}^*\big(z - \boldsymbol{L}^*\cdot\,\big) + \boldsymbol{h}\Big)(\boldsymbol{v}) \\
&\Leftrightarrow \boldsymbol{v} \text{ solves } (6.121),
\end{aligned} \tag{6.129}$$

which completes the proof. □

We conclude this section with an application to a composite best approximation problem.

Example 6.34 Let $z \in \mathcal{H}$, let C be a closed convex subset of $\mathcal{H}$, let $\alpha \in\,]0,+\infty[$, let $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$, let $(a_n)_{n\in\mathbb{N}}$ be an absolutely summable sequence in $\mathcal{H}$, and let m be a strictly positive integer. For every $i \in \{1,\ldots,m\}$, let $r_i \in \mathcal{G}_i$, let D_i be a closed convex subset of $\mathcal{G}_i$, let $v_{i,0} \in \mathcal{G}_i$, let $(b_{i,n})_{n\in\mathbb{N}}$ be an absolutely summable sequence in $\mathcal{G}_i$, let $(U_{i,n})_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{G}_i)$, and suppose that $0 \neq L_i \in \mathcal{B}(\mathcal{H},\mathcal{G}_i)$. The problem is

$$\underset{\substack{x\in C\\ L_1x\in r_1+D_1\\ \vdots\\ L_mx\in r_m+D_m}}{\text{minimize}}\ \ \|x - z\|. \tag{6.130}$$

Suppose that (6.106) holds, that $(\max_{1\leqslant i\leqslant m}\sup_{n\in\mathbb{N}}\|U_{i,n}\|)\sum_{i=1}^m \|L_i\|^2 < 2$, and that

$$(r_1,\ldots,r_m) \in \operatorname{sri}\big\{(L_ix - y_i)_{1\leqslant i\leqslant m}\ \big|\ x \in C \text{ and } (\forall i \in \{1,\ldots,m\})\ y_i \in D_i\big\}. \tag{6.131}$$

Set

$$(\forall n \in \mathbb{N})\quad \left\lfloor \begin{array}{l} s_n = z - \sum_{i=1}^m L_i^* v_{i,n} \\ x_n = P_C s_n + a_n \\ \text{For } i = 1,\ldots,m \\ \left\lfloor \begin{array}{l} w_{i,n} = v_{i,n} + U_{i,n}\big(L_i x_n - r_i\big) \\ v_{i,n+1} = w_{i,n} - U_{i,n}\Big(P_{D_i}^{U_{i,n}}\big(U_{i,n}^{-1}w_{i,n}\big) + b_{i,n}\Big). \end{array}\right. \end{array}\right. \tag{6.132}$$

Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to the unique solution $\overline{x}$ to (6.130).

Proof. Set $f = \iota_C$ and $(\forall i \in \{1,\ldots,m\})$ $g_i = \iota_{D_i}$, $\ell_i = \iota_{\{0\}}$, and $(\forall n \in \mathbb{N})$ $\gamma_n = \lambda_n = 1$ and $d_{i,n} = 0$. Then (6.131) and Proposition 6.29(iv)((b))i imply that (6.119) is satisfied. Moreover, in view of Example 6.17(iii), (6.132) is a special case of (6.122). Hence, the claim follows from Example 6.33(ii). □

6.2.6 Inclusions involving cocoercive operators

We revisit a primal-dual problem investigated first in [16], and then in [41] with the scenario described below.

Problem 6.35 Let $z \in \mathcal{H}$, let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone, let $\mu \in \left]0,+\infty\right[$, let $C\colon \mathcal{H} \to \mathcal{H}$ be μ-cocoercive, and let m be a strictly positive integer. For every $i \in \{1,\ldots,m\}$, let $r_i \in \mathcal{G}_i$, let $B_i\colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ be maximally monotone, let $\nu_i \in \left]0,+\infty\right[$, let $D_i\colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ be maximally monotone and ν_i-strongly monotone, and suppose that $0 \neq L_i \in \mathcal{B}\left(\mathcal{H},\mathcal{G}_i\right)$. The problem is to solve the primal inclusion

$$\text{find} \quad \overline{x} \in \mathcal{H} \quad \text{such that} \quad z \in A\overline{x} + \sum_{i=1}^{m} L_i^*\big((B_i \,\Box\, D_i)(L_i\overline{x} - r_i)\big) + C\overline{x}, \tag{6.133}$$

together with the dual inclusion

$$\begin{aligned}&\text{find } \overline{v_1} \in \mathcal{G}_1,\ \ldots,\ \overline{v_m} \in \mathcal{G}_m \text{ such that}\\ &\qquad (\exists\, x \in \mathcal{H}) \quad \begin{cases} z - \sum_{i=1}^{m} L_i^*\overline{v}_i \in Ax + Cx \\ (\forall i \in \{1,\ldots,m\})\ \overline{v}_i \in (B_i \,\Box\, D_i)(L_i x - r_i). \end{cases}\end{aligned} \tag{6.134}$$

Corollary 6.36 *In Problem 6.35, suppose that*

$$z \in \operatorname{ran}\bigg(A + \sum_{i=1}^{m} L_i^*\big((B_i \,\Box\, D_i)(L_i \cdot - r_i)\big) + C\bigg), \tag{6.135}$$

and set

$$\beta = \min\{\mu, \nu_1, \ldots, \nu_m\}. \tag{6.136}$$

Let $\varepsilon \in \left]0, \min\{1,\beta\}\right[$, let $\alpha \in \left]0,+\infty\right[$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, 1]$, let $x_0 \in \mathcal{H}$, let $(a_n)_{n\in\mathbb{N}}$ and $(c_n)_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{H}$, and let $(U_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{H})$ such that $(\forall n \in \mathbb{N})$ $U_{n+1} \succcurlyeq U_n$. For every $i \in \{1,\ldots,m\}$, let $v_{i,0} \in \mathcal{G}_i$, and let $(b_{i,n})_{n\in\mathbb{N}}$ and $(d_{i,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{G}_i$, and let $(U_{i,n})_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{G}_i)$ such that $(\forall n \in \mathbb{N})$ $U_{i,n+1} \succcurlyeq U_{i,n}$. For every $n \in \mathbb{N}$, set

$$\delta_n = \left(\sqrt{\sum_{i=1}^{m} \|\sqrt{U_{i,n}}L_i\sqrt{U_n}\|^2}\right)^{-1} - 1, \tag{6.137}$$

and suppose that

$$\zeta_n = \frac{\delta_n}{(1+\delta_n)\max\{\|U_n\|, \|U_{1,n}\|, \ldots, \|U_{m,n}\|\}} \geqslant \frac{1}{2\beta-\varepsilon}. \tag{6.138}$$

Set

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} p_n = J_{U_n A}\Big(x_n - U_n\big(\sum_{i=1}^m L_i^* v_{i,n} + Cx_n + c_n - z\big)\Big) + a_n \\ y_n = 2p_n - x_n \\ x_{n+1} = x_n + \lambda_n(p_n - x_n) \\ \text{For } i = 1, \ldots, m \\ \left\lfloor \begin{array}{l} q_{i,n} = J_{U_{i,n} B_i^{-1}}\Big(v_{i,n} + U_{i,n}\big(L_i y_n - D_i^{-1} v_{i,n} - d_{i,n} - r_i\big)\Big) + b_{i,n} \\ v_{i,n+1} = v_{i,n} + \lambda_n(q_{i,n} - v_{i,n}). \end{array} \right. \end{array} \right. \tag{6.139}$$

Then the following hold for some solution $\overline{x}$ to (6.133) *and some solution $(\overline{v_1}, \ldots, \overline{v_m})$ to* (6.134).

(i) $x_n \rightharpoonup \overline{x}$ *as* $n \to +\infty$.

(ii) $(\forall i \in \{1, \ldots, m\})$ $v_{i,n} \rightharpoonup \overline{v_i}$ *as* $n \to +\infty$.

(iii) *Suppose that C is demiregular at $\overline{x}$. Then $x_n \to \overline{x}$ as $n \to +\infty$.*

(iv) *Suppose that, for some $j \in \{1, \ldots, m\}$, D_j^{-1} is demiregular at $\overline{v_j}$. Then $v_{j,n} \to \overline{v_j}$ as $n \to +\infty$.*

Proof. Define $\boldsymbol{\mathcal{G}}$ as in Notation 6.8 and set $\boldsymbol{\mathcal{K}} = \mathcal{H} \oplus \boldsymbol{\mathcal{G}}$. We denote the scalar product and the norm of $\boldsymbol{\mathcal{K}}$ by $\langle\langle\langle \cdot \mid \cdot \rangle\rangle\rangle$ and $|||| \cdot ||||$, respectively. As shown in [16, 41], the operators

$$\begin{cases} \boldsymbol{A} : \boldsymbol{\mathcal{K}} \to 2^{\boldsymbol{\mathcal{K}}} \colon (x, v_1, \ldots, v_m) \mapsto (\sum_{i=1}^m L_i^* v_i - z + Ax) \times (r_1 - L_1 x + B_1^{-1} v_1) \times \ldots \times \\ \qquad\qquad\qquad\qquad\qquad\qquad (r_m - L_m x + B_m^{-1} v_m) \\ \boldsymbol{B} : \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}} \colon (x, v_1, \ldots, v_m) \mapsto (Cx, D_1^{-1} v_1, \ldots, D_m^{-1} v_m) \\ \boldsymbol{S} : \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}} \colon (x, v_1, \ldots, v_m) \mapsto \Big(\sum_{i=1}^m L_i^* v_i, -L_1 x, \ldots, -L_m x\Big) \end{cases} \tag{6.140}$$

are maximally monotone and, moreover, $\boldsymbol{B}$ is β-cocoercive [41, Eq. (3.12)]. Furthermore, as shown in [16, Section 3], under condition (6.135), zer$(\boldsymbol{A} + \boldsymbol{B}) \neq \varnothing$ and

$$(\overline{x}, \overline{\boldsymbol{v}}) \in \text{zer}(\boldsymbol{A} + \boldsymbol{B}) \quad \Rightarrow \quad \overline{x} \text{ solves (6.133) and } \overline{\boldsymbol{v}} \text{ solves (6.134)}. \tag{6.141}$$

Next, for every $n \in \mathbb{N}$, define

$$\begin{cases} \boldsymbol{U}_n \colon \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}} \colon (x, \boldsymbol{v}) \mapsto \Big(U_n x, U_{1,n} v_1, \ldots, U_{m,n} v_m\Big) \\ \boldsymbol{V}_n \colon \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}} \colon (x, \boldsymbol{v}) \mapsto \Big(U_n^{-1} x - \sum_{i=1}^m L_i^* v_i, \big(-L_i x + U_{i,n}^{-1} v_i\big)_{1 \leqslant i \leqslant m}\Big) \\ \boldsymbol{T}_n \colon \mathcal{H} \to \boldsymbol{\mathcal{G}} \colon x \mapsto \Big(\sqrt{U_{1,n}} L_1 x, \ldots, \sqrt{U_{m,n}} L_m x\Big). \end{cases} \tag{6.142}$$

It follows from our assumptions and Lemma 6.9(iii) that

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{U}_{n+1} \succcurlyeq \boldsymbol{U}_n \in \mathcal{P}_\alpha(\boldsymbol{\mathcal{K}}) \quad \text{and} \quad \|\boldsymbol{U}_n^{-1}\| \leqslant \frac{1}{\alpha}. \tag{6.143}$$

Moreover, for every $n \in \mathbb{N}$, $\boldsymbol{V}_n \in \mathcal{S}(\boldsymbol{\mathcal{K}})$ since $\boldsymbol{U}_n \in \mathcal{S}(\boldsymbol{\mathcal{K}})$. In addition, (6.142) and (6.143) yield

$$\big(\forall n \in \mathbb{N}\big) \quad \|\boldsymbol{V}_n\| \leqslant \|\boldsymbol{U}_n^{-1}\| + \|\boldsymbol{S}\| \leqslant \rho, \quad \text{where} \quad \rho = \frac{1}{\alpha} + \sqrt{\sum_{i=1}^m \|L_i\|^2}. \tag{6.144}$$

On the other hand,

$$\begin{aligned}(\forall n \in \mathbb{N})(\forall x \in \mathcal{H}) \quad |||\boldsymbol{T}_n x|||^2 &= \sum_{i=1}^m \big\|\sqrt{U_{i,n}} L_i \sqrt{U_n}\sqrt{U_n}^{-1} x\big\|^2 \\ &\leqslant \|x\|^2_{U_n^{-1}} \sum_{i=1}^m \big\|\sqrt{U_{i,n}} L_i \sqrt{U_n}\big\|^2 \\ &= \beta_n \|x\|^2_{U_n^{-1}}, \end{aligned} \tag{6.145}$$

where $(\forall n \in \mathbb{N})$ $\beta_n = \sum_{i=1}^m \big\|\sqrt{U_{i,n}} L_i \sqrt{U_n}\big\|^2$. Hence, (6.137) yields

$$(\forall n \in \mathbb{N}) \quad (1+\delta_n)\beta_n = \frac{1}{1+\delta_n}. \tag{6.146}$$

For every $n \in \mathbb{N}$, set

$$\kappa_n = 2\langle\langle \sqrt{(1+\delta_n)\beta_n}^{-1} \boldsymbol{T}_n x \mid \sqrt{(1+\delta_n)\beta_n}\big(\sqrt{U_{1,n}}^{-1} v_1, \ldots, \sqrt{U_{m,n}}^{-1} v_m\big)\rangle\rangle \tag{6.147}$$

Therefore, for every $n \in \mathbb{N}$ and every $\boldsymbol{x} = (x, v_1, \ldots, v_m) \in \boldsymbol{\mathcal{K}}$, using (6.142), (6.145), (6.146), Lemma 6.9(ii), and (6.138), we obtain

$$\begin{aligned}
\langle\!\langle\!\langle \boldsymbol{x} \mid \boldsymbol{V}_n \boldsymbol{x} \rangle\!\rangle\!\rangle &= \left\langle x \mid U_n^{-1} x \right\rangle + \sum_{i=1}^m \left\langle v_i \mid U_{i,n}^{-1} v_i \right\rangle - 2 \sum_{i=1}^m \langle L_i x \mid v_i \rangle \\
&= \|x\|^2_{U_n^{-1}} + \sum_{i=1}^m \|v_i\|^2_{U_{i,n}^{-1}} - 2 \sum_{i=1}^m \left\langle \sqrt{U_{i,n}} L_i x \mid \sqrt{U_{i,n}}^{\,-1} v_i \right\rangle \\
&= \|x\|^2_{U_n^{-1}} + \sum_{i=1}^m \|v_i\|^2_{U_{i,n}^{-1}} - \kappa_n \\
&\geqslant \|x\|^2_{U_n^{-1}} + \sum_{i=1}^m \|v_i\|^2_{U_{i,n}^{-1}} - \left(\frac{|||\boldsymbol{T}_n x|||^2}{(1+\delta_n)\beta_n} + (1+\delta_n)\beta_n \sum_{i=1}^m \|v_i\|^2_{U_{i,n}^{-1}} \right) \\
&\geqslant \|x\|^2_{U_n^{-1}} + \sum_{i=1}^m \|v_i\|^2_{U_{i,n}^{-1}} - \left(\frac{\|x\|^2_{U_n^{-1}}}{(1+\delta_n)} + (1+\delta_n)\beta_n \sum_{i=1}^m \|v_i\|^2_{U_{i,n}^{-1}} \right) \\
&= \frac{\delta_n}{1+\delta_n} \left(\|x\|^2_{U_n^{-1}} + \sum_{i=1}^m \|v_i\|^2_{U_{i,n}^{-1}} \right) \\
&\geqslant \frac{\delta_n}{1+\delta_n} \left(\|U_n\|^{-1} \|x\|^2 + \sum_{i=1}^m \|U_{i,n}\|^{-1} \|v_i\|^2 \right) \\
&\geqslant \zeta_n |||\boldsymbol{x}|||^2.
\end{aligned} \tag{6.148}$$

In turn, it follows from Lemma 6.9(iii) and (6.138) that

$$(\forall n \in \mathbb{N}) \quad \|\boldsymbol{V}_n^{-1}\| \leqslant \frac{1}{\zeta_n} \leqslant 2\beta - \varepsilon. \tag{6.149}$$

Moreover, by Lemma 6.9(i), $(\forall n \in \mathbb{N})$ $(\boldsymbol{U}_{n+1} \succcurlyeq \boldsymbol{U}_n \Rightarrow \boldsymbol{U}_n^{-1} \succcurlyeq \boldsymbol{U}_{n+1}^{-1} \Rightarrow \boldsymbol{V}_n \succcurlyeq \boldsymbol{V}_{n+1} \Rightarrow \boldsymbol{V}_{n+1}^{-1} \succcurlyeq \boldsymbol{V}_n^{-1})$. Furthermore, we derive from Lemma 6.9(ii) and (6.144) that

$$(\forall \boldsymbol{x} \in \boldsymbol{\mathcal{K}}) \quad \langle\!\langle\!\langle \boldsymbol{V}_n^{-1} \boldsymbol{x} \mid \boldsymbol{x} \rangle\!\rangle\!\rangle \geqslant \|\boldsymbol{V}_n\|^{-1} |||\boldsymbol{x}|||^2 \geqslant \frac{1}{\rho} |||\boldsymbol{x}|||^2. \tag{6.150}$$

Altogether,

$$\sup_{n \in \mathbb{N}} \|\boldsymbol{V}_n^{-1}\| \leqslant 2\beta - \varepsilon \qquad \text{and} \quad (\forall n \in \mathbb{N}) \quad \boldsymbol{V}_{n+1}^{-1} \succcurlyeq \boldsymbol{V}_n^{-1} \in \mathcal{P}_{1/\rho}(\boldsymbol{\mathcal{K}}). \tag{6.151}$$

Now set, for every $n \in \mathbb{N}$,

$$\begin{cases}
\boldsymbol{x}_n = (x_n, v_{1,n}, \ldots, v_{m,n}) \\
\boldsymbol{y}_n = (p_n, q_{1,n}, \ldots, q_{m,n}) \\
\boldsymbol{a}_n = (a_n, b_{1,n}, \ldots, b_{m,n}) \\
\boldsymbol{c}_n = (c_n, d_{1,n}, \ldots, d_{m,n}) \\
\boldsymbol{d}_n = (U_n^{-1} a_n, U_{1,n}^{-1} b_{1,n}, \ldots, U_{m,n}^{-1} b_{m,n})
\end{cases} \quad \text{and} \quad \boldsymbol{b}_n = (\boldsymbol{S} + \boldsymbol{V}_n)\boldsymbol{a}_n + \boldsymbol{c}_n - \boldsymbol{d}_n. \tag{6.152}$$

Then $\sum_{n\in\mathbb{N}} ||||\boldsymbol{a}_n|||| < +\infty$, $\sum_{n\in\mathbb{N}} ||||\boldsymbol{c}_n|||| < +\infty$, and $\sum_{n\in\mathbb{N}} ||||\boldsymbol{d}_n|||| < +\infty$. Therefore (6.144) implies that $\sum_{n\in\mathbb{N}} ||||\boldsymbol{b}_n|||| < +\infty$. Furthermore, using the same arguments as in [41, Eqs. (3.22)–(3.35)], we derive from (6.139) and (6.140) that

$$(\forall n\in\mathbb{N})\quad \boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda_n\Big(J_{\boldsymbol{V}_n^{-1}\boldsymbol{A}}\big(\boldsymbol{x}_n - \boldsymbol{V}_n^{-1}(\boldsymbol{B}\boldsymbol{x}_n + \boldsymbol{b}_n)\big) + \boldsymbol{a}_n - \boldsymbol{x}_n\Big). \tag{6.153}$$

We observe that (6.153) has the structure of the variable metric forward-backward splitting algorithm (6.62), where $(\forall n\in\mathbb{N})\ \gamma_n = 1$. Finally, (6.149) and (6.151) imply that all the conditions in Theorem 6.24 are satisfied.

(i)&(ii) : Theorem 6.24(i) asserts that there exists

$$\overline{\boldsymbol{x}} = (\overline{x}, \overline{v_1}, \ldots, \overline{v_m}) \in \operatorname{zer}(\boldsymbol{A}+\boldsymbol{B}) \tag{6.154}$$

such that $\boldsymbol{x}_n \rightharpoonup \overline{\boldsymbol{x}}$ as $n\to+\infty$. In view of (6.141), the assertions are proved.

(iii)&(iv) : It follows from Theorem 6.24(ii) that $\boldsymbol{B}\boldsymbol{x}_n \to \boldsymbol{B}\overline{\boldsymbol{x}}$ as $n\to+\infty$. Hence, (6.140), (6.152), and (6.154) yield

$$Cx_n \to C\overline{x} \quad\text{and}\quad \big(\forall i\in\{1,\ldots,m\}\big)\ \ D_i^{-1}v_{i,n} \to D_i^{-1}\overline{v_i} \quad\text{as}\quad n\to+\infty. \tag{6.155}$$

Hence the results follow from (i)&(ii) and Definition 6.22. ◻

Remark 6.37 In the case when $C = \rho\,\mathrm{Id}$ for some $\rho\in\,]0,+\infty[$, Problem 6.35 reduces to Problem 6.28. However, the algorithm obtained in Corollary 6.29 is quite different from that of Corollary 6.36. Indeed, the former was obtained by applying the forward-backward algorithm (6.62) to the dual inclusion, which was made possible by the strong monotonicity of the primal problem. By contrast, the latter relies on an application of (6.62) in a primal-dual product space.

Example 6.38 Let $z\in\mathcal{H}$, let $f\in\Gamma_0(\mathcal{H})$, let $\mu\in\,]0,+\infty[$, let $h\colon\mathcal{H}\to\mathbb{R}$ be convex and differentiable with a μ^{-1}-Lipschitzian gradient, let $(a_n)_{n\in\mathbb{N}}$ and $(c_n)_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{H}$, let $\alpha\in\,]0,+\infty[$, let m be a strictly positive integer, and let $(U_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{H})$ such that $(\forall n\in\mathbb{N})\ U_{n+1}\succcurlyeq U_n$. For every $i\in\{1,\ldots,m\}$, let $r_i\in\mathcal{G}_i$, let $g_i\in\Gamma_0(\mathcal{G}_i)$, let $\nu_i\in\,]0,+\infty[$, let $\ell_i\in\Gamma_0(\mathcal{G}_i)$ be ν_i-strongly convex, let $v_{i,0}\in\mathcal{G}_i$, let $(b_{i,n})_{n\in\mathbb{N}}$ and $(d_{i,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{G}_i$, suppose that $0\neq L_i\in\mathcal{B}(\mathcal{H},\mathcal{G}_i)$, and let $(U_{i,n})_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{G}_i)$ such that $(\forall n\in\mathbb{N})$ $U_{i,n+1}\succcurlyeq U_{i,n}$. Furthermore, suppose that

$$z\in\operatorname{ran}\Big(\partial f + \sum_{i=1}^m L_i^*(\partial g_i\,\square\,\partial\ell_i)(L_i\cdot - r_i) + \nabla h\Big). \tag{6.156}$$

The primal problem is

$$\underset{x\in\mathcal{H}}{\text{minimize}}\ f(x) + \sum_{i=1}^m (g_i\,\square\,\ell_i)(L_ix - r_i) + h(x) - \langle x \mid z\rangle, \tag{6.157}$$

and the dual problem is

$$\underset{v_1\in\mathcal{G}_1,\ldots,v_m\in\mathcal{G}_m}{\text{minimize}} \ (f^*\,\square\, h^*)\Big(z-\sum_{i=1}^m L_i^*v_i\Big)+\sum_{i=1}^m \big(g_i^*(v_i)+\ell_i^*(v_i)+\langle v_i \mid r_i\rangle\big). \tag{6.158}$$

Let $\beta=\min\{\mu,\nu_1,\ldots,\nu_m\}$, let $\varepsilon\in\,]0,\min\{1,\beta\}[$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon,1]$, suppose that (6.138) holds, and set

$$(\forall n\in\mathbb{N})\quad \left\lfloor \begin{array}{l} p_n=\text{prox}_f^{U_n^{-1}}\Big(x_n-U_n\big(\sum_{i=1}^m L_i^*v_{i,n}+\nabla h(x_n)+c_n-z\big)\Big)+a_n\\ y_n=2p_n-x_n\\ x_{n+1}=x_n+\lambda_n(p_n-x_n)\\ \text{For } i=1,\ldots,m\\ \left\lfloor \begin{array}{l} q_{i,n}=\text{prox}_{g_i^*}^{U_{i,n}^{-1}}\Big(v_{i,n}+U_{i,n}\big(L_iy_n-\nabla\ell_i^*(v_{i,n})-d_{i,n}-r_i\big)\Big)+b_{i,n}\\ v_{i,n+1}=v_{i,n}+\lambda_n(q_{i,n}-v_{i,n}). \end{array}\right. \end{array}\right. \tag{6.159}$$

Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a solution to (6.157), for every $i\in\{1,\ldots,m\}$ $(v_{i,n})_{n\in\mathbb{N}}$ converges weakly to some $\overline{v}_i\in\mathcal{G}_i$, and $(\overline{v}_1,\ldots,\overline{v}_m)$ is a solution to (6.158).

Proof. Set $A=\partial f$, $C=\nabla h$, and $(\forall i\in\{1,\ldots,m\})$ $B_i=\partial g_i$ and $D_i=\partial\ell_i$. In this setting, it follows from the analysis of [16, Section 4] that (6.157)–(6.158) is a special case of Problem 6.35 and, using (6.48), that (6.159) is a special case of (6.139). Thus, the claims follow from Corollary 6.36(i)&(ii). □

Remark 6.39 Suppose that, in Corollary 6.36 and Example 6.38, there exist τ and $(\sigma_i)_{1\leqslant i\leqslant m}$ in $]0,+\infty[$ such that $(\forall n\in\mathbb{N})$ $U_n=\tau\,\mathrm{Id}$ and $(\forall i\in\{1,\ldots,m\})$ $U_{i,n}=\sigma_i\,\mathrm{Id}$. Then (6.139) and (6.159) reduce to the fixed metric methods appearing in [41, Eq. (3.3)] and [41, Eq. (4.5)], respectively (see [41] for further connections with existing work).

6.3 Bibliographie

[1] H. Attouch, L. M. Briceño-Arias, and P. L. Combettes, A parallel splitting method for coupled monotone inclusions, *SIAM J. Control Optim.*, vol. 48, pp. 3246–3270, 2010.

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

[3] H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Bregman monotone optimization algorithms, *SIAM J. Control Optim.*, vol. 42, pp. 596–636, 2003.

[4] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York, 2011.

[5] J. F. Bonnans, J. Ch. Gilbert, C. Lemaréchal, and C. A. Sagastizábal, A family of variable metric proximal methods, *Math. Programming*, vol. 68, pp. 15–47, 1995.

[6] R. I. Boţ, *Conjugate Duality in Convex Optimization*, Lecture Notes in Economics and Mathematical Systems, vol. 637. Springer, New York, 2010.

[7] L. M. Briceño-Arias and P. L. Combettes, Monotone operator methods for Nash equilibria in non-potential games, in *Computational and Analytical Mathematics,* (D. Bailey, H. H. Bauschke, P. Borwein, F. Garvan, M. Théra, J. Vanderwerff, and H. Wolkowicz, eds.). Springer, New York, 2013.

[8] J. V. Burke and M. Qian, A variable metric proximal point algorithm for monotone operators, *SIAM J. Control Optim.,* vol. 37, pp. 353–375, 1999.

[9] J. V. Burke and M. Qian, On the superlinear convergence of the variable metric proximal point algorithm using Broyden and BFGS matrix secant updating, *Math. Programming*, vol. 88, pp. 157–181, 2000.

[10] G. H.-G. Chen and R. T. Rockafellar, Convergence rates in forward-backward splitting, *SIAM J. Optim.*, vol. 7, pp. 421–444, 1997.

[11] X. Chen and M. Fukushima, Proximal quasi-Newton methods for nondifferentiable convex optimization, *Math. Program.,* vol. 85, pp. 313–334, 1999.

[12] P. L. Combettes, Solving monotone inclusions via compositions of nonexpansive averaged operators, *Optimization,* vol. 53, pp. 475–504, 2004.

[13] P. L. Combettes, Đinh Dũng, and B. C. Vũ, Dualization of signal recovery problems, *Set-Valued Var. Anal.*, vol. 18, pp. 373–404, 2010.

[14] P. L. Combettes, Đinh Dũng, and B. C. Vũ, Proximity for sums of composite functions, *J. Math. Anal. Appl.*, vol. 380, pp. 680–688, 2011.

[15] P. L. Combettes and J.-C. Pesquet, Proximal splitting methods in signal processing, in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering,* (H. H. Bauschke et al., eds), pp. 185–212. Springer, New York, 2011.

[16] P. L. Combettes and J.-C. Pesquet, Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators, *Set-Valued Var. Anal.*, vol. 20, pp. 307–330, 2012.

[17] P. L. Combettes and B. C. Vũ, Variable metric quasi-Fejér monotonicity, *Nonlinear Anal.,* vol. 78, pp. 17–31, 2013.

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

[19] W. C. Davidon, Variable metric method for minimization, Argonne National Laboratory research and development report 5990, 1959 ; reprinted in *SIAM J. Optim.,* vol. 1, pp. 1–17, 1991.

[20] E. De Vito, V. Umanità, and S. Villa, A consistent algorithm to solve Lasso, elastic-net and Tikhonov regularization, *J. Complexity,* vol. 27, pp. 188–200, 2011.

[21] J. Duchi and Y. Singer, Efficient online and batch learning using forward backward splitting, *J. Mach. Learn. Res.,* vol. 10, pp. 2899–2934, 2009.

[22] J. Eckstein and M. C. Ferris, Smooth methods of multipliers for complementarity problems, *Math. Programming,* vol. 86, pp. 65–90, 1999.

[23] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems.* Springer-Verlag, New York, 2003.

[24] R. Glowinski and P. Le Tallec, *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics.* SIAM, Philadelphia, 1989.

[25] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms.* Springer-Verlag, New York, 1993.

[26] T. Kato, *Perturbation Theory for Linear Operators,* 2nd ed. Springer-Verlag, New York, 1980.

[27] C. Lemaréchal and C. Sagastizábal, Variable metric bundle methods : from conceptual to implementable forms, *Math. Programming*, vol. 76, pp. 393–410, 1997.

[28] P. A. Lotito, L. A. Parente, and M. V. Solodov, A class of variable metric decomposition methods for monotone variational inclusions, *J. Convex Anal.*, vol. 16, pp. 857–880, 2009.

[29] B. Mercier, *Topics in Finite Element Solution of Elliptic Problems* (Lectures on Mathematics, no. 63). Tata Institute of Fundamental Research, Bombay, 1979.

[30] B. Mercier, *Inéquations Variationnelles de la Mécanique* (Publications Mathématiques d'Orsay, no. 80.01). Université de Paris-XI, Orsay, France, 1980.

[31] U. Mosco, Dual variational inequalities, *J. Math. Anal. Appl.,* vol. 40, pp. 202–206, 1972.

[32] L. A. Parente, P. A. Lotito, and M. V. Solodov, A class of inexact variable metric proximal point algorithms, *SIAM J. Optim.*, vol. 19, pp. 240–260, 2008.

[33] J. D. Pearson, Variable metric methods of minimisation, *Comput. J.,* vol. 12, pp. 171–178, 1969.

[34] T. Pennanen, Dualization of generalized equations of maximal monotone type, *SIAM J. Optim.,* vol. 10, pp. 809–835, 2000.

[35] L. Qi and X. Chen, A preconditioning proximal Newton method for nondifferentiable convex optimization, *Math. Programming*, vol. 76, pp. 411–429, 1997.

[36] M. Qian, *The Variable Metric Proximal Point Algorithm : Theory and Application,* Ph.D. thesis, University of Washington, Seattle, WA, 1992.

[37] S. M. Robinson, Composition duality and maximal monotonicity, *Math. Programming,* vol. 85, pp. 1–13, 1999.

[38] R. T. Rockafellar, Duality and stability in extremum problems involving convex functions, *Pacific J. Math.,* vol. 21, pp. 167–187, 1967.

[39] P. Tseng, Further applications of a splitting algorithm to decomposition in variational inequalities and convex programming, *Math. Programming*, vol. 48, pp. 249–263, 1990.

[40] P. Tseng, Applications of a splitting algorithm to decomposition in convex programming and variational inequalities, *SIAM J. Control Optim.*, vol. 29, pp. 119–138, 1991.

[41] B. C. Vũ, A splitting algorithm for dual monotone inclusions involving cocoercive operators, *Adv. Comput. Math.*, vol. 38, pp. 667–681, 2013.

[42] D. L. Zhu and P. Marcotte, Co-coercivity and its role in the convergence of iterative schemes for solving variational inequalities, *SIAM J. Optim.,* vol. 6, pp. 714–726, 1996.

Chapitre 7

Méthode explicite-implicite-explicite à métrique variable

Nous proposons une extension avec métrique variable de l'algorithme explicite-implicite-explicite (1.6) pour trouver un zéro de la somme d'un opérateur maximalement monotone et d'un opérateur monotone et lipschitzien. Ce cadre nous donne un algorithme d'éclatement à métrique variable pour résoudre des inclusions monotones composites.

7.1 Description et résultats principaux

Le résultat principal de ce chapitre est le suivant. On note $\ell^1_+(\mathbb{N})$ l'ensemble des suites absolument sommables dans $[0,+\infty[$.

Théorème 7.1 *Soient α et β des réels strictement positifs, soit $(\eta_n)_{n\in\mathbb{N}}$ une suite dans $\ell^1_+(\mathbb{N})$, et soit $(\boldsymbol{U}_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{B}\,(\boldsymbol{\mathcal{K}})$ telle que*

$$\mu = \sup_{n\in\mathbb{N}} \|\boldsymbol{U}_n\| < +\infty \quad \textit{et} \quad (1+\eta_n)\boldsymbol{U}_{n+1} \succcurlyeq \boldsymbol{U}_n \in \mathcal{P}_\alpha(\boldsymbol{\mathcal{K}}). \tag{7.1}$$

Soient $\boldsymbol{A}\colon \boldsymbol{\mathcal{K}} \to 2^{\boldsymbol{\mathcal{K}}}$ un opérateur maximalement monotone, et $\boldsymbol{B}\colon \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}}$ un opérateur monotone et β-lipschitzien sur $\boldsymbol{\mathcal{K}}$. Supposons que

$$\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B}) \neq \varnothing. \tag{7.2}$$

Soient $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$, $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$, et $(\boldsymbol{c}_n)_{n\in\mathbb{N}}$ des suites absolument sommables dans $\boldsymbol{\mathcal{K}}$, $\boldsymbol{x}_0 \in \boldsymbol{\mathcal{K}}$, $\varepsilon \in\,]0, 1/(\beta\mu+1)[$, $(\gamma_n)_{n\in\mathbb{N}}$ une suite dans $[\varepsilon, (1-\varepsilon)/(\beta\mu)]$. Posons

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} \boldsymbol{y}_n = \boldsymbol{x}_n - \gamma_n \boldsymbol{U}_n(\boldsymbol{B}\boldsymbol{x}_n + \boldsymbol{a}_n) \\ \boldsymbol{p}_n = J_{\gamma_n \boldsymbol{U}_n \boldsymbol{A}} \boldsymbol{y}_n + \boldsymbol{b}_n \\ \boldsymbol{q}_n = \boldsymbol{p}_n - \gamma_n \boldsymbol{U}_n(\boldsymbol{B}\boldsymbol{p}_n + \boldsymbol{c}_n) \\ \boldsymbol{x}_{n+1} = \boldsymbol{x}_n - \boldsymbol{y}_n + \boldsymbol{q}_n. \end{array} \right. \tag{7.3}$$

Alors, on a les résultats suivants pour un point $\overline{\boldsymbol{x}} \in \mathrm{zer}(\boldsymbol{A}+\boldsymbol{B})$.

(i) $\sum_{n\in\mathbb{N}} ||||\boldsymbol{x}_n - \boldsymbol{p}_n||||^2 < +\infty$ *et* $\sum_{n\in\mathbb{N}} ||||\boldsymbol{y}_n - \boldsymbol{q}_n||||^2 < +\infty$.

(ii) $\boldsymbol{x}_n \rightharpoonup \overline{\boldsymbol{x}}$ *et* $\boldsymbol{p}_n \rightharpoonup \overline{\boldsymbol{x}}$.

(iii) *Supposons que l'une des conditions suivantes soit satisfaite.*

(a) $\underline{\lim}\, d_{\mathrm{zer}(\boldsymbol{A}+\boldsymbol{B})}(\boldsymbol{x}_n) = 0$.

(b) $\boldsymbol{A}+\boldsymbol{B}$ *est demirégulier en* $\overline{\boldsymbol{x}}$.

(c) $\boldsymbol{A}$ *ou* $\boldsymbol{B}$ *est uniformément monotone en* $\overline{\boldsymbol{x}}$.

(d) $\mathrm{int}\,\mathrm{zer}(\boldsymbol{A}+\boldsymbol{B}) \neq \varnothing$ *et il existe* $(\nu_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$ *telle que* $(\forall n \in \mathbb{N})\ (1+\nu_n)\boldsymbol{U}_n \succeq \boldsymbol{U}_{n+1}$.

Alors $\boldsymbol{x}_n \to \overline{\boldsymbol{x}}$.

Nous allons nous intéressons à la résolution d'inclusions monotones impliquant des opérateurs lipschitziens et monotones.

Problème 7.2 Soient $\mathcal{H}$ un espace hilbertien réel, $z \in \mathcal{H}$, m un entier strictement positif, $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ un opérateur maximalement monotone, $\nu_0 \in \,]0,+\infty[$, et $C\colon \mathcal{H} \to \mathcal{H}$ un opérateur ν_0-lipschitzien et monotone. Pour tout $i \in \{1,\ldots,m\}$, soient $\mathcal{G}_i$ un espace hilbertien réel, $r_i \in \mathcal{G}_i$, $\nu_i \in \,]0,+\infty[$, $B_i\colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ un opérateur maximalement monotone, $D_i\colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ un opérateur monotone tel que D_i^{-1} est ν_i-lipschitzien, et $0 \neq L_i \in \mathcal{B}(\mathcal{H},\mathcal{G}_i)$. Supposons que

$$z \in \mathrm{ran}\left(A + \sum_{i=1}^m L_i^*\big((B_i \,\square\, D_i)(L_i\cdot - r_i)\big) + C\right). \tag{7.4}$$

Le problème est de résoudre l'inclusion primale

$$\text{trouver } \overline{x} \in \mathcal{H} \text{ tel que } z \in A\overline{x} + \sum_{i=1}^m L_i^*\big((B_i \,\square\, D_i)(L_i\overline{x} - r_i)\big) + C\overline{x}, \tag{7.5}$$

et l'inclusion duale

$$\text{trouver } \overline{v}_1 \in \mathcal{G}_1,\ldots,\overline{v}_m \in \mathcal{G}_m \text{tel que}$$
$$(\exists\, x \in \mathcal{H}) \quad \begin{cases} z - \sum_{i=1}^m L_i^*\overline{v}_i \in Ax + Cx,\\ (\forall i \in \{1,\ldots,m\})\ \overline{v}_i \in (B_i \,\square\, D_i)(L_ix - r_i). \end{cases} \tag{7.6}$$

Corollaire 7.3 *Soit* $\alpha \in \,]0,+\infty[$*, soit* $(\eta_{0,n})_{n\in\mathbb{N}}$ *une suite dans* $\ell_+^1(\mathbb{N})$*, soit* $(U_n)_{n\in\mathbb{N}}$ *une suite dans* $\mathcal{P}_\alpha(\mathcal{H})$*, et pour tout* $i \in \{1,\ldots,m\}$*, soit* $(\eta_{i,n})_{n\in\mathbb{N}}$ *une suite dans* $\ell_+^1(\mathbb{N})$*, soit* $(U_{i,n})_{n\in\mathbb{N}}$ *une suite dans* $\mathcal{P}_\alpha(\mathcal{G}_i)$ *telle que* $\mu = \sup_{n\in\mathbb{N}}\{\|U_n\|, \|U_{1,n}\|,\ldots,\|U_{m,n}\|\} < +\infty$ *et*

$$(\forall n \in \mathbb{N}) \quad (1+\eta_{0,n})U_{n+1} \succcurlyeq U_n, \ \textit{et}\ (\forall i \in \{1,\ldots,m\}) \quad (1+\eta_{i,n})U_{i,n+1} \succcurlyeq U_{i,n}. \tag{7.7}$$

Soient $(a_{1,n})_{n\in\mathbb{N}}, (b_{1,n})_{n\in\mathbb{N}}$*, et* $(c_{1,n})_{n\in\mathbb{N}}$ *des suites absolument sommables dans* $\mathcal{H}$*, et pour tout* $i \in \{1,\ldots,m\}$*, soient* $(a_{2,i,n})_{n\in\mathbb{N}}, (b_{2,i,n})_{n\in\mathbb{N}}$*, et* $(c_{2,i,n})_{n\in\mathbb{N}}$ *des suites absolument sommables dans* $\mathcal{G}_i$*. De plus, posons*

$$\beta = \max\{\nu_0, \nu_1, \ldots, \nu_m\} + \sqrt{\sum_{i=1}^{m} \|L_i\|^2}, \tag{7.8}$$

soit $x_0 \in \mathcal{H}$*, soit* $(v_{1,0}, \ldots, v_{m,0}) \in \mathcal{G}_1 \oplus \ldots \oplus \mathcal{G}_m$*, soit* $\varepsilon \in \left]0, 1/(1+\beta\mu)\right[$*, et soit* $(\gamma_n)_{n\in\mathbb{N}}$ *une suite dans* $[\varepsilon, (1-\varepsilon)/(\beta\mu)]$*. Posons*

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} y_{1,n} = x_n - \gamma_n U_n\big(Cx_n + \sum_{i=1}^m L_i^* v_{i,n} + a_{1,n}\big) \\ p_{1,n} = J_{\gamma_n U_n A}(y_{1,n} + \gamma_n U_n z) + b_{1,n} \\ \text{pour } i = 1, \ldots, m \\ \left\lfloor \begin{array}{l} y_{2,i,n} = v_{i,n} + \gamma_n U_{i,n}\big(L_i x_n - D_i^{-1} v_{i,n} + a_{2,i,n}\big) \\ p_{2,i,n} = J_{\gamma_n U_{i,n} B_i^{-1}}(y_{2,i,n} - \gamma_n U_{i,n} r_i) + b_{2,i,n} \\ q_{2,i,n} = p_{2,i,n} + \gamma_n U_{i,n}\big(L_i p_{1,n} - D_i^{-1} p_{2,i,n} + c_{2,i,n}\big) \\ v_{i,n+1} = v_{i,n} - y_{2,i,n} + q_{2,i,n} \end{array} \right. \\ q_{1,n} = p_{1,n} - \gamma_n U_n\big(Cp_{1,n} + \sum_{i=1}^m L_i^* p_{2,i,n} + c_{1,n}\big) \\ x_{n+1} = x_n - y_{1,n} + q_{1,n}. \end{array} \right. \tag{7.9}$$

Alors on a les résultats suivants.

(i) $\sum_{n\in\mathbb{N}} \|x_n - p_{1,n}\|^2 < +\infty$ *et* $(\forall i \in \{1,\ldots,m\})$ $\sum_{n\in\mathbb{N}} \|v_{i,n} - p_{2,i,n}\|^2 < +\infty$.

(ii) *Il existe une solution* $\overline{x}$ *du problème* (7.5) *et une solution* $(\overline{v}_1, \ldots, \overline{v}_m)$ *du problème* (7.6) *telles que :*

  (a) $x_n \rightharpoonup \overline{x}$ *et* $p_{1,n} \rightharpoonup \overline{x}$.

  (b) $(\forall i \in \{1,\ldots,m\})$ $v_{i,n} \rightharpoonup \overline{v}_i$ *et* $p_{2,i,n} \rightharpoonup \overline{v}_i$.

  (c) *Si* A *ou* C *est uniformément monotone en* $\overline{x}$*, alors* $x_n \to \overline{x}$ *et* $p_{1,n} \to \overline{x}$.

  (d) *Si* B_j^{-1} *ou* D_j^{-1} *est uniformément monotone en* $\overline{v}_j$*, pour* $j \in \{1,\ldots,m\}$*, alors* $v_{j,n} \to \overline{v}_j$ *et* $p_{2,j,n} \to \overline{v}_j$.

7.2 Article en anglais

A VARIABLE METRIC EXTENSION OF THE FORWARD-BACKWARD-FORWARD ALGORITHM FOR MONOTONE OPERATORS [1]

Abstract :

We propose a variable metric extension of the forward–backward-forward algorithm for finding a zero of the sum of a maximally monotone operator and a Lipschitzian monotone operator in Hilbert spaces. In turn, this framework provides a variable metric splitting algorithm for solving monotone inclusions involving sums of composite operators. Several splitting algorithms recently proposed in the literature are recovered as special cases.

7.2.1 Introduction

A basic problem in applied monotone operator theory is to find a zero of a maximally monotone operator A on a real Hilbert space $\mathcal{H}$. This problem can be solved by the proximal point algorithm proposed in [17] which requires only the resolvent of A, provided it is easy to implement numerically. In order to get more efficient proximal algorithms, some authors have proposed the use of variable metric or preconditioning in such algorithms [3, 5, 6, 10, 13, 15, 16].

This problem was then extended to the problem of finding a zero of the sum of a maximally monotone operator A and a cocoercive operator B (i.e., B^{-1} is strongly monotone). In such instances, the forward-backward splitting algorithm [1, 8, 12, 18] can be used. Recently, this algorithm has been investigated in the context of variable metric [11]. In the case when B is only Lipschitzian and not cocoercive, the problem can be solved by the forward-backward-forward splitting algorithm [4, 19]. New applications of this basic algorithm to more complex monotone inclusions are presented in [4, 9].

In the present paper, we propose a variable metric version of the forward-backward-forward splitting algorithm. In Section 7.2.2, we recall notation and background on convex analysis and monotone operator theory. In Section 7.2.3, we present our variable metric forward-backward-forward splitting algorithm. In Section 7.2.4, the results of Section 7.2.3 are used to develop a variable metric primal–dual algorithm for solving the type of composite inclusions considered in [9].

1. B. C. Vũ, A variable metric extension of the forward–backward–forward algorithm for monotone operators, *Numerical Functional Analysis and Optimization*, à paraître.

7.2.2 Notation and background

Throughout, $\mathcal{H}$, $\mathcal{G}$, and $(\mathcal{G}_i)_{1\leqslant i\leqslant m}$ are real Hilbert spaces. Their scalar products and associated norms are respectively denoted by $\langle\cdot\mid\cdot\rangle$ and $\|\cdot\|$. We denote by $\mathcal{B}(\mathcal{H},\mathcal{G})$ the space of bounded linear operators from $\mathcal{H}$ to $\mathcal{G}$. The adjoint of $L\in\mathcal{B}(\mathcal{H},\mathcal{G})$ is denoted by L^*. We set $\mathcal{B}(\mathcal{H})=\mathcal{B}(\mathcal{H},\mathcal{H})$. The symbols $\rightharpoonup$ and $\to$ denote respectively weak and strong convergence, and Id denotes the identity operator, and $B(x;\rho)$ denotes the closed ball of center $x\in\mathcal{H}$ and radius $\rho\in\,]0,+\infty[$. The interior of $C\subset\mathcal{H}$ is denoted by $\operatorname{int}C$. We denote by $\ell^1_+(\mathbb{N})$ the set of summable sequences in $[0,+\infty[$.

Let M_1 and M_2 be self-adjoint operators in $\mathcal{B}(\mathcal{H})$, we write $M_1\succcurlyeq M_2$ if and only if $(\forall x\in\mathcal{H})\ \langle M_1x\mid x\rangle\geqslant\langle M_2x\mid x\rangle$. Let $\alpha\in\,]0,+\infty[$. We set

$$\mathcal{P}_\alpha(\mathcal{H})=\big\{M\in\mathcal{B}(\mathcal{H})\;\big|\;M^*=M\quad\text{and}\quad M\succcurlyeq\alpha\operatorname{Id}\big\}. \tag{7.10}$$

Moreover, for every $M\in\mathcal{P}_\alpha(\mathcal{H})$, we define respectively a scalar product and a norm by

$$(\forall x\in\mathcal{H})(\forall y\in\mathcal{H})\quad\langle x\mid y\rangle_M=\langle Mx\mid y\rangle\quad\text{and}\quad\|x\|_M=\sqrt{\langle Mx\mid x\rangle}. \tag{7.11}$$

Let $A\colon\mathcal{H}\to2^{\mathcal{H}}$ be a set-valued operator. The domain is $\operatorname{dom}A=\big\{x\in\mathcal{H}\;\big|\;Ax\neq\varnothing\big\}$, and the graph of A is $\operatorname{gra}A=\big\{(x,u)\in\mathcal{H}\times\mathcal{H}\;\big|\;u\in Ax\big\}$. The set of zeros of A is $\operatorname{zer}A=\big\{x\in\mathcal{H}\;\big|\;0\in Ax\big\}$, and the range of A is $\operatorname{ran}A=\big\{u\in\mathcal{H}\;\big|\;(\exists\,x\in\mathcal{H})\;u\in Ax\big\}$. The inverse of A is $A^{-1}\colon\mathcal{H}\mapsto2^{\mathcal{H}}\colon u\mapsto\big\{x\in\mathcal{H}\;\big|\;u\in Ax\big\}$, and the resolvent of A is

$$J_A=(\operatorname{Id}+A)^{-1}. \tag{7.12}$$

Moreover, A is monotone if

$$(\forall(x,y)\in\mathcal{H}\times\mathcal{H})(\forall(u,v)\in Ax\times Ay)\quad\langle x-y\mid u-v\rangle\geqslant0, \tag{7.13}$$

and maximally monotone if it is monotone and there exists no monotone operator $B\colon\mathcal{H}\to2^{\mathcal{H}}$ such that $\operatorname{gra}A\subset\operatorname{gra}B$ and $A\neq B$. We say that A is uniformly monotone at $x\in\operatorname{dom}A$ if there exists an increasing function $\phi_A\colon\ [0,+\infty[\,\to[0,+\infty]$ vanishing only at 0 such that

$$\big(\forall u\in Ax\big)\big(\forall(y,v)\in\operatorname{gra}A\big)\quad\langle x-y\mid u-v\rangle\geqslant\phi_A(\|x-y\|). \tag{7.14}$$

7.2.3 Variable metric forward-backward-forward splitting algorithm

The forward-backward-forward splitting algorithm was first proposed in [19] to solve inclusion involving the sum of a maximally monotone operator and a Lipschitzian monotone operator. In [4], it was revisited to include computational errors. Below, we extend it to a variable metric setting.

Theorem 7.4 *Let $\boldsymbol{\mathcal{K}}$ be a real Hilbert space with the scalar product $\langle\langle\langle \cdot \mid \cdot \rangle\rangle\rangle$ and the associated norm $|||\cdot|||$. Let α and β be in $]0,+\infty[$, let $(\eta_n)_{n\in\mathbb{N}}$ be a sequence in $\ell_+^1(\mathbb{N})$, and let $(\boldsymbol{U}_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{B}(\boldsymbol{\mathcal{K}})$ such that*

$$\mu = \sup_{n\in\mathbb{N}} \|\boldsymbol{U}_n\| < +\infty \quad \text{and} \quad (1+\eta_n)\boldsymbol{U}_{n+1} \succcurlyeq \boldsymbol{U}_n \in \mathcal{P}_\alpha(\boldsymbol{\mathcal{K}}). \tag{7.15}$$

Let $\boldsymbol{A}\colon \boldsymbol{\mathcal{K}} \to 2^{\boldsymbol{\mathcal{K}}}$ be maximally monotone, let $\boldsymbol{B}\colon \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}}$ be a monotone and β-Lipschitzian operator on $\boldsymbol{\mathcal{K}}$ such that $\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B}) \neq \varnothing$. Let $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$, $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$, and $(\boldsymbol{c}_n)_{n\in\mathbb{N}}$ be absolutely summable sequences in $\boldsymbol{\mathcal{K}}$. Let $\boldsymbol{x}_0 \in \boldsymbol{\mathcal{K}}$, let $\varepsilon \in]0, 1/(\beta\mu+1)[$, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, (1-\varepsilon)/(\beta\mu)]$, and set

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} \boldsymbol{y}_n = \boldsymbol{x}_n - \gamma_n \boldsymbol{U}_n(\boldsymbol{B}\boldsymbol{x}_n + \boldsymbol{a}_n) \\ \boldsymbol{p}_n = J_{\gamma_n \boldsymbol{U}_n \boldsymbol{A}} \boldsymbol{y}_n + \boldsymbol{b}_n \\ \boldsymbol{q}_n = \boldsymbol{p}_n - \gamma_n \boldsymbol{U}_n(\boldsymbol{B}\boldsymbol{p}_n + \boldsymbol{c}_n) \\ \boldsymbol{x}_{n+1} = \boldsymbol{x}_n - \boldsymbol{y}_n + \boldsymbol{q}_n. \end{array} \right. \tag{7.16}$$

Then the following hold for some $\overline{\boldsymbol{x}} \in \operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$.

(i) $\sum_{n\in\mathbb{N}} |||\boldsymbol{x}_n - \boldsymbol{p}_n|||^2 < +\infty$ *and* $\sum_{n\in\mathbb{N}} |||\boldsymbol{y}_n - \boldsymbol{q}_n|||^2 < +\infty$.

(ii) $\boldsymbol{x}_n \rightharpoonup \overline{\boldsymbol{x}}$ *and* $\boldsymbol{p}_n \rightharpoonup \overline{\boldsymbol{x}}$.

(iii) *Suppose that one of the following is satisfied :*

  (a) $\underline{\lim}\, d_{\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})}(\boldsymbol{x}_n) = 0$.

  (b) $\boldsymbol{A}+\boldsymbol{B}$ *is demiregular (see [1, Definition 2.3]) at* $\overline{\boldsymbol{x}}$.

  (c) $\boldsymbol{A}$ *or* $\boldsymbol{B}$ *is uniformly monotone at* $\overline{\boldsymbol{x}}$.

  (d) $\operatorname{int}\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B}) \neq \varnothing$ *and there exists* $(\nu_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$ *such that* $(\forall n \in \mathbb{N})\ (1+\nu_n)\boldsymbol{U}_n \succcurlyeq \boldsymbol{U}_{n+1}$.

  Then $\boldsymbol{x}_n \to \overline{\boldsymbol{x}}$ *and* $\boldsymbol{p}_n \to \overline{\boldsymbol{x}}$.

Proof. It follows from [11, Lemma 3.7] that the sequences $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$, $(\boldsymbol{y}_n)_{n\in\mathbb{N}}$, $(\boldsymbol{p}_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{q}_n)_{n\in\mathbb{N}}$ are well defined. Moreover, using [10, Lemma 2.1(i)(ii)] and (7.15), we obtain

$$(\forall (\boldsymbol{z}_n)_{n\in\mathbb{N}} \in \boldsymbol{\mathcal{K}}^{\mathbb{N}}) \quad \sum_{n\in\mathbb{N}} |||\boldsymbol{z}_n||| < +\infty \quad \Leftrightarrow \quad \sum_{n\in\mathbb{N}} |||\boldsymbol{z}_n|||_{\boldsymbol{U}_n^{-1}} < +\infty \tag{7.17}$$

and

$$(\forall (\boldsymbol{z}_n)_{n\in\mathbb{N}} \in \boldsymbol{\mathcal{K}}^{\mathbb{N}}) \quad \sum_{n\in\mathbb{N}} |||\boldsymbol{z}_n||| < +\infty \quad \Leftrightarrow \quad \sum_{n\in\mathbb{N}} |||\boldsymbol{z}_n|||_{\boldsymbol{U}_n} < +\infty. \tag{7.18}$$

Let us set, for every $n \in \mathbb{N}$,

$$\begin{cases} \widetilde{\boldsymbol{y}}_n = \boldsymbol{x}_n - \gamma_n \boldsymbol{U}_n \boldsymbol{B}\boldsymbol{x}_n \\ \widetilde{\boldsymbol{p}}_n = J_{\gamma_n \boldsymbol{U}_n \boldsymbol{A}} \widetilde{\boldsymbol{y}}_n \\ \widetilde{\boldsymbol{q}}_n = \widetilde{\boldsymbol{p}}_n - \gamma_n \boldsymbol{U}_n \boldsymbol{B}\widetilde{\boldsymbol{p}}_n \\ \widetilde{\boldsymbol{x}}_{n+1} = \boldsymbol{x}_n - \widetilde{\boldsymbol{y}}_n + \widetilde{\boldsymbol{q}}_n, \end{cases} \quad \text{and} \quad \begin{cases} \boldsymbol{u}_n = \gamma_n^{-1}\boldsymbol{U}_n^{-1}(\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n) + \boldsymbol{B}\widetilde{\boldsymbol{p}}_n - \boldsymbol{B}\boldsymbol{x}_n \\ \boldsymbol{e}_n = \widetilde{\boldsymbol{x}}_{n+1} - \boldsymbol{x}_{n+1} \\ \boldsymbol{d}_n = \boldsymbol{q}_n - \widetilde{\boldsymbol{q}}_n + \widetilde{\boldsymbol{y}}_n - \boldsymbol{y}_n. \end{cases} \tag{7.19}$$

Then (7.19) yields

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{u}_n = \gamma_n^{-1}\boldsymbol{U}_n^{-1}(\,\widetilde{\boldsymbol{y}}_n - \widetilde{\boldsymbol{p}}_n) + \boldsymbol{B}\widetilde{\boldsymbol{p}}_n \in \boldsymbol{A}\widetilde{\boldsymbol{p}}_n + \boldsymbol{B}\widetilde{\boldsymbol{p}}_n, \tag{7.20}$$

and (7.19), (7.16), Lemma [11, Lemma 3.7(ii)], and the Lipschitzianity of $\boldsymbol{B}$ on $\boldsymbol{\mathcal{K}}$ yield

$$(\forall n \in \mathbb{N}) \quad \begin{cases} |||\boldsymbol{y}_n - \widetilde{\boldsymbol{y}}_n|||_{\boldsymbol{U}_n^{-1}} & \leqslant (\beta\mu)^{-1}|||\boldsymbol{a}_n|||_{\boldsymbol{U}_n} \\ |||\boldsymbol{p}_n - \widetilde{\boldsymbol{p}}_n|||_{\boldsymbol{U}_n^{-1}} & \leqslant |||\boldsymbol{b}_n|||_{\boldsymbol{U}_n^{-1}} + (\beta\mu)^{-1}|||\boldsymbol{a}_n|||_{\boldsymbol{U}_n} \\ |||\boldsymbol{q}_n - \widetilde{\boldsymbol{q}}_n|||_{\boldsymbol{U}_n^{-1}} & \leqslant 2\Big(|||\boldsymbol{b}_n|||_{\boldsymbol{U}_n^{-1}} + (\beta\mu)^{-1}|||\boldsymbol{a}_n|||_{\boldsymbol{U}_n}\Big) \\ & \quad +(\beta\mu)^{-1}|||\boldsymbol{c}_n|||_{\boldsymbol{U}_n}. \end{cases} \tag{7.21}$$

Since $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$, $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$, and $(\boldsymbol{c}_n)_{n\in\mathbb{N}}$ are absolutely summable sequences in $\boldsymbol{\mathcal{K}}$, we derive from (7.17), (7.18), (7.19), and (7.21) that

$$\begin{cases} \sum_{n\in\mathbb{N}} |||\boldsymbol{p}_n - \widetilde{\boldsymbol{p}}_n||| < +\infty \quad \text{and} \quad \sum_{n\in\mathbb{N}} |||\boldsymbol{p}_n - \widetilde{\boldsymbol{p}}_n|||_{\boldsymbol{U}_n^{-1}} < +\infty \\ \sum_{n\in\mathbb{N}} |||\boldsymbol{q}_n - \widetilde{\boldsymbol{q}}_n||| < +\infty \quad \text{and} \quad \sum_{n\in\mathbb{N}} |||\boldsymbol{q}_n - \widetilde{\boldsymbol{q}}_n|||_{\boldsymbol{U}_n^{-1}} < +\infty \\ \sum_{n\in\mathbb{N}} |||\boldsymbol{d}_n||| < +\infty \quad \text{and} \quad \sum_{n\in\mathbb{N}} |||\boldsymbol{d}_n|||_{\boldsymbol{U}_n^{-1}} < +\infty. \end{cases} \tag{7.22}$$

Now, let $\boldsymbol{x} \in \operatorname{zer}(\boldsymbol{A} + \boldsymbol{B})$. Then, for every $n \in \mathbb{N}$, $(\boldsymbol{x}, -\gamma_n\boldsymbol{U}_n\boldsymbol{B}\boldsymbol{x}) \in \operatorname{gra}(\gamma_n\boldsymbol{U}_n\boldsymbol{A})$ and (7.19) yields $(\widetilde{\boldsymbol{p}}_n, \widetilde{\boldsymbol{y}}_n - \widetilde{\boldsymbol{p}}_n) \in \operatorname{gra}(\gamma_n\boldsymbol{U}_n\boldsymbol{A})$. Hence, by monotonicity of $\boldsymbol{U}_n\boldsymbol{A}$ with respect to the scalar product $\langle\langle\langle\cdot \mid \cdot\rangle\rangle\rangle_{\boldsymbol{U}_n^{-1}}$, we have $\langle\langle\langle\widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \widetilde{\boldsymbol{p}}_n - \widetilde{\boldsymbol{y}}_n - \gamma_n\boldsymbol{U}_n\boldsymbol{B}\boldsymbol{x}\rangle\rangle\rangle_{\boldsymbol{U}_n^{-1}} \leqslant 0$. Moreover, by monotonicity of $\boldsymbol{U}_n\boldsymbol{B}$ with respect to the scalar product $\langle\langle\langle\cdot \mid \cdot\rangle\rangle\rangle_{\boldsymbol{U}_n^{-1}}$, we also have $\langle\langle\langle\widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \gamma_n\boldsymbol{U}_n\boldsymbol{B}\boldsymbol{x} - \gamma_n\boldsymbol{U}_n\boldsymbol{B}\widetilde{\boldsymbol{p}}_n\rangle\rangle\rangle_{\boldsymbol{U}_n^{-1}} \leqslant 0$. By adding the last two inequalities, we obtain

$$(\forall n \in \mathbb{N}) \quad \langle\langle\langle\widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \widetilde{\boldsymbol{p}}_n - \widetilde{\boldsymbol{y}}_n - \gamma_n\boldsymbol{U}_n\boldsymbol{B}\widetilde{\boldsymbol{p}}_n\rangle\rangle\rangle_{\boldsymbol{U}_n^{-1}} \leqslant 0. \tag{7.23}$$

In turn, we derive from (7.19) that

$$\begin{aligned} (\forall n \in \mathbb{N}) \quad & 2\gamma_n\langle\langle\langle\widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \boldsymbol{U}_n\boldsymbol{B}\boldsymbol{x}_n - \boldsymbol{U}_n\boldsymbol{B}\widetilde{\boldsymbol{p}}_n\rangle\rangle\rangle_{\boldsymbol{U}_n^{-1}} \\ & = 2\langle\langle\langle\widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \widetilde{\boldsymbol{p}}_n - \widetilde{\boldsymbol{y}}_n - \gamma_n\boldsymbol{U}_n\boldsymbol{B}\widetilde{\boldsymbol{p}}_n\rangle\rangle\rangle_{\boldsymbol{U}_n^{-1}} \\ & \quad + 2\langle\langle\langle\widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \gamma_n\boldsymbol{U}_n\boldsymbol{B}\boldsymbol{x}_n + \widetilde{\boldsymbol{y}}_n - \widetilde{\boldsymbol{p}}_n\rangle\rangle\rangle_{\boldsymbol{U}_n^{-1}} \\ & \leqslant 2\langle\langle\langle\widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \gamma_n\boldsymbol{U}_n\boldsymbol{B}\boldsymbol{x}_n + \widetilde{\boldsymbol{y}}_n - \widetilde{\boldsymbol{p}}_n\rangle\rangle\rangle_{\boldsymbol{U}_n^{-1}} \\ & = 2\langle\langle\langle\widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n\rangle\rangle\rangle_{\boldsymbol{U}_n^{-1}} \\ & = |||\boldsymbol{x}_n - \boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}} - |||\widetilde{\boldsymbol{p}}_n - \boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}} - |||\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n|||^2_{\boldsymbol{U}_n^{-1}}. \end{aligned} \tag{7.24}$$

Hence, using (7.19), (7.24), the β-Lipschitz continuity of $\boldsymbol{B}$, (7.15), and [10, Lemma 2.1(ii)], for every $n \in \mathbb{N}$, we obtain

$$\begin{aligned}
||||\widetilde{\boldsymbol{x}}_{n+1} - \boldsymbol{x}||||^2_{\boldsymbol{U}_n^{-1}} &= ||||\widetilde{\boldsymbol{q}}_n + \boldsymbol{x}_n - \widetilde{\boldsymbol{y}}_n - \boldsymbol{x}||||^2_{\boldsymbol{U}_n^{-1}} \\
&= ||||(\widetilde{\boldsymbol{p}}_n - \boldsymbol{x}) + \gamma_n \boldsymbol{U}_n(\boldsymbol{B}\boldsymbol{x}_n - \boldsymbol{B}\widetilde{\boldsymbol{p}}_n)||||^2_{\boldsymbol{U}_n^{-1}} \\
&= ||||\widetilde{\boldsymbol{p}}_n - \boldsymbol{x}||||^2_{\boldsymbol{U}_n^{-1}} + 2\gamma_n \langle\langle\langle \widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \boldsymbol{B}\boldsymbol{x}_n - \boldsymbol{B}\widetilde{\boldsymbol{p}}_n \rangle\rangle\rangle \\
&\qquad + \gamma_n^2 ||||\boldsymbol{U}_n(\boldsymbol{B}\boldsymbol{x}_n - \boldsymbol{B}\widetilde{\boldsymbol{p}}_n)||||^2_{\boldsymbol{U}_n^{-1}} \\
&\leqslant ||||\boldsymbol{x}_n - \boldsymbol{x}||||^2_{\boldsymbol{U}_n^{-1}} - ||||\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n||||^2_{\boldsymbol{U}_n^{-1}} \\
&\qquad + \gamma_n^2 \mu \beta^2 ||||\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n||||^2 \\
&\leqslant ||||\boldsymbol{x}_n - \boldsymbol{x}||||^2_{\boldsymbol{U}_n^{-1}} - \mu^{-1} ||||\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n||||^2 \\
&\qquad + \gamma_n^2 \mu \beta^2 ||||\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n||||^2.
\end{aligned} \tag{7.25}$$

Hence, it follows from (7.15) and [10, Lemma 2.1(i)] that

$$\begin{aligned}
(\forall n \in \mathbb{N}) \quad ||||\widetilde{\boldsymbol{x}}_{n+1} - \boldsymbol{x}||||^2_{\boldsymbol{U}_{n+1}^{-1}} &\leqslant (1+\eta_n) ||||\boldsymbol{x}_n - \boldsymbol{x}||||^2_{\boldsymbol{U}_n^{-1}} \\
&\qquad - \mu^{-1}(1 - \gamma_n^2 \beta^2 \mu^2) ||||\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n||||^2.
\end{aligned} \tag{7.26}$$

Consequently,

$$(\forall n \in \mathbb{N}) \quad ||||\widetilde{\boldsymbol{x}}_{n+1} - \boldsymbol{x}||||_{\boldsymbol{U}_{n+1}^{-1}} \leqslant (1+\eta_n) ||||\boldsymbol{x}_n - \boldsymbol{x}||||_{\boldsymbol{U}_n^{-1}}. \tag{7.27}$$

For every $n \in \mathbb{N}$, set

$$\varepsilon_n = \sqrt{\mu\alpha^{-1}} \Big(2\big(||||\boldsymbol{b}_n||||_{\boldsymbol{U}_n^{-1}} + (\beta\mu)^{-1} ||||\boldsymbol{a}_n||||_{\boldsymbol{U}_n} \big) + (\beta\mu)^{-1} ||||\boldsymbol{c}_n||||_{\boldsymbol{U}_n} + (\beta\mu)^{-1} ||||\boldsymbol{a}_n||||_{\boldsymbol{U}_n} \Big). \tag{7.28}$$

Then $(\varepsilon_n)_{n\in\mathbb{N}}$ is summable by (7.17) and we derive from [10, Lemma 2.1(ii)(iii)], and (7.22) that

$$\begin{aligned}
(\forall n \in \mathbb{N}) \quad ||||\boldsymbol{e}_n||||_{\boldsymbol{U}_{n+1}^{-1}} &= ||||\widetilde{\boldsymbol{x}}_{n+1} - \boldsymbol{x}_{n+1}||||_{\boldsymbol{U}_{n+1}^{-1}} \\
&\leqslant \sqrt{\alpha^{-1}} ||||\widetilde{\boldsymbol{x}}_{n+1} - \boldsymbol{x}_{n+1}|||| \\
&\leqslant \sqrt{\mu\alpha^{-1}} ||||\widetilde{\boldsymbol{x}}_{n+1} - \boldsymbol{x}_{n+1}||||_{\boldsymbol{U}_n^{-1}} \\
&\leqslant \sqrt{\mu\alpha^{-1}} (||||\widetilde{\boldsymbol{y}}_n - \boldsymbol{y}_n||||_{\boldsymbol{U}_n^{-1}} + ||||\widetilde{\boldsymbol{q}}_n - \boldsymbol{q}_n||||_{\boldsymbol{U}_n^{-1}}) \\
&\leqslant \varepsilon_n.
\end{aligned} \tag{7.29}$$

In turn, we derive from (7.27) that

$$\begin{aligned}
(\forall n \in \mathbb{N}) \quad ||||\boldsymbol{x}_{n+1} - \boldsymbol{x}||||_{\boldsymbol{U}_{n+1}^{-1}} &\leqslant ||||\widetilde{\boldsymbol{x}}_{n+1} - \boldsymbol{x}||||_{\boldsymbol{U}_{n+1}^{-1}} + ||||\widetilde{\boldsymbol{x}}_{n+1} - \boldsymbol{x}_{n+1}||||_{\boldsymbol{U}_{n+1}^{-1}} \\
&\leqslant ||||\widetilde{\boldsymbol{x}}_{n+1} - \boldsymbol{x}||||_{\boldsymbol{U}_{n+1}^{-1}} + \varepsilon_n \\
&\leqslant (1+\eta_n) ||||\boldsymbol{x}_n - \boldsymbol{x}||||_{\boldsymbol{U}_n^{-1}} + \varepsilon_n.
\end{aligned} \tag{7.30}$$

This shows that $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ is $|\cdot|$–quasi-Fejér monotone with respect to the target set $\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$ relative to $(\boldsymbol{U}_n^{-1})_{n\in\mathbb{N}}$. Moreover, by [10, Proposition 3.2], $(|||\boldsymbol{x}_n-\boldsymbol{x}|||_{\boldsymbol{U}_n^{-1}})_{n\in\mathbb{N}}$ is bounded. In turn, since $\boldsymbol{B}$ and $(J_{\gamma_n\boldsymbol{U}_n\boldsymbol{A}})_{n\in\mathbb{N}}$ are Lipschitzian, and $(\forall n\in\mathbb{N})$ $\boldsymbol{x}=J_{\gamma_n\boldsymbol{U}_n\boldsymbol{A}}(\boldsymbol{x}-\gamma_n\boldsymbol{U}_n\boldsymbol{B}\boldsymbol{x})$, we deduce from (7.19) that $(\widetilde{\boldsymbol{y}}_n)_{n\in\mathbb{N}}$, $(\widetilde{\boldsymbol{p}}_n)_{n\in\mathbb{N}}$, and $(\widetilde{\boldsymbol{q}}_n)_{n\in\mathbb{N}}$ are bounded. Therefore,

$$\tau=\sup_{n\in\mathbb{N}}\{|||\boldsymbol{x}_n-\widetilde{\boldsymbol{y}}_n+\widetilde{\boldsymbol{q}}_n-\boldsymbol{x}|||_{\boldsymbol{U}_n^{-1}},|||\boldsymbol{x}_n-\boldsymbol{x}|||_{\boldsymbol{U}_n^{-1}},1+\eta_n\}<+\infty. \tag{7.31}$$

Hence, using (7.19), Cauchy-Schwarz for the norms $(|||\cdot|||_{\boldsymbol{U}_n^{-1}})_{n\in\mathbb{N}}$, and (7.25), we get, for every $n\in\mathbb{N}$,

$$\begin{aligned}|||\boldsymbol{x}_{n+1}-\boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}}&=|||\boldsymbol{x}_n-\boldsymbol{y}_n+\boldsymbol{q}_n-\boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}}\\&=|||\widetilde{\boldsymbol{q}}_n+\boldsymbol{x}_n-\widetilde{\boldsymbol{y}}_n-\boldsymbol{x}+\boldsymbol{d}_n|||^2_{\boldsymbol{U}_n^{-1}}\\&\leqslant|||\widetilde{\boldsymbol{q}}_n+\boldsymbol{x}_n-\widetilde{\boldsymbol{y}}_n-\boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}}+2\tau|||\boldsymbol{d}_n|||_{\boldsymbol{U}_n^{-1}}+|||\boldsymbol{d}_n|||^2_{\boldsymbol{U}_n^{-1}}\\&\leqslant|||\boldsymbol{x}_n-\boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}}-\mu^{-1}(1-\gamma_n^2\beta^2\mu^2)|||\boldsymbol{x}_n-\widetilde{\boldsymbol{p}}_n|||^2+\varepsilon_{1,n},\end{aligned} \tag{7.32}$$

where $(\forall n\in\mathbb{N})$ $\varepsilon_{1,n}=2\tau|||\boldsymbol{d}_n|||_{\boldsymbol{U}_n^{-1}}+|||\boldsymbol{d}_n|||^2_{\boldsymbol{U}_n^{-1}}$. In turn, for every $n\in\mathbb{N}$, by (7.15) and [10, Lemma 2.1(i)],

$$\begin{aligned}|||\boldsymbol{x}_{n+1}-\boldsymbol{x}|||^2_{\boldsymbol{U}_{n+1}^{-1}}&\leqslant(1+\eta_n)|||\boldsymbol{x}_{n+1}-\boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}}\\&\leqslant|||\boldsymbol{x}_n-\boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}}-\mu^{-1}(1-\gamma_n^2\beta^2\mu^2)|||\boldsymbol{x}_n-\widetilde{\boldsymbol{p}}_n|||^2\\&\quad+\tau\varepsilon_{1,n}+\tau^2\eta_n.\end{aligned} \tag{7.33}$$

Since $(\tau\varepsilon_{1,n}+\tau^2\eta_n)_{n\in\mathbb{N}}\in\ell_+^1(\mathbb{N})$ by (7.22), it follows from [7, Lemma 3.1] that

$$\sum_{n\in\mathbb{N}}|||\boldsymbol{x}_n-\widetilde{\boldsymbol{p}}_n|||^2<+\infty. \tag{7.34}$$

(i) : It follows from (7.34) and (7.22) that

$$\sum_{n\in\mathbb{N}}|||\boldsymbol{x}_n-\boldsymbol{p}_n|||^2\leqslant2\sum_{n\in\mathbb{N}}|||\boldsymbol{x}_n-\widetilde{\boldsymbol{p}}_n|||^2+2\sum_{n\in\mathbb{N}}|||\boldsymbol{p}_n-\widetilde{\boldsymbol{p}}_n|||^2<+\infty. \tag{7.35}$$

Furthermore, we derive from (7.22) and (7.19) that

$$\begin{aligned}\sum_{n\in\mathbb{N}}|||\boldsymbol{y}_n-\boldsymbol{q}_n|||^2&=\sum_{n\in\mathbb{N}}|||\widetilde{\boldsymbol{q}}_n-\widetilde{\boldsymbol{y}}_n+\boldsymbol{d}_n|||^2\\&=\sum_{n\in\mathbb{N}}|||\widetilde{\boldsymbol{p}}_n-\boldsymbol{x}_n+\gamma_n\boldsymbol{U}_n(\boldsymbol{B}\boldsymbol{x}_n-\boldsymbol{B}\widetilde{\boldsymbol{p}}_n)+\boldsymbol{d}_n|||^2\\&\leqslant3\Big(\sum_{n\in\mathbb{N}}|||\boldsymbol{x}_n-\widetilde{\boldsymbol{p}}_n|||^2+|||\gamma_n\boldsymbol{U}_n(\boldsymbol{B}\boldsymbol{x}_n-\boldsymbol{B}\widetilde{\boldsymbol{p}}_n)|||^2+|||\boldsymbol{d}_n|||^2\Big)\\&<+\infty.\end{aligned} \tag{7.36}$$

(ii) : Let $\boldsymbol{x}$ be a weak cluster point of $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$. Then there exists a subsequence $(\boldsymbol{x}_{k_n})_{n\in\mathbb{N}}$ that converges weakly to $\boldsymbol{x}$. Therefore $\widetilde{\boldsymbol{p}}_{k_n} \rightharpoonup \boldsymbol{x}$ by (7.34). Furthermore, it follows from (7.19) that $\boldsymbol{u}_{k_n} \to 0$. Hence, since $(\forall n \in \mathbb{N})$ $(\widetilde{\boldsymbol{p}}_{k_n}, \boldsymbol{u}_{k_n}) \in \operatorname{gra}(\boldsymbol{A}+\boldsymbol{B})$, we obtain, $\boldsymbol{x} \in \operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$ [2, Proposition 20.33(ii)]. Altogether, it follows [10, Lemma 2.3(ii)] and [10, Theorem 3.3] that $\boldsymbol{x}_n \rightharpoonup \overline{\boldsymbol{x}}$ and hence that $\boldsymbol{p}_n \rightharpoonup \overline{\boldsymbol{x}}$ by (i).

(iii)(a) : Since $\boldsymbol{A}$ and $\boldsymbol{B}$ are maximally monotone and $\operatorname{dom}\boldsymbol{B} = \boldsymbol{\mathcal{K}}$, $\boldsymbol{A}+\boldsymbol{B}$ is maximally monotone [2, Corollary 24.4(i)], $\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$ is therefore closed [2, Proposition 23.39]. Hence, the claims follow from (i), (7.30), and [10, Proposition 3.4].

(iii)(b) : By (i), $\boldsymbol{x}_n \rightharpoonup \overline{\boldsymbol{x}}$, and hence (7.34) implies that $\widetilde{\boldsymbol{p}}_n \rightharpoonup \overline{\boldsymbol{x}}$. Furthermore, it follows from (7.19) that $\boldsymbol{u}_n \to 0$. Hence, since $(\forall n \in \mathbb{N})$ $(\widetilde{\boldsymbol{p}}_n, \boldsymbol{u}_n) \in \operatorname{gra}(\boldsymbol{A}+\boldsymbol{B})$ and since $\boldsymbol{A}+\boldsymbol{B}$ is demiregular at $\overline{\boldsymbol{x}}$, by [1, Definition 2.3], $\widetilde{\boldsymbol{p}}_n \to \overline{\boldsymbol{x}}$, and therefore (7.34) implies that $\boldsymbol{x}_n \to \overline{\boldsymbol{x}}$.

(iii)(c) : If $\boldsymbol{A}$ or $\boldsymbol{B}$ is uniformly monotone at $\overline{\boldsymbol{x}}$, then $\boldsymbol{A}+\boldsymbol{B}$ is uniformly monotone at $\overline{x}$. Therefore, the result follows from [1, Proposition 2.4(i)].

(iii)(d) : Suppose that $\boldsymbol{z} \in \operatorname{int}\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$ and fix $\rho \in \left]0,+\infty\right[$ such that $B(\boldsymbol{z};\rho) \subset \operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$. It follows from (7.30) and [10, Proposition 3.2] that

$$\varepsilon = \sup_{\boldsymbol{x}\in B(\boldsymbol{z};\rho)} \sup_{n\in\mathbb{N}} |||\boldsymbol{x}_n - \boldsymbol{x}|||_{\boldsymbol{U}_n^{-1}} \leqslant (1/\sqrt{\alpha})\Big(\sup_{n\in\mathbb{N}} |||\boldsymbol{x}_n - \boldsymbol{z}||| + \sup_{\boldsymbol{x}\in B(\boldsymbol{z};\rho)} |||\boldsymbol{x}-\boldsymbol{z}|||\Big) < +\infty \tag{7.37}$$

and from (7.30) that

$$(\forall n \in \mathbb{N})(\forall \boldsymbol{x} \in B(\boldsymbol{z};\rho))\ |||\boldsymbol{x}_{n+1} - \boldsymbol{x}|||^2_{\boldsymbol{U}_{n+1}^{-1}} \leqslant |||\boldsymbol{x}_n - \boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}} + 2\varepsilon(\varepsilon\eta_n + \varepsilon_n) + (\varepsilon\eta_n + \varepsilon_n)^2. \tag{7.38}$$

Hence, the claim follows from (i), [10, Lemma 2.1], and [10, Proposition 4.3]. □

Remark 7.5 Here are some remarks.

(i) In the case when $(\forall n \in \mathbb{N})$ $\boldsymbol{U}_n = \mathrm{Id}$, the standard forward-backward-forward splitting algorithm (7.16) reduces to algorithm proposed in [4, Eq. (2.3)], which was proposed initially in the error-free setting in [19].

(ii) An alternative variable metric splitting algorithm proposed in [14] can be used to find a zero of the sum of a maximally monotone operator $\boldsymbol{A}$ and a Lipschitzian monotone operator $\boldsymbol{B}$ in instance when $\boldsymbol{\mathcal{K}}$ is finite-dimensional. This algorithm uses a different error model and involves more iteration-dependent variables than (7.16).

Example 7.6 Let $\boldsymbol{f}\colon \boldsymbol{\mathcal{K}} \to [-\infty,+\infty]$ be a proper lower semicontinuous convex function, let $\alpha \in \left]0,+\infty\right[$, let $\beta \in \left]0,+\infty\right[$, let $\boldsymbol{B}\colon \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}}$ be a monotone and β-Lipschitzian operator, let $(\eta_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N})$, and let $(\boldsymbol{U}_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\boldsymbol{\mathcal{K}})$ that satisfies

(7.15). Furthermore, let $\boldsymbol{x}_0 \in \boldsymbol{\mathcal{K}}$, let $\varepsilon \in]0, \min\{1, 1/(\mu\beta+1)\}[$, where μ is defined as in (7.15), let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, (1-\varepsilon)/(\beta\mu)]$. Suppose that the variational inequality

$$\text{find} \quad \bar{\boldsymbol{x}} \in \boldsymbol{\mathcal{K}} \quad \text{such that} \quad (\forall \boldsymbol{y} \in \boldsymbol{\mathcal{K}}) \quad \langle \bar{\boldsymbol{x}} - \boldsymbol{y} \mid \boldsymbol{B}\bar{\boldsymbol{x}} \rangle + \boldsymbol{f}(\bar{\boldsymbol{x}}) \leqslant \boldsymbol{f}(\boldsymbol{y}) \tag{7.39}$$

admits at least one solution and set

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} \boldsymbol{y}_n = \boldsymbol{x}_n - \gamma_n \boldsymbol{U}_n \boldsymbol{B}\boldsymbol{x}_n \\ \boldsymbol{p}_n = \arg\min\limits_{\boldsymbol{x}\in\boldsymbol{\mathcal{K}}} \big(\boldsymbol{f}(\boldsymbol{x}) + \frac{1}{2\gamma_n}||||\boldsymbol{x} - \boldsymbol{y}_n||||^2_{\boldsymbol{U}_n^{-1}}\big) \\ \boldsymbol{q}_n = \boldsymbol{p}_n - \gamma_n \boldsymbol{U}_n \boldsymbol{B}\boldsymbol{p}_n \\ \boldsymbol{x}_{n+1} = \boldsymbol{x}_n - \boldsymbol{y}_n + \boldsymbol{q}_n. \end{array} \right. \tag{7.40}$$

Then $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ converges weakly to a solution $\bar{\boldsymbol{x}}$ to (7.39).

Proof. Set $\boldsymbol{A} = \partial \boldsymbol{f}$ and $(\forall n \in \mathbb{N})$ $\boldsymbol{a}_n = 0, \boldsymbol{b}_n = 0, \boldsymbol{c}_n = 0$ in Theorem 7.4(ii). □

7.2.4 Monotone inclusions involving Lipschitzian operators

The applications of the forward-backward-forward splitting algorithm considered in [4, 9, 19] can be extended to a variable metric setting using Theorem 7.4. As an illustration, we present a variable metric version of the algorithm proposed in [9, Eq. (3.1)]. Recall that the parallel sum of $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ and $B\colon \mathcal{H} \to 2^{\mathcal{H}}$ is [2]

$$A \,\square\, B = (A^{-1} + B^{-1})^{-1}. \tag{7.41}$$

Problem 7.7 Let $\mathcal{H}$ be a real Hilbert space, let m be a strictly positive integer, let $z \in \mathcal{H}$, let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone operator, let $C\colon \mathcal{H} \to \mathcal{H}$ be monotone and μ-Lipschitzian for some $\mu \in]0, +\infty[$. For every $i \in \{1, \ldots, m\}$, let $\mathcal{G}_i$ be a real Hilbert space, let $r_i \in \mathcal{G}_i$, let $B_i\colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ be maximally monotone operator, let $D_i\colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ be monotone and such that D_i^{-1} is ν_i-Lipschitzian for some $\nu_i \in]0, +\infty[$, and let $L_i\colon \mathcal{H} \to \mathcal{G}_i$ is a nonzero bounded linear operator. Suppose that

$$z \in \operatorname{ran}\bigg(A + \sum_{i=1}^{m} L_i^*\big((B_i \,\square\, D_i)(L_i \cdot - r_i)\big) + C\bigg). \tag{7.42}$$

The problem is to solve the primal inclusion

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } z \in A\overline{x} + \sum_{i=1}^{m} L_i^*\big((B_i \,\square\, D_i)(L_i\overline{x} - r_i)\big) + C\overline{x}, \tag{7.43}$$

and the dual inclusion find $\overline{v}_1 \in \mathcal{G}_1, \ldots, \overline{v}_m \in \mathcal{G}_m$ such that

$$(\exists x \in \mathcal{H}) \begin{cases} z - \sum_{i=1}^{m} L_i^*\overline{v}_i \in Ax + Cx, \\ (\forall i \in \{1, \ldots, m\})\ \overline{v}_i \in (B_i \,\square\, D_i)(L_i x - r_i). \end{cases} \tag{7.44}$$

As shown in [9], Problem 7.7 covers a wide class of problems in nonlinear analysis and convex optimization problems. However, the algorithm in [9, Theorem 3.1] is studied in the context of a fixed metric. The following result extends this result to a variable metric setting.

Corollary 7.8 *Let α be in $]0,+\infty[$, let $(\eta_{0,n})_{n\in\mathbb{N}}$ be a sequence in $\ell_+^1(\mathbb{N})$, let $(U_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{H})$, and for every $i \in \{1,\ldots,m\}$, let $(\eta_{i,n})_{n\in\mathbb{N}}$ be a sequence in $\ell_+^1(\mathbb{N})$, let $(U_{i,n})_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{G}_i)$ such that $\mu = \sup_{n\in\mathbb{N}}\{\|U_n\|, \|U_{1,n}\|,\ldots,\|U_{m,n}\|\} < +\infty$ and*

$$(\forall n \in \mathbb{N}) \quad (1+\eta_{0,n})U_{n+1} \succcurlyeq U_n, \quad \text{and} \quad (\forall i \in \{1,\ldots,m\}) \quad (1+\eta_{i,n})U_{i,n+1} \succcurlyeq U_{i,n}. \tag{7.45}$$

Let $(a_{1,n})_{n\in\mathbb{N}}$, $(b_{1,n})_{n\in\mathbb{N}}$, and $(c_{1,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{H}$, and for every $i \in \{1,\ldots,m\}$, let $(a_{2,i,n})_{n\in\mathbb{N}}$, $(b_{2,i,n})_{n\in\mathbb{N}}$, and $(c_{2,i,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{G}_i$. Furthermore, set

$$\beta = \max\{\nu_0,\nu_1,\ldots,\nu_m\} + \sqrt{\sum_{i=1}^m \|L_i\|^2}, \tag{7.46}$$

let $x_0 \in \mathcal{H}$, let $(v_{1,0},\ldots,v_{m,0}) \in \mathcal{G}_1 \oplus \ldots \oplus \mathcal{G}_m$, let $\varepsilon \in]0, 1/(1+\beta\mu)[$, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon,(1-\varepsilon)/(\beta\mu)]$. Set

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} y_{1,n} = x_n - \gamma_n U_n\big(Cx_n + \sum_{i=1}^m L_i^* v_{i,n} + a_{1,n}\big) \\ p_{1,n} = J_{\gamma_n U_n A}(y_{1,n} + \gamma_n U_n z) + b_{1,n} \\ \text{for } i = 1,\ldots,m \\ \left\lfloor \begin{array}{l} y_{2,i,n} = v_{i,n} + \gamma_n U_{i,n}\big(L_i x_n - D_i^{-1} v_{i,n} + a_{2,i,n}\big) \\ p_{2,i,n} = J_{\gamma_n U_{i,n} B_i^{-1}}(y_{2,i,n} - \gamma_n U_{i,n} r_i) + b_{2,i,n} \\ q_{2,i,n} = p_{2,i,n} + \gamma_n U_{i,n}\big(L_i p_{1,n} - D_i^{-1} p_{2,i,n} + c_{2,i,n}\big) \\ v_{i,n+1} = v_{i,n} - y_{2,i,n} + q_{2,i,n} \end{array}\right. \\ q_{1,n} = p_{1,n} - \gamma_n U_n\big(Cp_{1,n} + \sum_{i=1}^m L_i^* p_{2,i,n} + c_{1,n}\big) \\ x_{n+1} = x_n - y_{1,n} + q_{1,n}. \end{array}\right. \tag{7.47}$$

Then the following hold.

(i) $\sum_{n\in\mathbb{N}} \|x_n - p_{1,n}\|^2 < +\infty$ *and* $(\forall i \in \{1,\ldots,m\})$ $\sum_{n\in\mathbb{N}} \|v_{i,n} - p_{2,i,n}\|^2 < +\infty$.

(ii) *There exist a solution $\overline{x}$ to (7.43) and a solution $(\overline{v}_1,\ldots,\overline{v}_m)$ to (7.44) such that the following hold.*

  (a) $x_n \rightharpoonup \overline{x}$ *and* $p_{1,n} \rightharpoonup \overline{x}$.

  (b) $(\forall i \in \{1,\ldots,m\})$ $v_{i,n} \rightharpoonup \overline{v}_i$ *and* $p_{2,i,n} \rightharpoonup \overline{v}_i$.

  (c) *Suppose that A or C is uniformly monotone at $\overline{x}$, then $x_n \to \overline{x}$ and $p_{1,n} \to \overline{x}$.*

  (d) *Suppose that B_j^{-1} or D_j^{-1} is uniformly monotone at $\overline{v}_j$, for some $j \in \{1,\ldots,m\}$, then $v_{j,n} \to \overline{v}_j$ and $p_{2,j,n} \to \overline{v}_j$.*

Proof. All sequences generated by algorithm (7.47) are well defined by [11, Lemma 3.7]. We define $\boldsymbol{\mathcal{K}} = \mathcal{H} \oplus \mathcal{G}_1 \oplus \cdots \oplus \mathcal{G}_m$ the Hilbert direct sum of the Hilbert spaces $\mathcal{H}$ and $(\mathcal{G}_i)_{1 \leqslant i \leqslant m}$, the scalar product and the associated norm of $\boldsymbol{\mathcal{K}}$ respectively defined by

$$\langle\langle\langle \cdot \mid \cdot \rangle\rangle\rangle \colon ((x, \boldsymbol{v}), (y, \boldsymbol{w})) \mapsto \langle x \mid y\rangle + \sum_{i=1}^{m} \langle v_i \mid w_i\rangle \text{ and } |||\cdot||| \colon (x, \boldsymbol{v}) \mapsto \sqrt{\|x\|^2 + \sum_{i=1}^{m} \|v_i\|^2}, \tag{7.48}$$

where $\boldsymbol{v} = (v_1, \ldots, v_m)$ and $\boldsymbol{w} = (w_1, \ldots, w_m)$ are generic elements in $\mathcal{G}_1 \oplus \cdots \oplus \mathcal{G}_m$. Set

$$\begin{cases} \boldsymbol{A} \colon \boldsymbol{\mathcal{K}} \to 2^{\boldsymbol{\mathcal{K}}} \colon (x, v_1, \ldots, v_m) \mapsto (-z + Ax) \times (r_1 + B_1^{-1} v_1) \times \ldots \times (r_m + B_m^{-1} v_m) \\ \boldsymbol{B} \colon \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}} \colon (x, v_1, \ldots, v_m) \mapsto \Big(Cx + \sum_{i=1}^{m} L_i^* v_i, D_1^{-1} v_1 - L_1 x, \ldots, D_m^{-1} v_m - L_m x \Big) \\ (\forall n \in \mathbb{N}) \quad \boldsymbol{U}_n \colon \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}} \colon (x, v_1, \ldots, v_m) \mapsto \big(U_n x, U_{1,n} v_1, \ldots U_{m,n} v_m \big). \end{cases} \tag{7.49}$$

Since $\boldsymbol{A}$ is maximally monotone [2, Propositions 20.22 and 20.23], $\boldsymbol{B}$ is monotone and β-Lipschitzian [9, Eq. (3.10)] with $\operatorname{dom} \boldsymbol{B} = \boldsymbol{\mathcal{K}}$, $\boldsymbol{A} + \boldsymbol{B}$ is maximally monotone [2, Corollary 24.24(i)]. Now set $(\forall n \in \mathbb{N})$ $\eta_n = \max\{\eta_{0,n}, \eta_{1,n}, \ldots, \eta_{m,n}\}$. Then $(\eta_n)_{n \in \mathbb{N}} \in \ell_+^1(\mathbb{N})$. Moreover, we derive from our assumptions on the sequences $(U_n)_{n \in \mathbb{N}}$ and $(U_{1,n})_{n \in \mathbb{N}}, \ldots, (U_{m,n})_{n \in \mathbb{N}}$ that

$$\mu = \sup_{n \in \mathbb{N}} \|\boldsymbol{U}_n\| < +\infty \quad \text{and} \quad (1 + \eta_n) \boldsymbol{U}_{n+1} \succcurlyeq \boldsymbol{U}_n \in \mathcal{P}_\alpha(\boldsymbol{\mathcal{K}}). \tag{7.50}$$

In addition, [2, Propositions 23.15(ii) and 23.16] yield $(\forall \gamma \in \left]0, +\infty\right[)(\forall n \in \mathbb{N})(\forall (x, v_1, \ldots, v_m) \in \boldsymbol{\mathcal{K}})$

$$J_{\gamma \boldsymbol{U}_n \boldsymbol{A}}(x, v_1, \ldots, v_m) = \Big(J_{\gamma U_n A}(x + \gamma U_n z), \big(J_{\gamma U_{i,n} B_i^{-1}}(v_i - \gamma U_{i,n} r_i) \big)_{1 \leqslant i \leqslant m} \Big). \tag{7.51}$$

It is shown in [9, Eq. (3.12)] and [9, Eq. (3.13)] that under the condition (7.42), $\operatorname{zer}(\boldsymbol{A} + \boldsymbol{B}) \neq \varnothing$. Moreover, [9, Eq. (3.21)] and [9, Eq. (3.22)] yield

$$(\overline{x}, \overline{v}_1, \ldots, \overline{v}_m) \in \operatorname{zer}(\boldsymbol{A} + \boldsymbol{B}) \Rightarrow \overline{x} \text{ solves (7.43) and } (\overline{v}_1, \ldots, \overline{v}_m) \text{ solves (7.44)}. \tag{7.52}$$

Let us next set, for every $n \in \mathbb{N}$,

$$\begin{cases} \boldsymbol{x}_n = (x_n, v_{1,n}, \ldots, v_{m,n}) \\ \boldsymbol{y}_n = (y_{1,n}, y_{2,1,n}, \ldots, y_{2,m,n}) \\ \boldsymbol{p}_n = (p_{1,n}, p_{2,1,n}, \ldots, p_{2,m,n}) \\ \boldsymbol{q}_n = (q_{1,n}, q_{2,1,n}, \ldots, q_{2,m,n}) \end{cases} \quad \text{and} \quad \begin{cases} \boldsymbol{a}_n = (a_{1,n}, a_{2,1,n}, \ldots, a_{2,m,n}) \\ \boldsymbol{b}_n = (b_{1,n}, b_{2,1,n}, \ldots, b_{2,m,n}) \\ \boldsymbol{c}_n = (c_{1,n}, c_{2,1,n}, \ldots, c_{2,m,n}). \end{cases} \tag{7.53}$$

Then our assumptions imply that

$$\sum_{n\in\mathbb{N}} ||||\boldsymbol{a}_n|||| < \infty, \quad \sum_{n\in\mathbb{N}} ||||\boldsymbol{b}_n|||| < \infty, \quad \text{and} \quad \sum_{n\in\mathbb{N}} ||||\boldsymbol{c}_n|||| < \infty. \tag{7.54}$$

Furthermore, it follows from the definition of $\boldsymbol{B}$, (7.51), and (7.53) that (7.47) can be rewritten in $\boldsymbol{\mathcal{K}}$ as

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} \boldsymbol{y}_n = \boldsymbol{x}_n - \gamma_n \boldsymbol{U}_n(\boldsymbol{B}\boldsymbol{x}_n + \boldsymbol{a}_n) \\ \boldsymbol{p}_n = J_{\gamma_n \boldsymbol{U}_n \boldsymbol{A}} \boldsymbol{y}_n + \boldsymbol{b}_n \\ \boldsymbol{q}_n = \boldsymbol{p}_n - \gamma_n \boldsymbol{U}_n(\boldsymbol{B}\boldsymbol{p}_n + \boldsymbol{c}_n) \\ \boldsymbol{x}_{n+1} = \boldsymbol{x}_n - \boldsymbol{y}_n + \boldsymbol{q}_n, \end{array} \right. \tag{7.55}$$

which is (7.16). Moreover, every specific conditions in Theorem 7.4 are satisfied.

(i) : By Theorem 7.4(i), $\sum_{n\in\mathbb{N}} ||||\boldsymbol{x}_n - \boldsymbol{p}_n||||^2 < \infty$.

(ii)(a)&(ii)(b) : These assertions follow from Theorem 7.4(ii).

(ii)(c) : Theorem 7.4(ii) shows that $(\overline{x}, \overline{v}_1, \ldots, \overline{v}_m) \in \operatorname{zer}(\boldsymbol{A} + \boldsymbol{B})$. Hence, it follows from [9, Eq (3.19)] that $(\overline{x}, \overline{v}_1, \ldots, \overline{v}_m)$ satisfies the inclusions

$$\begin{cases} -\sum_{i=1}^m L_i^* \overline{v}_i - C\overline{x} \in -z + A\overline{x} \\ (\forall i \in \{1, \ldots, m\})\ L_i\overline{x} - D_i^{-1}\overline{v}_i \in r_i + B_i^{-1}\overline{v}_i. \end{cases} \tag{7.56}$$

For every $n \in \mathbb{N}$ and every $i \in \{1, \ldots, m\}$, set

$$\begin{cases} \widetilde{y}_{1,n} = x_n - \gamma_n U_n\big(Cx_n + \sum_{i=1}^m L_i^* v_{i,n}\big) \\ \widetilde{p}_{1,n} = J_{\gamma_n U_n A}(\widetilde{y}_{1,n} + \gamma_n U_n z) \end{cases} \quad \text{and} \quad \begin{cases} \widetilde{y}_{2,i,n} = v_{i,n} + \gamma_n U_{i,n}\big(L_i x_n - D_i^{-1} v_{i,n}\big) \\ \widetilde{p}_{2,i,n} = J_{\gamma_n U_{i,n} B_i^{-1}}(\widetilde{y}_{2,i,n} - \gamma_n U_{i,n} r_i). \end{cases} \tag{7.57}$$

Then, using [11, Lemma 3.7], we get

$$\widetilde{p}_{1,n} - p_{1,n} \to 0 \quad \text{and} \quad (\forall i \in \{1, \ldots, m\}) \quad \widetilde{p}_{2,i,n} - p_{2,i,n} \to 0, \tag{7.58}$$

in turn, by (i),(ii)(a), and (ii)(b), we obtain

$$\begin{cases} \widetilde{p}_{1,n} - x_n \to 0, \quad \widetilde{p}_{1,n} \rightharpoonup \overline{x}, \\ (\forall i \in \{1, \ldots, m\}) \quad \widetilde{p}_{2,i,n} - v_{i,n} \to 0, \quad \widetilde{p}_{2,i,n} \rightharpoonup \overline{v}_i. \end{cases} \tag{7.59}$$

Furthermore, we derive from (7.57) that

$$(\forall n \in \mathbb{N}) \quad \begin{cases} \gamma_n^{-1} U_n^{-1}(x_n - \widetilde{p}_{1,n}) - \sum_{i=1}^m L_i^* v_{i,n} - Cx_n \in -z + A\widetilde{p}_{1,n} \\ (\forall i \in \{1, \ldots, m\})\ \gamma_n^{-1} U_{i,n}^{-1}(v_{i,n} - \widetilde{p}_{2,i,n}) + L_i x_n - D_i^{-1} v_{i,n} \\ \qquad\qquad\qquad \in r_i + B_i^{-1}\widetilde{p}_{2,i,n}. \end{cases} \tag{7.60}$$

Since A is uniformly monotone at $\overline{x}$, using (7.56) and (7.60), there exists an increasing function $\phi_A\colon\ [0,+\infty[\ \to [0,+\infty]$ vanishing only at 0 such that, for every $n\in\mathbb{N}$,

$$\begin{aligned}\phi_A(\|\widetilde{p}_{1,n}-\overline{x}\|) &\leqslant \left\langle \widetilde{p}_{1,n}-\overline{x} \;\middle|\; \gamma_n^{-1}U_n^{-1}(x_n-\widetilde{p}_{1,n})-\sum_{i=1}^m(L_i^*v_{i,n}-L_i^*\overline{v}_i)\right\rangle-\chi_n\\ &=\big\langle \widetilde{p}_{1,n}-\overline{x}\mid \gamma_n^{-1}U_n^{-1}(x_n-\widetilde{p}_{1,n})\big\rangle-\sum_{i=1}^m\big\langle \widetilde{p}_{1,n}-\overline{x}\mid L_i^*v_{i,n}-L_i^*\overline{v}_i\big\rangle\\ &\quad-\chi_n,\end{aligned}\tag{7.61}$$

where we denote $\big(\forall n\in\mathbb{N}\big)\ \chi_n=\langle \widetilde{p}_{1,n}-\bar{x}\mid Cx_n-C\bar{x}\rangle$. Since $(B_i^{-1})_{1\leqslant i\leqslant m}$ are monotone, for every $i\in\{1,\ldots,m\}$, we obtain

$$\begin{aligned}(\forall n\in\mathbb{N})\quad 0&\leqslant \big\langle \widetilde{p}_{2,i,n}-\overline{v}_i\mid L_ix_n+\gamma_n^{-1}U_{i,n}^{-1}(v_{i,n}-\widetilde{p}_{2,i,n})-L_i\overline{x}\big\rangle-\beta_{i,n}\\ &=\big\langle \widetilde{p}_{2,i,n}-\overline{v}_i\mid L_i(x_n-\overline{x})+\gamma_n^{-1}U_{i,n}^{-1}(v_{i,n}-\widetilde{p}_{2,i,n})\big\rangle-\beta_{i,n},\end{aligned}\tag{7.62}$$

where $\big(\forall n\in\mathbb{N}\big)\ \beta_{i,n}=\big\langle \widetilde{p}_{2,i,n}-\overline{v}_i\mid D_i^{-1}v_{i,n}-D_i^{-1}\overline{v}_i\big\rangle$. Now, adding (7.62) from $i=1$ to $i=m$ and (7.61), we obtain, for every $n\in\mathbb{N}$,

$$\begin{aligned}\phi_A(\|\widetilde{p}_{1,n}-\overline{x}\|) &\leqslant \big\langle \widetilde{p}_{1,n}-\overline{x}\mid \gamma_n^{-1}U_n^{-1}(x_n-\widetilde{p}_{1,n})\big\rangle+\left\langle \widetilde{p}_{1,n}-\overline{x}\;\middle|\;\sum_{i=1}^m L_i^*(\widetilde{p}_{2,i,n}-v_{i,n})\right\rangle\\ &\quad+\sum_{i=1}^m\big\langle \widetilde{p}_{2,i,n}-\overline{v}_i\mid L_i(x_n-\widetilde{p}_{1,n})+\gamma_n^{-1}U_{i,n}^{-1}(v_{i,n}-\widetilde{p}_{2,i,n})\big\rangle\\ &\quad-\chi_n-\sum_{i=1}^m\beta_{i,n}.\end{aligned}\tag{7.63}$$

For every $n\in\mathbb{N}$ and every $i\in\{1,\ldots,m\}$, we expand χ_n and $\beta_{i,n}$ as

$$\begin{cases}\chi_n=\langle x_n-\overline{x}\mid Cx_n-C\overline{x}\rangle+\langle \widetilde{p}_{1,n}-x_n\mid Cx_n-C\overline{x}\rangle,\\ \beta_{i,n}=\big\langle v_{i,n}-\overline{v}_i\mid D_i^{-1}v_{i,n}-D_i^{-1}\overline{v}_i\big\rangle+\big\langle \widetilde{p}_{2,i,n}-v_{i,n}\mid D_i^{-1}v_{i,n}-D_i^{-1}\overline{v}_i\big\rangle.\end{cases}\tag{7.64}$$

By monotonicity of C and $(D_i^{-1})_{1\leqslant i\leqslant m}$,

$$(\forall n\in\mathbb{N})\quad\begin{cases}\langle x_n-\overline{x}\mid Cx_n-C\overline{x}\rangle\geqslant 0,\\ (\forall i\in\{1,\ldots,m\})\ \big\langle v_{i,n}-\overline{v}_i\mid D_i^{-1}v_{i,n}-D_i^{-1}\overline{v}_i\big\rangle\geqslant 0.\end{cases}\tag{7.65}$$

Therefore, for every $n \in \mathbb{N}$, we derive from (7.64) and (7.63) that

$$\begin{aligned}
\phi_A(\|\widetilde{p}_{1,n} - \overline{x}\|) &\leqslant \phi_A(\|\widetilde{p}_{1,n} - \overline{x}\|) + \langle x_n - \overline{x} \mid Cx_n - C\overline{x}\rangle \\
&\qquad + \sum_{i=1}^{m} \big\langle v_{i,n} - \overline{v}_i \mid D_i^{-1} v_{i,n} - D_i^{-1}\overline{v}_i \big\rangle \\
&\leqslant \big\langle \widetilde{p}_{1,n} - \overline{x} \mid \gamma_n^{-1} U_n^{-1}(x_n - \widetilde{p}_{1,n})\big\rangle + \Big\langle \widetilde{p}_{1,n} - \overline{x} \;\Big|\; \sum_{i=1}^{m} L_i^*(\widetilde{p}_{2,i,n} - v_{i,n})\Big\rangle \\
&\quad + \sum_{i=1}^{m} \big\langle \widetilde{p}_{2,i,n} - \overline{v}_i \mid L_i(x_n - \widetilde{p}_{1,n}) + \gamma_n^{-1} U_{i,n}^{-1}(v_{i,n} - \widetilde{p}_{2,i,n})\big\rangle \\
&\quad - \langle \widetilde{p}_{1,n} - x_n \mid Cx_n - C\overline{x}\rangle - \sum_{i=1}^{m} \big\langle \widetilde{p}_{2,i,n} - v_{i,n} \mid D_i^{-1} v_{i,n} - D_i^{-1}\overline{v}_i \big\rangle.
\end{aligned} \tag{7.66}$$

We set

$$\zeta = \max_{1\leqslant i\leqslant m} \sup_{n\in\mathbb{N}} \{\|x_n - \overline{x}\|, \|\widetilde{p}_{1,n} - \overline{x}\|, \|v_{i,n} - \overline{v}_i\|, \|\widetilde{p}_{2,i,n} - \overline{v}_i\|\}. \tag{7.67}$$

Then it follows from (ii)(a), (ii)(b), and (7.59) that $\zeta < \infty$, and from [10, Lemma 2.1(ii)] that $(\forall n \in \mathbb{N})$ $\|\gamma_n^{-1} U_n^{-1}\| \leqslant (\varepsilon\alpha)^{-1}$ and $(\forall i \in \{1,\ldots,m\})$ $\|\gamma_n^{-1} U_{i,n}^{-1}\| \leqslant (\varepsilon\alpha)^{-1}$. Therefore, using the Cauchy-Schwarz inequality, and the Lipschitzianity of C and $(D_i^{-1})_{1\leqslant i\leqslant m}$, we derive from (7.66) that

$$\begin{aligned}
\phi_A(\|\widetilde{p}_{1,n} - \overline{x}\|) &\leqslant (\varepsilon\alpha)^{-1}\zeta\|x_n - \widetilde{p}_{1,n}\| + \zeta \sum_{i=1}^{m} \big(\|L_i\|\,\|x_n - \widetilde{p}_{1,n}\| \\
&\quad + (\varepsilon\alpha)^{-1}\|v_{i,n} - \widetilde{p}_{2,i,n}\|\big) + \zeta\Big(\sum_{i=1}^{m} \|L_i^*\|\|\widetilde{p}_{2,i,n} - v_{i,n}\| \\
&\quad + \nu_0\|\widetilde{p}_{1,n} - x_n\| + \sum_{i=1}^{m} \nu_i \|\widetilde{p}_{2,i,n} - v_{i,n}\|\Big) \\
&\to 0.
\end{aligned} \tag{7.68}$$

We deduce from (7.68) and (7.59) that $\phi_A(\|\widetilde{p}_{1,n} - \overline{x}\|) \to 0$, which implies that $\widetilde{p}_{1,n} \to \overline{x}$. In turn, $x_n \to \overline{x}$ and $p_n \to \overline{x}$. Likewise, if C is uniformly monotone at $\overline{x}$, there exists an increasing function $\phi_C\colon\ [0,+\infty[\ \to [0,+\infty]$ that vanishes only at 0 such that

$$\begin{aligned}\phi_C(\|x_n-\overline{x}\|) \leqslant (\varepsilon\alpha)^{-1}\zeta\|x_n-\widetilde{p}_{1,n}\| + \zeta\sum_{i=1}^m \big(\|L_i\|\,\|x_n-\widetilde{p}_{1,n}\| \\ + (\varepsilon\alpha)^{-1}\|v_{i,n}-\widetilde{p}_{2,i,n}\|\big) + \zeta\bigg(\sum_{i=1}^m \|L_i^*\|\|\widetilde{p}_{2,i,n}-v_{i,n}\| \\ + \nu_0\|\widetilde{p}_{1,n}-x_n\| + \sum_{i=1}^m \nu_i\|\widetilde{p}_{2,i,n}-v_{i,n}\|\bigg) \\ \to 0,\end{aligned} \tag{7.69}$$

in turn, $x_n \to \overline{x}$ and $p_n \to \overline{x}$.

(ii)(d) : Proceeding as in the proof of (ii)(c), we obtain the conclusions. □

Acknowledgement. I thank Professor Patrick L. Combettes for bringing this problem to my attention and for helpful discussions.

7.3 Bibliographie

[1] H. Attouch, L. M. Briceño-Arias, and P. L. Combettes, A parallel splitting method for coupled monotone inclusions, *SIAM J. Control Optim.*, vol. 48, pp. 3246–3270, 2010.

[2] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York, 2011.

[3] J. F. Bonnans, J. Ch. Gilbert, C. Lemaréchal, and C. A. Sagastizábal, A family of variable metric proximal methods, *Math. Programming*, vol. 68, pp. 15–47, 1995.

[4] 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.

[5] J. V. Burke and M. Qian, A variable metric proximal point algorithm for monotone operators, *SIAM J. Control Optim.,* vol. 37, pp. 353–375, 1999.

[6] J. V. Burke and M. Qian, On the superlinear convergence of the variable metric proximal point algorithm using Broyden and BFGS matrix secant updating, *Math. Program.*, vol. 88, pp. 157–181, 2000.

[7] P. L. Combettes, Quasi-Fejérian analysis of some optimization algorithms, *in Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications* (D. Butnariu, Y. Censor, and S. Reich, Eds.), pp. 115–152. New York : Elsevier, 2001.

[8] P. L. Combettes, Solving monotone inclusions via compositions of nonexpansive averaged operators, *Optimization*, vol. 53, pp. 475–504, 2004.

[9] P. L. Combettes and J.-C. Pesquet, Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators, *Set-Valued Var. Anal.*, vol. 20, pp. 307–330, 2012.

[10] P. L. Combettes and B. C. Vũ, Variable metric quasi-Fejér monotonicity, *Nonlinear Anal.,* vol. 78, pp. 17–31, 2013.

[11] P. L. Combettes and B. C. Vũ, Variable metric forward-backward splitting with applications to monotone inclusions in duality, *Optimization,* to appear.

[12] B. Mercier, *Topics in Finite Element Solution of Elliptic Problems* (Lectures on Mathematics, no. 63). Tata Institute of Fundamental Research, Bombay, 1979.

[13] C. Lemaréchal and C. Sagastizábal, Variable metric bundle methods : from conceptual to implementable forms, *Math. Program.*, vol. 76, pp. 393–410, 1997.

[14] P. A. Lotito, L. A. Parente, and M. V. Solodov, A class of variable metric decomposition methods for monotone variational inclusions, *J. Convex Anal.*, vol. 16, pp. 857–880, 2009.

[15] L. A. Parente, P. A. Lotito, and M. V. Solodov, A class of inexact variable metric proximal point algorithms, *SIAM J. Optim.*, vol. 19, pp. 240–260, 2008.

[16] L. Qi and X. Chen, A preconditioning proximal Newton's method for nondifferentiable convex optimization, *Math. Program.*, vol. 76, pp. 411–430, 1995.

[17] R. T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM J. Control Optimization*, vol. 14, pp. 877–898, 1976.

[18] P. Tseng, Applications of a splitting algorithm to decomposition in convex programming and variational inequalities, *SIAM J. Control Optim.*, vol. 29, pp. 119–138, 1991.

[19] P. Tseng, A modified forward-backward splitting method for maximal monotone mappings, *SIAM J. Control Optim.*, vol. 38, pp. 431–446, 2000.

Chapitre 8

Conclusions et perspectives

8.1 Conclusions

Nous avons proposé de nouvelles méthodes primales et primales–duales d'éclatement d'opérateurs pour résoudre divers types de problèmes d'analyse non-linéaire. En particulier, afin d'analyser dans un cadre unifié le comportement asymptotique des méthodes à métrique variable, nous avons introduit une nouvelle notion de suite quasi-fejérienne. Les résultats obtenus ont été appliqué à divers schémas itératifs de construction de zéros et d'optimisation.

8.2 Perspectives

Les résultats de la thèse suggèrent l'étude des problèmes ouverts suivants.

- Afin d'utiliser efficacement les méthodes à métrique variable proposées dans la thèse, il faut calculer des résolvantes $(\mathrm{Id}\, + UA)^{-1}$ où A est multivoque maximalement monotone, et $U \in \mathcal{P}_\alpha(\mathcal{H})$. En particulier, il serait intéressant de trouver des fonctions $f \in \Gamma_0(\mathcal{H})$ telles que on a des formules explicites de prox_f^U avec $f \in \Gamma_0(\mathcal{H})$ et $U \in \mathcal{P}_\alpha(\mathcal{H})$, i.e., des formules explicites de la solution du problème

$$\underset{x\in\mathcal{H}}{\text{minimiser}}\ f(x) + \frac{1}{2}\|x - z\|_U^2, \quad \text{où } z \in \mathcal{H}. \tag{8.1}$$

  Dans le cas où $U = \mathrm{Id}$, les formules explicites sont données dans [3]. Le choix des métriques optimales dans certains cas simples sont également à étudier, même si c'est un problème complexe en général (problèmes de pré-conditionnement en particulier).

- Les méthodes à métrique variable proposées dans la thèse sont appliquées aux problèmes d'inclusions monotones, des problèmes variationnels, des problèmes inverses, des inégalités variationnelles, des problèmes de traitement du signal, des problèmes d'admissibilité et de meilleure approximation. Des applications de ces méthodes aux problèmes d'inclusions d'évolution [1], aux équations aux dérivées partielles [4], aux équilibres de Nash [2] sont à explorer.

- Nous avons montré la convergence de quelques algorithmes à métrique variable dans les Chapitres 5, 6 et 7, mais le problème de montrer la convergence de la méthode d'éclatement de Douglas-Rachford à métrique variable est encore ouvert.
- Soient $\mathcal{H}$ un espace hilbertien réel, $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ un opérateur maximalement monotone, $(\mu,\nu) \in \left]0,+\infty\right[^2$, $C\colon \mathcal{H} \to \mathcal{H}$ un opérateur μ-cocoercif ou μ-lipschitzien monotone, $\mathcal{G}$ et $\mathcal{Y}$ des espaces hilbertiens réels, $r \in \mathcal{G}$, $B\colon \mathcal{Y} \to 2^{\mathcal{Y}}$, $D\colon \mathcal{G} \to 2^{\mathcal{G}}$ des opérateurs maximalement monotones tels que D^{-1} est ν-cocoercif ou μ-lipschitzien et monotone, et $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$, $M \in \mathcal{B}(\mathcal{G},\mathcal{Y})$. Le problème est de résoudre l'inclusion primale

$$\text{trouver } x \in \mathcal{H} \text{ tel que } z \in Ax + L^*\Big(\big((M^*BM)^{-1} + D^{-1}\big)^{-1}(Lx - r)\Big) + Cx, \tag{8.2}$$

et l'inclusion duale

$$\text{trouver } v \in \mathcal{G} \text{ tel que } -r \in (M^*BM)^{-1}v - L\big((A+C)^{-1}(z - L^*v)\big) + D^{-1}v. \tag{8.3}$$

Dans le cas où $\mathcal{G} = \mathcal{Y}$ et $M = \mathrm{Id}$, on peut utiliser la méthode du Chapitre 7 pour résoudre ce problème. Dans le cas général, il reste ouvert.

Paris, le 15 avril 2013.

8.3 Bibliographie

[1] H. Attouch, L. M. Briceño-Arias, and P. L. Combettes, A parallel splitting method for coupled monotone inclusions, *SIAM J. Control Optim.*, vol. 48, pp. 3246–3270, 2010.

[2] L. M. Briceño-Arias and P. L. Combettes, Monotone operator methods for Nash equilibria in non-potential games, in *Computational and Analytical Mathematics,* (D. Bailey, H. H. Bauschke, P. Borwein, F. Garvan, M. Théra, J. Vanderwerff, and H. Wolkowicz, eds.). Springer, New York, 2013.

[3] P. L. Combettes and J.-C. Pesquet, Proximal splitting methods in signal processing, in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering,* (H. H. Bauschke et al., eds), pp. 185–212. Springer, New York, 2011.

[4] B. Mercier, *Topics in Finite Element Solution of Elliptic Problems* (Lectures on Mathematics, no. 63). Tata Institute of Fundamental Research, Bombay, 1979.