ABSTRACT

CORNEJO, DIEGO JAVIER. Resolvent and Proximal Compositions: Theory and Applications. (Under the direction of Patrick L. Combettes.)

This dissertation analyzes the mathematical properties of parametrized versions of resolvent and proximal compositions, as well as resolvent and proximal mixtures, and explores their applications to data science. These new operations combine, respectively, set-valued operators with linear operators and functions with linear operators. First, we establish variational properties of proximal compositions and integral proximal mixtures. Second, we study parametrized resolvent compositions and investigate their connections with classical constructions for combining set-valued and linear operators. We then focus on the special case of resolvent compositions for positive linear operators. Finally, we propose a minimization model based on proximal comixtures and illustrate its use in image recovery and data analysis applications.

This dissertation was supported by the U.S. National Science Foundation grants CCF-2211123 and DMS-2513409 (PI: P. L. Combettes).

Resolvent and Proximal Compositions: Theory and Applications

by

Diego Javier Cornejo

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Mathematics

Raleigh, North Carolina

2026

APPROVED BY:

Lorena Bociu

Tien Khai Nguyen

Ryan Murray

Sara Shashaani

Patrick L. Combettes
Chair of Advisory Committee

RESUMEN

CORNEJO, DIEGO JAVIER. Composiciones de Tipo Resolvente y Proximal: Teoría y Aplicaciones. (Bajo la dirección de Patrick L. Combettes.)

Esta tesis analiza las propiedades matemáticas de las versiones parametrizadas de las composiciones de tipo resolvente y proximal, así como de las mezclas de tipo resolvente y proximal, y explora sus aplicaciones en ciencia de datos. Estas nuevas operaciones combinan, respectivamente, operadores multivaluados con operadores lineales y funciones con operadores lineales. En primer lugar, establecemos propiedades variacionales de las composiciones de tipo proximal y las mezclas de tipo proximal integral. En segundo lugar, estudiamos las composiciones parametrizadas de tipo resolvente e investigamos sus conexiones con construcciones clásicas para combinar operadores multivaluados y lineales. A continuación, nos centramos en el caso especial de las composiciones de tipo resolvente para operadores lineales positivos. Finalmente, proponemos un modelo de minimización basado en mezclas de tipo proximal e ilustramos su uso en aplicaciones de recuperación de imágenes y análisis de datos.

BIOGRAPHY

Diego J. Cornejo received his Bachelor of Science from Universidad Nacional de Ingeniería in Lima, Peru, in 2020. He received his Master of Science in Mathematics from North Carolina State University in December 2023, as part of his Ph.D. studies, which he began in August 2021, under the supervision of Professor Patrick L. Combettes.

DEDICATION

A mi futura esposa y mejor amiga, Jeannette, por su amor incondicional, por ser mi soporte en los días difíciles y por celebrar conmigo cada pequeño logro. Gracias por tu comprensión en los momentos de ausencia y por ser siempre mi refugio. Este camino fue más ligero porque lo recorrí contigo.

A mis padres, Javier Cornejo y Gina Koc, por cada sacrificio, por su respaldo y amor infinito. Han sido mi motivación y ejemplo en cada paso que he dado. Todo lo que soy se lo debo a ustedes, y mi gratitud y afecto no podrían caber en palabras.

A mis hermanos, Gianfranco y Fernando, por su cariño, consejos y apoyo. A mi hermana Gabriela, por siempre estar presente, por acompañarme en cada momento, y por las largas conversaciones que me hacían sentir cerca de casa.

A Roberto Cornejo y María Koc, que son como mis segundos padres, por su cariño y su constante presencia en mi vida.

ACKNOWLEDGMENTS

I would like to express my deepest and most sincere gratitude to Professor Patrick L. Combettes for the privilege of being his doctoral student at North Carolina State University. Your knowledge, rigor, and advising have profoundly shaped my academic development. Throughout my doctoral studies, you have been an inspiring mentor and a constant source of thoughtful guidance. Your insightful perspectives continually challenged me to think more critically, work more carefully, and strive for greater clarity and depth in my research. Beyond your guidance, I am especially grateful for your time, your openness to discussion, and your steady support during both the rewarding and challenging moments of my doctoral journey. It has been a true honor to learn from you, and I will always be grateful for your mentorship and support.

I thank Prof. Bociu, Prof. Murray, Prof. Nguyen, and Prof. Shashaani for agreeing to serve on my advisory committee.

To Minh Bùi, Javier Madariaga, Julien Mayrand, Cristóbal Vivar, Matthew Broussard, Evangelia Ftaka, Rachel Morris, Adam Pickarski, Andrew Shedlock, and Tiancheng Xue, I am grateful for having shared my graduate school journey with you and for the many valuable experiences and memories along the way.

A la Dra. Yboon García, gracias por sentar las bases de mi formación académica. Su acompañamiento y ejemplo despertaron en mí el interés por esta area de investigacion y dejaron una huella duradera en mi formación.

This dissertation was supported by the U.S. National Science Foundation grants CCF-2211123 and DMS-2513409 (PI: P. L. Combettes).

TABLE OF CONTENTS

LIST OF FIGURES

AUTHORSHIP STATEMENT

Chapter 1

Diego J. Cornejo: sole author of Chapter 1.

Chapter 2

P. L. Combettes and D. J. Cornejo, Variational analysis of proximal compositions and integral proximal mixtures, *Evolution Equations and Control Theory*, vol. 17, pp. 106–139, 2026.

Chapter 3

D. J. Cornejo, Parametrized families of resolvent compositions, *Set-Valued and Variational Analysis*, vol. 33, art. 6, 24 pp., 2025.

Chapter 4

D. J. Cornejo, Resolvent compositions for positive linear operators, *Positivity*, to appear.

Chapter 5

P. L. Combettes and D. J. Cornejo, Proximal comixture minimization models for image recovery and data analysis, submitted.

Chapter 6

Diego J. Cornejo: sole author of Chapter 6.

Use of generative artificial intelligence

No generative artificial intelligence was used to write this dissertation.

NOTATION AND DEFINITIONS

General notation

- $\mathcal{H}$, $\mathcal{G}$, $\mathcal{G}_k$, $\mathcal{K}$: Real Hilbert spaces.
- $\langle \cdot \mid \cdot \rangle_{\mathcal{H}}$: Scalar product on the real Hilbert space $\mathcal{H}$.
- $\|\cdot\|_{\mathcal{H}}$: Norm on the real Hilbert space $\mathcal{H}$.
- $\mathscr{Q}_{\mathcal{H}}$: Quadratic kernel, defined by $\mathscr{Q}_{\mathcal{H}} = \|\cdot\|_{\mathcal{H}}^2/2$.
- $\mathcal{H} \oplus \mathcal{G}$: Hilbert direct sum of $\mathcal{H}$ and $\mathcal{G}$, defined as $\mathcal{H} \times \mathcal{G}$ equipped with the scalar product

$$\big((x_1, y_1), (x_2, y_2)\big) \mapsto \langle x_1 \mid x_2 \rangle_{\mathcal{H}} + \langle y_1 \mid y_2 \rangle_{\mathcal{G}}.$$

- $\mathrm{Id}_{\mathcal{H}}$: Identity operator on $\mathcal{H}$.
- $2^{\mathcal{H}}$: Power set of $\mathcal{H}$.
- $\mathcal{B}(\mathcal{H}, \mathcal{G})$: Space of bounded linear operators from $\mathcal{H}$ to $\mathcal{G}$.
- $\mathcal{B}(\mathcal{H})$: Space of bounded linear operators from $\mathcal{H}$ to $\mathcal{H}$.
- L^*: Adjoint operator of a bounded linear operator L.
- $L^\dagger$: Generalized inverse of a bounded linear operator L.
- $\|L\|$: Norm of a bounded linear operator L.
- $\to$: Strong convergence.
- $\rightharpoonup$: Weak convergence.
- $\xrightarrow{e}$: Epi-convergence.
- $\xrightarrow{g}$: Graph-convergence.
- $(\Omega, \mathcal{F}, \mu)$: Measure space.

Notation and definitions for a function $f\colon \mathcal{H} \to [-\infty, +\infty]$

- $\operatorname{dom} f = \{x \in \mathcal{H} \mid f(x) < +\infty\}$: Domain of f.
- $\operatorname{epi} f = \{(x, \xi) \in \mathcal{H} \times \mathbb{R} \mid f(x) \leqslant \xi\}$: Epigraph of f.
- $\operatorname{Argmin} f$: Set of global minimizers of f.
- $\operatorname{argmin} f$: The unique minimizers of f.
- f is proper if $-\infty \notin f(\mathcal{H})$ and $\operatorname{dom} f \neq \varnothing$.
- f is lower semicontinuous if $\operatorname{epi} f$ is a closed subset of $\mathcal{H} \oplus \mathbb{R}$.
- f is convex if $\operatorname{epi} f$ is a convex subset of $\mathcal{H} \oplus \mathbb{R}$.

- $\Gamma_\rho(\mathcal{H}) = \big\{f\colon \mathcal{H} \to \left]-\infty,+\infty\right] \mid f + \rho\mathscr{Q}_{\mathcal{H}}$ is proper lower semicontinuous and convex$\big\}$, where $\rho \in \mathbb{R}$.
- f^*: Conjugate of f, defined by

$$f^*\colon \mathcal{H} \to [-\infty,+\infty]\colon x^* \mapsto \sup_{x\in\mathcal{H}}\big(\langle x \mid x^*\rangle_{\mathcal{H}} - f(x)\big).$$

- $\operatorname{lenv}_\gamma f$ or ${}^\gamma f$: (Lower) Moreau envelope of f of parameter $\gamma \in \left]0,+\infty\right[$, defined by

$$\operatorname{lenv}_\gamma f = {}^\gamma f\colon \mathcal{H} \to [-\infty,+\infty]\colon x \mapsto \inf_{y\in\mathcal{H}}\Big(f(y) + \frac{1}{2\gamma}\|x-y\|_{\mathcal{H}}^2\Big).$$

- $\operatorname{uenv}_\gamma f$: Upper Moreau envelope of f of parameter $\gamma \in \left]0,+\infty\right[$, defined by

$$\operatorname{uenv}_\gamma f\colon \mathcal{H} \to [-\infty,+\infty]\colon x \mapsto \sup_{y\in\mathcal{H}}\Big(f(y) - \frac{1}{2\gamma}\|x-y\|_{\mathcal{H}}^2\Big).$$

- ∂f: Subdifferential of f, defined (when f is proper) by

$$\partial f\colon \mathcal{H} \to 2^{\mathcal{H}}\colon x \mapsto \big\{x^* \in \mathcal{H} \mid (\forall y \in \mathcal{H})\ \ \langle y - x \mid x^*\rangle_{\mathcal{H}} + f(x) \leqslant f(y)\big\}.$$

- prox_f: Proximity operator of a function $f \in \Gamma_0(\mathcal{H})$, defined by

$$\operatorname{prox}_f\colon \mathcal{H} \to \mathcal{H}\colon x \mapsto \operatorname{argmin}_{y\in\mathcal{H}}\Big(f(y) + \frac{1}{2}\|x-y\|_{\mathcal{H}}^2\Big).$$

Notation and definitions for a subset $C \subset \mathcal{H}$

- ι_C: Indicator function of C, defined by

$$\iota_C\colon \mathcal{H} \to [0,+\infty]\colon x \mapsto \begin{cases} 0, & \text{if } x \in C;\\ +\infty, & \text{if } x \notin C. \end{cases}$$

- $\overline{C}$: Closure of C.
- $\operatorname{span} C$: Span of C.
- $\operatorname{cone} C$: Conical hull of C.
- $\operatorname{sri} C = \big\{x \in C \mid \operatorname{cone}(C-x) = \overline{\operatorname{span}}(C-x)\big\}$: Strong relative interior of a convex set C.
- d_C: Distance function to C, defined by

$$d_C\colon \mathcal{H} \to [0,+\infty]\colon x \mapsto \inf\|C-x\|_{\mathcal{H}}.$$

- proj_C: Projection operator onto a nonempty closed and convex set C, defined by

$$\mathrm{proj}_C = \mathrm{prox}_{\iota_C}.$$

Notation and definitions for an operator $T\colon \mathcal{H} \to \mathcal{H}$

- $\mathrm{Fix}\, T = \big\{x \in \mathcal{H} \mid Tx = x\big\}$: Set of fixed points of T.
- T is Lipschitzian with constant $\beta \in [0, +\infty[$ if

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \|Tx - Ty\|_{\mathcal{H}} \leqslant \beta \|x - y\|_{\mathcal{H}}.$$

- T is cocoercive with constant $\beta \in \,]0, +\infty[$ if

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

- T is nonexpansive if it is Lipschitzian with constant 1.
- T is firmly nonexpansive if it is cocoercive with constant 1.

Notation and definitions for a set-valued operator $A\colon \mathcal{H} \to 2^{\mathcal{H}}$

- $\mathrm{dom}\, A = \big\{x \in \mathcal{H} \mid Ax \neq \varnothing\big\}$: Domain of A.
- $\mathrm{ran}\, A = \bigcup_{x \in \mathcal{H}} Ax$: Range of A.
- $\mathrm{zer}\, A = \big\{x \in \mathcal{H} \mid 0 \in Ax\big\}$: Set of zeros of A.
- $\mathrm{gra}\, A = \big\{(x, x^*) \in \mathcal{H} \times \mathcal{H} \mid x^* \in Ax\big\}$: Graph of A.
- A^{-1}: Inverse of A, defined by

$$A^{-1}\colon \mathcal{H} \to 2^{\mathcal{H}}\colon x^* \mapsto \big\{x \in \mathcal{H} \mid x^* \in Ax\big\}.$$

- J_A: Resolvent of A, defined by

$$J_A = (\mathrm{Id}_{\mathcal{H}} + A)^{-1}.$$

- ${}^{\gamma}A$: Yosida approximation of A of index $\gamma \in \,]0, +\infty[$, defined by

$${}^{\gamma}A = \frac{1}{\gamma}(\mathrm{Id}_{\mathcal{H}} - J_{\gamma A}).$$

- A is monotone if

$$\big(\forall (x, x^*) \in \mathrm{gra}\, A\big)\big(\forall (y, y^*) \in \mathrm{gra}\, A\big) \quad \langle x - y \,|\, x^* - y^* \rangle_{\mathcal{H}} \geqslant 0.$$

- A is maximally monotone if it is monotone and there exists no monotone operator $\widetilde{A}\colon \mathcal{H} \to 2^{\mathcal{H}}$ such that gra $\widetilde{A}$ properly contains gra A.

Notation for proximal compositions with parameter $\gamma \in \left]0, +\infty\right[$

- $L \overset{\gamma}{\diamond} g$: Proximal composition of g and L.
- $L \overset{\gamma}{\blacklozenge} g$: Proximal cocomposition of g and L.
- $\overset{\diamond}{\mathsf{M}}_{\gamma}(L_k, g_k)_{k\in K}$: Proximal mixture of $(g_k)_{k\in K}$ and $(L_k)_{k\in K}$.
- $\overset{\blacklozenge}{\mathsf{M}}_{\gamma}(L_k, g_k)_{k\in K}$: Proximal comixture of $(g_k)_{k\in K}$ and $(L_k)_{k\in K}$.
- $\mathsf{pav}_{\gamma}(g_k)_{k\in K}$: Proximal average of $(g_k)_{k\in K}$.

Notation for resolvent compositions with parameter $\gamma \in \left]0, +\infty\right[$

- $L \overset{\gamma}{\diamond} B$: Resolvent composition of B and L.
- $L \overset{\gamma}{\blacklozenge} B$: Resolvent cocomposition of B and L.
- $\overset{\diamond}{\mathsf{M}}_{\gamma}(L_k, B_k)_{k\in K}$: Resolvent mixture of $(B_k)_{k\in K}$ and $(L_k)_{k\in K}$.
- $\overset{\blacklozenge}{\mathsf{M}}_{\gamma}(L_k, B_k)_{k\in K}$: Resolvent comixture of $(B_k)_{k\in K}$ and $(L_k)_{k\in K}$.
- $\mathsf{rav}_{\gamma}(B_k)_{k\in K}$: Resolvent average of $(B_k)_{k\in K}$.

Chapter 1

INTRODUCTION

1.1 Overview

Throughout, $\mathcal{H}$ is a real Hilbert space with power set $2^{\mathcal{H}}$, identity operator $\mathrm{Id}_{\mathcal{H}}$, scalar product $\langle\cdot\mid\cdot\rangle_{\mathcal{H}}$, and associated norm $\|\cdot\|_{\mathcal{H}}$. The class of proper lower semicontinuous convex functions from $\mathcal{H}$ to $]-\infty,+\infty]$ is denoted by $\Gamma_0(\mathcal{H})$. In addition, $\mathcal{G}$ is a real Hilbert space, and the space of bounded linear operators from $\mathcal{H}$ to $\mathcal{G}$ is denoted by $\mathcal{B}(\mathcal{H},\mathcal{G})$.

Given a monotone operator $A\colon\mathcal{H}\to 2^{\mathcal{H}}$, a fundamental problem in nonlinear analysis is to

$$\text{find}\quad \overline{x}\in\mathcal{H}\quad\text{such that}\quad 0\in A\overline{x}. \tag{1.1}$$

This problem arises in a wide variety of applications [28], including optimization [8, 25, 33, 44, 45, 70], minimax problems [70], signal processing [32, 36–38], image recovery [14, 15, 20, 34], machine learning [62, 80], neural networks [35, 48, 49], game theory [13, 17], and partial differential equations [52, 74, 82]. From an algorithmic perspective, a fundamental approach to solve (1.1) is the proximal point algorithm [71], which is formulated in terms of the resolvent operators of A, defined by

$$(\forall\gamma\in\,]0,+\infty[)\quad J_{\gamma A}=(\mathrm{Id}_{\mathcal{H}}+\gamma A)^{-1}. \tag{1.2}$$

More precisely, given a family of parameters $(\gamma_n)_{n\in\mathbb{N}}$ in $]0,+\infty[$, the operators $(J_{\gamma_n A})_{n\in\mathbb{N}}$ are single-valued when A is maximally monotone, and the proximal point algorithm generates a sequence $(x_n)_{n\in\mathbb{N}}$ according to

$$(\forall n\in\mathbb{N})\quad x_{n+1}=J_{\gamma_n A}x_n. \tag{1.3}$$

In many scenarios, the operator A in (1.1) can be expressed as a composition of a set-valued operator $B\colon\mathcal{G}\to 2^{\mathcal{G}}$ and a linear operator $L\in\mathcal{B}(\mathcal{H},\mathcal{G})$. The most common construction is the standard composition

$$L^* \circ B \circ L, \tag{1.4}$$

which has been studied extensively in the literature; see, for instance [2, 3, 5, 6, 8, 12, 24, 28, 33, 39, 45, 65, 67, 75, 77, 79]. Another important instance is the parallel composition

$$L^* \triangleright B = (L^* \circ B^{-1} \circ L)^{-1} \tag{1.5}$$

of B by L^*, which was introduced in [11] and further investigated in [8, 16, 28, 79]. In general, operators obtained through such composite constructions do not admit closed form expressions of their resolvent operators, which prevents the direct implementation of the proximal point algorithm (1.3), as it relies on the computability of the resolvent.

In this dissertation, we focus on two monotonicity-preserving operations, called *resolvent compositions*, recently introduced in [27], which provide a framework for implementing the proximal point algorithm in terms of the linear operator L and the resolvent of B.

Definition 1.1 ([27, Definition 1.1]) Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$, let $B\colon \mathcal{G} \to 2^{\mathcal{G}}$, and let $\gamma \in \left]0, +\infty\right[$. The *resolvent composition* of B and L with parameter γ is the operator $L \overset{\gamma}{\diamond} B\colon \mathcal{H} \to 2^{\mathcal{H}}$ given by

$$L \overset{\gamma}{\diamond} B = L^* \triangleright (B + \gamma^{-1}\mathrm{Id}_{\mathcal{G}}) - \gamma^{-1}\mathrm{Id}_{\mathcal{H}}, \tag{1.6}$$

and the *resolvent cocomposition* of B and L with parameter γ is

$$L \overset{\gamma}{\blacklozenge} B = \big(L \overset{1/\gamma}{\diamond} B^{-1}\big)^{-1}. \tag{1.7}$$

A motivation for studying resolvent compositions stems from the fact that, contrary to the standard operations (1.4) and (1.5), their resolvent operator can be computed explicitly. This feature, in turn, significantly facilitates the design and implementation of algorithms for monotone inclusion and convex optimization problems [18, 27–29, 31, 41]. Special cases can be implicitly found in concrete applications such as computed tomography [21], signal recovery [29, 31], neural networks [48, 49], inverse problems [51], and machine learning [73, 80]. More precisely, [27, Propositions 1.2 and 4.1(v)] state that

$$J_{\gamma(L \overset{\gamma}{\diamond} B)} = L^* \circ J_{\gamma B} \circ L \quad \text{and} \quad J_{\gamma(L \overset{\gamma}{\blacklozenge} B)} = \mathrm{Id}_{\mathcal{H}} - L^* \circ (\mathrm{Id}_{\mathcal{G}} - J_{\gamma B}) \circ L. \tag{1.8}$$

At the variational level, resolvent compositions of subdifferentials of functions in $\Gamma_0(\mathcal{G})$ give rise to new functions the subdifferential of which is precisely the corresponding resolvent composition.

Definition 1.2 ([27, Definition 1.4]) Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$, let $g \in \Gamma_0(\mathcal{G})$, and let $\gamma \in \left]0, +\infty\right[$. The *proximal composition* of g and L with parameter γ is

$$L \overset{\gamma}{\diamond} g = \Big({}^{\frac{1}{\gamma}}(g^*) \circ L \Big)^* - \frac{1}{2\gamma} \| \cdot \|_{\mathcal{H}}^2, \tag{1.9}$$

and the *proximal cocomposition* of g and L with parameter γ is

$$L \overset{\gamma}{\blacklozenge} g = (L \overset{1/\gamma}{\diamond} g^*)^*. \tag{1.10}$$

As established in [27, Examples 3.6(iii) and 3.10(ii)], under the assumption that $0 < \|L\| \leqslant 1$, the resolvent and proximal compositions are related via

$$L \overset{1}{\diamond} \partial g = \partial (L \overset{1}{\diamond} g) \quad \text{and} \quad L \overset{1}{\blacklozenge} \partial g = \partial (L \overset{1}{\blacklozenge} g). \tag{1.11}$$

Thanks to (1.8) and the fact that the proximity operator of a function in $\Gamma_0(\mathcal{G})$ is the resolvent of its subdifferential operator (see [8, Section 23.1]), the proximity operators of the proximal compositions are computable explicitly in terms of the proximity operator of g and L, namely,

$$\operatorname{prox}_{\gamma(L \overset{\gamma}{\diamond} g)} = L^* \circ \operatorname{prox}_{\gamma g} \circ L \quad \text{and} \quad \operatorname{prox}_{\gamma(L \overset{\gamma}{\blacklozenge} g)} = \mathrm{Id}_{\mathcal{H}} - L^* \circ (\mathrm{Id}_{\mathcal{G}} - \operatorname{prox}_{\gamma g}) \circ L. \tag{1.12}$$

A particular instance of a proximal composition is the *proximal comixture*, defined as follows.

Definition 1.3 ([27, Example 5.9]) Let K be a nonempty finite set and, for every $k \in K$, let $\mathcal{G}_k$ be a real Hilbert space, let $g_k \in \Gamma_0(\mathcal{G}_k)$, let $L_k \in \mathcal{B}(\mathcal{H}, \mathcal{G}_k)$, and let $\alpha_k \in \left]0, +\infty\right[$. The *proximal comixture* of $(g_k)_{k \in K}$ and $(L_k)_{k \in K}$ with parameter $\gamma \in \left]0, +\infty\right[$ is

$$\overset{\blacklozenge}{\mathsf{M}}_\gamma (L_k, g_k)_{k \in K} = \left(\left(\sum_{k \in K} \alpha_k \big({}^{\gamma} g_k \big) \circ L_k \right)^* - \frac{1}{2} \| \cdot \|_{\mathcal{H}}^2 \right)^*. \tag{1.13}$$

The proximal comixture operation is a generalization of the proximal average operation [9], which arises as the special case in which $\sum_{k=1}^p \alpha_k = 1$ and, for every $k \in K$, $\mathcal{G}_k = \mathcal{H}$ and $L_k = \mathrm{Id}_{\mathcal{H}}$, namely,

$$\mathsf{pav}_\gamma (g_k)_{k \in K} = \left(\left(\sum_{k \in K} \alpha_k \big({}^{\gamma} g_k \big) \right)^* - \frac{1}{2} \| \cdot \|_{\mathcal{H}}^2 \right)^*. \tag{1.14}$$

In this specific context, the benefits of using proximal averages in lieu of standard averages have been documented in [50, 62, 73, 80].

Despite progress on resolvent and proximal compositions [18, 27], the state of the art

- focuses exclusively on the static case $\gamma = 1$;
- provides limited comparisons with the standard compositions (1.4) and (1.5);
- does not study the interplay with other types of function and operator approximations;

- does not address concrete applications.

This leaves several important open questions, which we address in this thesis:

(Q1) What are the variational properties of the proximal compositions of Definition 1.2 for $\gamma \in \,]0, +\infty[$? In particular, in which way do these compositions relate to classical ones, how do they behave asymptotically as the parameter γ varies, and in what sense does this convergence occur?

(Q2) How can the resolvent compositions of Definition 1.1 for $\gamma \in \,]0, +\infty[$ be interpreted and related to the composite methods (1.4) and (1.5)? Moreover, what is the asymptotic behavior of resolvent compositions as the parameter γ varies, and in what sense does this behavior occur? These questions remain open even for the *resolvent averages* [7,10,78], which correspond to a particular instance of resolvent compositions in which, for every $k \in K$, $\mathcal{G}_k = \mathcal{H}$ and $L_k = \mathrm{Id}_{\mathcal{H}}$.

(Q3) It has been observed in [56, 59] that the *arithmetic average*, *harmonic average*, and *resolvent average* are nonexpansive in the space of self-adjoint strongly monotone operators endowed with the *Thompson metric* [76]. This property guarantees the stability of the averaging processes and is particularly useful in the study of nonlinear equations [56]. Moreover, in the finite-dimensional setting, [10] shows that the resolvent average interpolates between the arithmetic average as $\gamma \to 0$ and the harmonic average as $\gamma \to +\infty$. These observations naturally lead to the following question: are resolvent compositions nonexpansive with respect to the Thompson metric, and can they be interpreted as interpolations of the operations (1.4) and (1.5)?

(Q4) In minimization models for image recovery and data analysis, loss functions and linear operators are commonly aggregated as an average of composite terms. Each term in the aggregate models a desired property of the ideal solution, reflecting both *a priori* knowledge and observed data. When proximal comixtures of Definition 1.3 are used as an aggregation method, what benefits do they offer in terms of modeling, variational structure, and algorithmic implementation?

1.2 Contributions and organization

The main contributions of this dissertation are the following:

- In Chapter 2, we focus on **(Q1)** and establish various variational properties of the parametrized proximal compositions. We study in particular convexity, Legendre conjugacy, differentiability, Moreau envelopes, coercivity, minimizers, recession functions, and perspective functions of these constructs. We also analyze their asymptotic behavior as the parameter varies, considering both pointwise convergence and epi-convergence.
- Chapter 3 is devoted to **(Q2)**, where we provide an in-depth analysis of the parametrized resolvent compositions. We show that these compositions can be interpreted as parallel

compositions of perturbed operators. Additionally, we derive asymptotic results concerning operator convergence, with a particular focus on graph-convergence and the ρ-Hausdorff distance.

- We address **(Q3)** in Chapter 4, where we present several new results, including Löwner partial order relations, concavity, nonexpansiveness, and asymptotic behavior. These results generalize in particular the asymptotic properties established in [10, 41, 54, 56]. Furthermore, we introduce geometric interpolations of (1.4) and (1.5) and investigate their connections with resolvent compositions.
- In Chapter 5, we focus on **(Q4)** and propose an alternative minimization model based on proximal comixtures. We analyze the mathematical properties of these aggregation operations and compare them with standard composite averages. We also examine the interplay between proximal cocompositions and classical function approximations. Numerical illustrations of the benefits of this model are provided in the context of image recovery and data analysis applications.
- We conclude the dissertation in Chapter 6 with future research directions.

1.3 Publications

The findings resulting from the research carried out in the dissertation have been disseminated in the following conference article and four Q1-ranked journal articles:

1. P. L. Combettes and D. J. Cornejo, Signal recovery with proximal comixtures, *Proceedings of the European Signal Processing Conference*, pp. 2637–2641. Lyon, France, August 26–30, 2024.
2. P. L. Combettes and D. J. Cornejo, Variational analysis of proximal compositions and integral proximal mixtures, *Evolution Equations and Control Theory*, vol. 17, pp. 106–139, 2026.
3. D. J. Cornejo, Parametrized families of resolvent compositions, *Set-Valued and Variational Analysis*, vol. 33, art. 6, 24 pp., 2025.
4. D. J. Cornejo, Resolvent compositions for positive linear operators, *Positivity*, to appear.
5. P. L. Combettes and D. J. Cornejo, Proximal comixture minimization models for image recovery and data analysis, submitted.

References

[1] T. Ando, On some operator inequalities, *Math. Ann.*, vol. 279, pp. 157–159, 1987.

[2] F. J. Aragón-Artacho, R. Campoy, and M. K. Tam, Strengthened splitting methods for computing resolvents, *Comput. Optim. Appl.*, vol. 80, pp. 549–585, 2021.

[3] H. Attouch, *Variational Convergence for Functions and Operators*. Pitman, Boston, MA, 1984.

[4] H. Attouch, R. Lucchetti, and R. Wets, The topology of the ρ-Hausdorff distance, *Ann. Mat. Pura Appl.*, vol. 160, pp. 303–320, 1991.

[5] H. Attouch, A. Moudafi, and H. Riahi, Quantitative stability analysis for maximal monotone operators and semi-groups of contractions, *Nonlinear Anal.*, vol. 21, pp. 697–723, 1993.

[6] D. Azé and J. P. Penot, Operations on convergent families of sets and functions, *Optimization*, vol. 21, pp. 521–534, 1990.

[7] S. Bartz, H. H. Bauschke, S. M. Moffat, and X. Wang, The resolvent average of monotone operators: Dominant and recessive properties, *SIAM J. Optim.*, vol. 26, pp. 602–634, 2016.

[8] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed. Springer, New York, 2017.

[9] H. H. Bauschke, R. Goebel, Y. Lucet, and X. Wang, The proximal average: Basic theory, *SIAM J. Optim.*, vol. 19, pp. 766–785, 2008.

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

[11] S. R. Becker and P. L. Combettes, An algorithm for splitting parallel sums of linearly composed monotone operators, with applications to signal recovery, *J. Nonlinear Convex Anal.*, vol. 15, pp. 137–159, 2014.

[12] 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.

[13] 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 et al., Eds.), pp. 143–159. Springer, New York, 2013.

[14] L. M. Briceño-Arias, P. L. Combettes, J.-C. Pesquet, and N. Pustelnik, Proximal algorithms for multicomponent image recovery problems, *J. Math. Imaging Vision*, vol. 41, pp. 3–22, 2011.

[15] L. M. Briceño-Arias and N. Pustelnik, Infimal post-composition approach for composite convex optimization applied to image restoration, *Signal Process.*, vol. 223, art. 109549, 2024.

[16] L. M. Briceño-Arias and F. Roldán, Resolvent of the parallel composition and the proximity operator of the infimal postcomposition, *Optim. Lett.*, vol. 17, pp. 339–412, 2023.

[17] M. N. Bùi and P. L. Combettes, Analysis and numerical solution of a modular convex Nash equilibrium problem, *J. Convex Anal.*, vol. 29, pp. 1007–1021, 2022.

[18] M. N. Bùi and P. L. Combettes, Integral resolvent and proximal mixtures, *J. Optim. Theory Appl.*, vol. 203, pp. 2328–2353, 2024.

[19] M. N. Bùi and P. L. Combettes, Hilbert direct integrals of monotone operators, *Canad. J. Math.*, vol. 77, pp. 1425–1456, 2025.

[20] A. Chambolle and T. Pock, An introduction to continuous optimization for imaging, *Acta Numer.*, vol. 25, pp. 161–319, 2016.

[21] E. P. Chandler, S. Shoushtari, B. Wohlberg, and U. S. Kamilov, Closed-form approximation of the total variation proximal operator, *IEEE Trans. Comput. Imaging*, vol. 11, pp. 1217–1228, 2025.

[22] R. R. Coifman and D. L. Donoho, Translation-invariant de-noising, *Lect. Notes Stat.*, vol. 103, pp 125–150, 1995.

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

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

[25] P. L. Combettes, Monotone operator theory in convex optimization, *Math. Program.*, vol. 170, pp. 177–206, 2018.

[26] P. L. Combettes, Perspective functions: properties, constructions, and examples, *Set-Valued Var. Anal.*, vol. 26, pp. 247–264, 2018.

[27] P. L. Combettes, Resolvent and proximal compositions, *Set-Valued Var. Anal.*, vol. 31, art. 22, 29 pp., 2023.

[28] P. L. Combettes, The geometry of monotone operator splitting methods, *Acta Numer.*, vol. 33, pp. 487–632, 2024.

[29] P. L. Combettes and D. J. Cornejo, Signal recovery with proximal comixtures, *Proc. Europ. Signal Process. Conf.*, pp. 2637–2641, 2024.

[30] P. L. Combettes and D. J. Cornejo, Variational analysis of proximal compositions and integral proximal mixtures, *Evol. Equ. Control Theory*, vol. 17, pp. 106–139, 2026.

[31] P. L. Combettes and D. J. Cornejo, Proximal comixture minimization models for image recovery and data analysis, submitted.

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

[33] 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.

[34] P. L. Combettes and L. E. Glaudin, Proximal activation of smooth functions in splitting algorithms for convex image recovery, *SIAM J. Imaging Sci.*, vol. 12, pp. 1905–1935, 2019.

[35] P. L. Combettes and J.-C. Pesquet, Deep neural network structures solving variational inequalities, *Set-Valued Var. Anal.*, vol. 28, pp. 491–518, 2020.

[36] P. L. Combettes and J.-C. Pesquet, Fixed point strategies in data science, *IEEE Trans. Signal Process.*, vol. 69, pp. 3878–3905, 2021.

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

[38] P. L. Combettes and Z. C. Woodstock, A variational inequality model for the construction of signals from inconsistent nonlinear equations, *SIAM J. Imaging Sci.*, vol. 15, pp. 84–109, 2022.

[39] L. Condat, D. Kitahara, A. Contreras, and A. Hirabayashi, Proximal splitting algorithms for convex optimization: A tour of recent advances, with new twists, *SIAM Rev.*, vol. 65, pp. 375–435, 2023.

[40] G. Corach, H. Porta, and L. Recht, Convexity of the geodesic distance on spaces of positive operators, *Illinois J. Math.*, vol. 38, pp. 87–94, 1994.

[41] D. J. Cornejo, Parametrized families of resolvent compositions, *Set-Valued Var. Anal.*, vol. 33, art. 6, 24 pp., 2025.

[42] D. J. Cornejo, Resolvent compositions for positive linear operators, *Positivity*, to appear.

[43] J. Dixmier, *Les Algèbres d'Opérateurs dans l'Espace Hilbertien (Algèbres de von Neumann)*, 2e éd. Gauthier–Villars, Paris, 1969.

[44] J. Eckstein, Some saddle-function splitting methods for convex programming, *Optim. Methods Softw.,* vol. 4, pp. 75–83, 1994.

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

[46] S. Gaubert and Z. Qu, The contraction rate in Thompson's part metric of order-preserving flows on a cone – Application to generalized Riccati equations, *J. Differential Equations*, vol. 256, pp. 2902–2948, 2014.

[47] R. Glowinski and P. L. Tallec (eds.), *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics.* SIAM, Philadelphia, PA, 1989.

[48] M. Hasannasab, J. Hertrich, S. Neumayer, G. Plonka, S. Setzer, and G. Steidl, Parseval proximal neural networks, *J. Fourier Anal. Appl.*, vol. 26, art. 59, 31 pp., 2020.

[49] J. Hertrich, S. Neumayer, and G. Steidl, Convolutional proximal neural networks and plug-and-play algorithms, *Linear Algebra Appl.*, vol. 631, pp. 203–234, 2021.

[50] Z. Hu, K. Shaloudegi, G. Zhang, and Y. Yu, Federated learning meets multi-objective optimization, *IEEE Trans. Network Sci. Eng.*, vol. 9, pp. 2039–2051, 2022.

[51] U. S. Kamilov, A parallel proximal algorithm for anisotropic total variation minimization, *IEEE Trans. Image Process.*, vol. 26, pp. 539–548, 2017.

[52] A. Kaplan and R. Tichatschke, Proximal point method and elliptic regularization, *Nonlinear Anal.*, vol. 71, pp. 4525–4543, 2009.

[53] S. Kim, Mixture and interpolation of the parametrized ordered means, *J. Inequal. Appl.*, vol. 2022, art. 119, 16 pp., 2022.

[54] S. Kim, J. Lawson, and Y. Lim, The matrix geometric mean of parametrized, weighted arithmetic and harmonic means, *Linear Algebra Appl.*, vol. 435, pp. 2114–2131, 2011.

[55] F. Kubo and T. Ando, Means of positive linear operators, *Math. Ann.*, vol. 246, pp. 205–224, 1980.

[56] S. Kum and Y. Lim, Nonexpansiveness of the resolvent average, *J. Math. Anal. Appl.*, vol. 432, pp. 918–927, 2015.

[57] J. Lawson and Y. Lim, The geometric mean, matrices, metrics, and more, *Amer. Math. Monthly*, vol. 108, pp. 797–812, 2001.

[58] J. Lawson and Y. Lim, A Birkhoff contraction formula with applications to Riccati equations, *SIAM J. Control Optim.*, vol. 46, pp. 930–951, 2007.

[59] J. Lawson and Y. Lim, Karcher means and Karcher equations of positive definite operators, *Trans. Amer. Math. Soc. Ser. B*, vol. 1, pp. 1–22, 2014.

[60] H. Lee and Y. Lim, Invariant metrics, contractions and nonlinear matrix equations, *Nonlinearity*, vol. 21, pp. 857–878, 2008.

[61] B. Lemmens and M. Roelands, Unique geodesics for Thompson's metric, *Ann. Inst. Fourier (Grenoble)*, vol. 65, pp. 315–348, 2015.

[62] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, Federated optimization in heterogeneous networks, *Proc. Machine Learn. Syst.*, vol. 2, pp. 429–450, 2020.

[63] Y. Lim, Solving the nonlinear matrix equation $X = Q + \sum_{i=1}^{m} M_i X^{\delta_i} M_i^*$ via a contraction principle, *Linear Algebra Appl.*, vol. 430, pp. 1380–1383, 2009.

[64] Y. Lim and M. Pálfia, Matrix power means and the Karcher mean, *J. Funct. Anal.*, vol. 262, pp. 1498–1514, 2012.

[65] Y. Malitsky and M. K. Tam, Resolvent splitting for sums of monotone operators with minimal lifting, *Math. Program.*, vol. 201, pp. 231–262, 2023.

[66] L. Montrucchio, Thompson metric, contraction property and differentiability of policy functions, *J. Econ. Behav. Organ.*, vol. 33, pp. 449–466, 1998.

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

[68] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. I: Functional Analysis*, Academic Press, New York, 1980.

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

[70] R. T. Rockafellar, Monotone operators associated with saddle-functions and minimax problems, in: *Nonlinear Functional Analysis, Part 1*, (F. E. Browder, Ed.), pp. 241–250. AMS, Providence, RI, 1970.

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

[72] R. T. Rockafellar and R. J. B. Wets, *Variational Analysis*, 3rd printing. Springer, New York, 2009.

[73] L. Shen, W. Liu, J. Huang, Y.-G. Jiang, and S. Ma, Adaptive proximal average approximation for composite convex minimization, *Proc. 31st AAAI Conf. Artificial Intell.*, pp. 2513–2519, 2017.

[74] R. E. Showalter, *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations.* AMS, Providence, RI, 1997.

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

[76] A. C. Thompson, On certain contraction mappings in a partially ordered vector space, *Proc. Amer. Math. Soc.*, vol. 14, pp. 438–443, 1963.

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

[78] X. Wang, Self-dual regularization of monotone operators via the resolvent average, *SIAM J. Optim.*, vol. 21, pp. 438–462, 2011.

[79] F. Xue, The operator splitting schemes revisited: Primal-dual gap and degeneracy reduction by a unified analysis, *Optim. Lett.*, vol. 18, pp. 155–194, 2024.

[80] Y.-L. Yu, Better approximation and faster algorithm using the proximal average, *Proc. Conf. Adv. Neural Inform. Process. Syst.*, pp. 458–466, 2013.

[81] M. Yukawa, H. Kaneko, K. Suzuki, and I. Yamada, Linearly-involved Moreau-enhanced-over-subspace model: Debiased sparse modeling and stable outlier-robust regression, *IEEE Trans. Signal Process.*, vol. 71, pp. 1232–1247, 2023.

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

Chapter 2

VARIATIONAL ANALYSIS OF PROXIMAL COMPOSITIONS AND INTEGRAL PROXIMAL MIXTURES

2.1 Introduction and context

We address question **(Q1)** of Chapter 1 by providing variational properties of proximal compositions and integral proximal mixtures.

This chapter presents the following journal article:

> P. L. Combettes and D. J. Cornejo, Variational analysis of proximal compositions and integral proximal mixtures, *Evolution Equations and Control Theory*, vol. 17, pp. 106–139, 2026.

2.2 Article: Variational analysis of proximal compositions and integral proximal mixtures

Abstract. This paper establishes various variational properties of parametrized versions of two convexity-preserving constructs that were recently introduced in the literature: the proximal composition of a function and a linear operator, and the integral proximal mixture of arbitrary families of functions and linear operators. We study in particular convexity, Legendre conjugacy, differentiability, Moreau envelopes, coercivity, minimizers, recession functions, and perspective functions of these constructs, as well as their asymptotic behavior as the parameter varies. The special case of the proximal expectation of a family of functions is also discussed.

2.2.1 Introduction

Throughout, $\mathcal{H}$ is a real Hilbert space with power set $2^{\mathcal{H}}$, identity operator $\mathrm{Id}_{\mathcal{H}}$, scalar product $\langle\cdot\,|\,\cdot\rangle_{\mathcal{H}}$, associated norm $\|\cdot\|_{\mathcal{H}}$, and quadratic kernel $\mathscr{Q}_{\mathcal{H}}=\|\cdot\|_{\mathcal{H}}^2/2$. In addition, $\mathcal{G}$ is a real Hilbert space, the space of bounded linear operators from $\mathcal{H}$ to $\mathcal{G}$ is denoted by $\mathcal{B}\,(\mathcal{H},\mathcal{G})$, and we set $\mathcal{B}\,(\mathcal{H})=\mathcal{B}\,(\mathcal{H},\mathcal{H})$. The Legendre conjugate of $f\colon\mathcal{H}\to[-\infty,+\infty]$ is

$$f^*\colon\mathcal{H}\to[-\infty,+\infty]\colon x^*\mapsto\sup_{x\in\mathcal{H}}\big(\langle x\,|\,x^*\rangle_{\mathcal{H}}-f(x)\big),\tag{2.1}$$

the Moreau envelope of index $\gamma\in\,]0,+\infty[$ of $f\colon\mathcal{H}\to[-\infty,+\infty]$ is

$$^{\gamma}f\colon\mathcal{H}\to[-\infty,+\infty]\colon x\mapsto\inf_{y\in\mathcal{H}}\Big(f(y)+\frac{1}{\gamma}\mathscr{Q}_{\mathcal{H}}(x-y)\Big),\tag{2.2}$$

and the adjoint of $L\in\mathcal{B}\,(\mathcal{H},\mathcal{G})$ is denoted by L^*.

In analysis, there are several ways to compose a function $g\colon\mathcal{G}\to[-\infty,+\infty]$ and an operator $L\in\mathcal{B}\,(\mathcal{H},\mathcal{G})$ in order to construct a function from $\mathcal{H}$ to $[-\infty,+\infty]$. The most common is the standard composition

$$g\circ L\colon\mathcal{H}\to[-\infty,+\infty]\colon x\mapsto g(Lx).\tag{2.3}$$

Another instance is the infimal postcomposition of g by L^*, that is (see [2, Section 12.5] and [16, Section I.5], and, for applications, [4, 5, 19]),

$$L^*\rhd g\colon\mathcal{H}\to[-\infty,+\infty]\colon x\mapsto\inf_{\substack{y\in\mathcal{G}\\ L^*y=x}}g(y).\tag{2.4}$$

These two operations are dually related by the identities $(L^*\rhd g)^*=g^*\circ L$ and, under certain qualification conditions, $(g\circ L)^*=L^*\rhd g^*$ [2, Corollary 15.28]. The focus of the present paper is on the following alternative operations introduced in [9], where they were shown to manifest themselves in various variational models.

Definition 2.1 Let $L\in\mathcal{B}\,(\mathcal{H},\mathcal{G})$, $g\colon\mathcal{G}\to[-\infty,+\infty]$, and $\gamma\in\,]0,+\infty[$. The *proximal composition* of g and L with parameter γ is the function $L\overset{\gamma}{\diamond}g\colon\mathcal{H}\to[-\infty,+\infty]$ given by

$$L\overset{\gamma}{\diamond}g=\Big(\tfrac{1}{\gamma}(g^*)\circ L\Big)^*-\frac{1}{\gamma}\mathscr{Q}_{\mathcal{H}},\tag{2.5}$$

and the *proximal cocomposition* of g and L with parameter γ is $L\overset{\gamma}{\blacklozenge}g=(L\overset{1/\gamma}{\diamond}g^*)^*$.

In [9], proximal compositions were studied only in the case when $\gamma=1$ and few of their properties were explored. The goal of this paper is to carry out an in-depth analysis of these compositions, leading to results which are new even when $\gamma = 1$. We study in particular convexity, Legendre conjugacy, differentiability, subdifferentiability, Moreau envelopes, minimizers, recession functions, perspective functions, as well as the preservation of properties such as coercivity, supercoercivity, and Lipschitzianity. We also investigate the behavior of $L \overset{\gamma}{\diamond} g$ and $L \overset{\gamma}{\blacklozenge} g$ as γ varies. Another contribution of our work is to derive from these results a systematic analysis of the notions of integral proximal mixtures and comixtures. These operations, recently introduced in [7], combine arbitrary families of convex functions and linear operators acting in different spaces in such a way that the proximity operator of the mixture is explicitly computable in terms of those of the individual functions. In turn, this analysis leads to new results on the proximal expectation of a family of convex functions.

The remainder of the paper is organized as follows. In Section 2.2.2, we provide our notation and the necessary mathematical background. In Section 2.2.3, we investigate various variational properties of proximal compositions. Finally, Section 2.2.4 is devoted to applications to integral proximal mixtures and proximal expectations.

2.2.2 Notation and background

We first present our notation, which follows [2] (see also the first paragraph of Section 2.2.1).

Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$. The range of L is denoted by $\operatorname{ran} L$ and, if it is closed, the generalized inverse of L is denoted by $L^\dagger$. Further, L is called an isometry if $L^* \circ L = \mathrm{Id}_{\mathcal{H}}$ and a coisometry if $L \circ L^* = \mathrm{Id}_{\mathcal{G}}$. Let $f\colon \mathcal{H} \to [-\infty, +\infty]$. We set

$$\begin{cases} \operatorname{cam} f = \big\{h\colon \mathcal{H} \to \mathbb{R} \mid h \text{ is continuous, affine, and } h \leqslant f\big\} \\ \overline{f} = \sup\big\{h\colon \mathcal{H} \to [-\infty, +\infty] \mid h \text{ is lower semicontinuous and } h \leqslant f\big\} \\ \breve{f} = \sup\big\{h\colon \mathcal{H} \to [-\infty, +\infty] \mid h \text{ is lower semicontinuous, convex, and } h \leqslant f\big\}. \end{cases} \tag{2.6}$$

The infimal postcomposition of f by $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ (see (2.4)) is denoted by $L \rhd f$ if, for every $y \in L(\operatorname{dom} f)$, there exists $x \in \mathcal{H}$ such that $Lx = y$ and $(L \rhd f)(y) = f(x) \in \left]-\infty, +\infty\right]$. The function f is proper if $\operatorname{dom} f = \big\{x \in \mathcal{H} \mid f(x) < +\infty\big\} \neq \varnothing$ and $-\infty \notin f(\mathcal{H})$. If f is proper, its subdifferential is

$$\partial f\colon \mathcal{H} \to 2^{\mathcal{H}}\colon x \mapsto \big\{x^* \in \mathcal{H} \mid (\forall y \in \mathcal{H})\ \langle y - x \mid x^*\rangle_{\mathcal{H}} + f(x) \leqslant f(y)\big\} \tag{2.7}$$

and, if f is also convex, its recession function at $x \in \mathcal{H}$ is

$$(\operatorname{rec} f)(x) = \sup_{y \in \operatorname{dom} f} \big(f(x+y) - f(y)\big). \tag{2.8}$$

If f and $g\colon \mathcal{H} \to \left]-\infty,+\infty\right]$ are proper, their infimal convolution is

$$f \,\Box\, g\colon \mathcal{H} \to [-\infty,+\infty]\colon x \mapsto \inf_{y\in\mathcal{H}} \big(f(y)+g(x-y)\big). \tag{2.9}$$

We denote by $\Gamma_0(\mathcal{H})$ the class of functions from $\mathcal{H}$ to $\left]-\infty,+\infty\right]$ which are proper, lower semicontinuous, and convex. If $f \in \Gamma_0(\mathcal{H})$, its proximity operator is

$$\operatorname{prox}_f\colon \mathcal{H} \to \mathcal{H}\colon x \mapsto \underset{y\in\mathcal{H}}{\operatorname{argmin}} \big(f(y)+\mathscr{Q}_{\mathcal{H}}(x-y)\big). \tag{2.10}$$

Let $C \subset \mathcal{H}$. Then ι_C denotes the indicator function of C and σ_C the support function of C. If C is convex, its normal cone is denoted by N_C and its strong relative interior is the set $\operatorname{sri} C$ of points $x \in C$ such that the smallest cone containing $C-x$ is a closed vector subspace of $\mathcal{H}$. If C is nonempty, closed, and convex, its projection operator is denoted by proj_C. Finally, the closed ball with center $x \in \mathcal{H}$ and radius $\rho \in \left]0,+\infty\right[$ is denoted by $B(x;\rho)$.

The following facts will be frequently used in the paper.

Lemma 2.2 *Let f and g be functions from $\mathcal{H}$ to $[-\infty,+\infty]$. Then the following hold:*

(i) $f^{**} \leqslant f$.

(ii) $f \leqslant g \Rightarrow g^* \leqslant f^*$.

(iii) $f^{***} = f^*$.

(iv) $f^* \equiv +\infty \Leftrightarrow \operatorname{cam} f = \varnothing$.

(v) $f^* \in \Gamma_0(\mathcal{H}) \Leftrightarrow [\, f$ *is proper and* $\operatorname{cam} f \neq \varnothing\,]$.

Proof. (i)–(iii): [2, Proposition 13.16].

(iv): [2, Proposition 13.12(ii)].

(v): Combine [2, Proposition 13.10(ii)] and (iv). ☐

Lemma 2.3 [2, Propositions 13.10(ii) and 13.23(i)–(ii)] *Let $f\colon \mathcal{H} \to [-\infty,+\infty]$ and let $\rho \in \left]0,+\infty\right[$. Then the following hold:*

(i) $(\rho f)^* = \rho f^*(\cdot/\rho)$.

(ii) $(\rho f(\cdot/\rho))^* = \rho f^*$.

(iii) $(f(\rho\cdot))^* = f^*(\cdot/\rho)$.

The next lemma follows easily from (2.2).

Lemma 2.4 *Let $f\colon \mathcal{H} \to [-\infty,+\infty]$, $\gamma \in \left]0,+\infty\right[$, and $\rho \in \left]0,+\infty\right[$. Then the following hold:*

(i) $\rho({}^{\gamma}f) = {}^{\frac{\gamma}{\rho}}(\rho f)$.

(ii) $({}^{\gamma}f)(\rho\cdot) = {}^{\frac{\gamma}{\rho^2}}(f(\rho\cdot))$.

Lemma 2.5 *Let $f \in \Gamma_0(\mathcal{H})$ and $\gamma \in \left]0,+\infty\right[$. Then the following hold:*

(i) [2, Theorem 9.20] $\operatorname{cam} f \neq \varnothing$.

(ii) [2, Corollary 13.38] $f^* \in \Gamma_0(\mathcal{H})$ *and* $f^{**} = f$.

(iii) [2, Corollary 16.30] $\partial f^* = (\partial f)^{-1}$.

(iv) [2, Remark 14.4] ${}^1 f + {}^1(f^*) = \mathscr{Q}_{\mathcal{H}}$ *and* $\operatorname{prox}_f + \operatorname{prox}_{f^*} = \mathrm{Id}_{\mathcal{H}}$.

(v) [2, Theorem 13.49] $\operatorname{rec}(f^*) = \sigma_{\operatorname{dom} f}$ *and* $\operatorname{rec} f = \sigma_{\operatorname{dom} f^*}$.

(vi) [2, Propositions 12.15 and 12.30] ${}^\gamma f \colon \mathcal{H} \to \mathbb{R}$ *is convex and Fréchet differentiable.*

(vii) [2, Proposition 12.30] $\nabla({}^\gamma f) = (\mathrm{Id}_{\mathcal{H}} - \operatorname{prox}_{\gamma f})/\gamma$.

(viii) [2, Proposition 14.1] $(f + \gamma \mathscr{Q}_{\mathcal{H}})^* = {}^\gamma(f^*)$.

Lemma 2.6 *Let* $f \colon \mathcal{H} \to \left]-\infty, +\infty\right]$, $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$, *and* $\gamma \in \left]0, +\infty\right[$. *Then the following hold:*

(i) [2, Proposition 13.24(iii)] $({}^\gamma f)^* = f^* + \gamma \mathscr{Q}_{\mathcal{H}}$.

(ii) [2, Proposition 13.24(iv)] $(L \triangleright f)^* = f^* \circ L^*$.

(iii) [2, Corollary 15.28(i)] *Suppose that* $f \in \Gamma_0(\mathcal{H})$ *and* $0 \in \operatorname{sri}(\operatorname{dom} f - \operatorname{ran} L^*)$. *Then* $(f \circ L^*)^* = L \mathbin{\triangleright\!\!\!\cdot} f^*$.

Lemma 2.7 *Let* $f \in \Gamma_0(\mathcal{H})$, $g \in \Gamma_0(\mathcal{H})$, *and* $\gamma \in \left]0, +\infty\right[$ *be such that* ${}^\gamma f = {}^\gamma g$. *Then* $f = g$.

Proof. By Lemma 2.6(i), $f^* = ({}^\gamma f)^* - \gamma \mathscr{Q}_{\mathcal{H}} = ({}^\gamma g)^* - \gamma \mathscr{Q}_{\mathcal{H}} = g^*$. Therefore, we deduce from Lemma 2.5(ii) that $f = f^{**} = g^{**} = g$. ☐

Lemma 2.8 *Let* $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ *and set* $\Phi = \mathscr{Q}_{\mathcal{G}} - \mathscr{Q}_{\mathcal{H}} \circ L^*$. *Then* Φ *is convex if and only if* $\|L\| \leqslant 1$.

Proof. Since $\operatorname{dom} \Phi = \mathcal{G}$ and $\nabla \Phi = \mathrm{Id}_{\mathcal{G}} - L \circ L^*$, we deduce from [2, Proposition 17.7] that Φ is convex $\Leftrightarrow$ $\mathrm{Id}_{\mathcal{G}} - L \circ L^*$ is monotone $\Leftrightarrow$ $\|L^* \cdot\|_{\mathcal{H}}^2 \leqslant \|\cdot\|_{\mathcal{G}}^2 \Leftrightarrow \|L^*\| \leqslant 1 \Leftrightarrow \|L\| \leqslant 1$. ☐

Lemma 2.9 [2, Proposition 17.36(iii)] *Let* $A \in \mathcal{B}(\mathcal{H})$ *be monotone and self-adjoint. Suppose that* $\operatorname{ran} A$ *is closed, set* $q_A \colon \mathcal{H} \to \mathbb{R} \colon x \mapsto \langle x \mid Ax \rangle_{\mathcal{H}}/2$, *and define* $q_{A^\dagger}$ *likewise. Then* $q_A^* = \iota_{\operatorname{ran} A} + q_{A^\dagger}$.

2.2.3 Proximal compositions

2.2.3.1 General properties

We start with direct consequences of Definition 2.1.

Proposition 2.10 *Let* $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$, $g \colon \mathcal{G} \to [-\infty, +\infty]$, $\gamma \in \left]0, +\infty\right[$, *and* $\rho \in \left]0, +\infty\right[$. *Then the following hold:*

(i) *Let* $h \colon \mathcal{G} \to [-\infty, +\infty]$ *be such that* $g^{**} \leqslant h \leqslant g$. *Then* $L \overset{\gamma}{\diamond} h = L \overset{\gamma}{\diamond} g$ *and* $L \overset{\gamma}{\blacklozenge} h = L \overset{\gamma}{\blacklozenge} g$.

(ii) $(L \overset{\gamma}{\diamond} g)^* = L \overset{1/\gamma}{\blacklozenge} g^*$.

(iii) $(L \overset{\gamma}{\blacklozenge} g)^* = (L \overset{1/\gamma}{\diamond} g^*)^{**}$.

(iv) $(L \overset{\gamma}{\diamond} g)^{**} = (L \overset{1/\gamma}{\blacklozenge} g^*)^*$.

(v) $\rho(L \overset{\gamma}{\diamond} g) = L \overset{\gamma/\rho}{\diamond} (\rho g)$.

(vi) $(L \overset{\gamma}{\diamond} g)(\rho\cdot) = L \overset{\gamma/\rho^2}{\diamond} (g(\rho\cdot))$.

(vii) $\rho(L \overset{\gamma}{\blacklozenge} g) = L \overset{\gamma/\rho}{\blacklozenge} (\rho g)$.

(viii) $(L \overset{\gamma}{\blacklozenge} g)(\rho\cdot) = L \overset{\gamma/\rho^2}{\blacklozenge} (g(\rho\cdot))$.

Proof. (i): By Lemma 2.2(ii)–(iii), $g^* = g^{***} \geqslant h^* \geqslant g^*$. Therefore, $h^* = g^*$, and the claims follow from Definition 2.1.

(ii): It follows from Definition 2.1 and (i) that $L \overset{1/\gamma}{\blacklozenge} g^* = (L \overset{\gamma}{\diamond} g^{**})^* = (L \overset{\gamma}{\diamond} g)^*$.

(iii): An immediate consequence of Definition 2.1.

(iv): This follows from (ii).

(v): Combining Lemmas 2.3(ii), 2.4(i)–(ii), and 2.3(i), we obtain

$$\rho\Big({}^{\frac{1}{\gamma}}(g^*) \circ L\Big)^* = \Big(\rho\,{}^{\frac{1}{\gamma}}(g^*) \circ (L/\rho)\Big)^* = \Big({}^{\frac{\rho}{\gamma}}\big(\rho g^*(\cdot/\rho)\big) \circ L\Big)^* = \Big({}^{\frac{\rho}{\gamma}}\big((\rho g)^*\big) \circ L\Big)^*. \tag{2.11}$$

The assertion therefore follows from Definition 2.1.

(vi): We deduce from Lemmas 2.3(iii) and 2.4(ii) that

$$\Big({}^{\frac{1}{\gamma}}(g^*) \circ L\Big)^*(\rho\cdot) = \Big({}^{\frac{1}{\gamma}}(g^*) \circ (L/\rho)\Big)^* = \Big({}^{\frac{\rho^2}{\gamma}}\big(g^*(\cdot/\rho)\big) \circ L\Big)^* = \Big({}^{\frac{\rho^2}{\gamma}}\Big(\big(g(\rho\cdot)\big)^*\Big) \circ L\Big)^*. \tag{2.12}$$

In view of Definition 2.1, the assertion is established.

(vii): We invoke Definition 2.1, Lemma 2.3(ii), (v), (vi), and Lemma 2.3(i) to get

$$\rho(L \overset{\gamma}{\blacklozenge} g) = \rho(L \overset{1/\gamma}{\diamond} g^*)^* = \Big(\rho(L \overset{1/\gamma}{\diamond} g^*)(\cdot/\rho)\Big)^* = \Big(L \overset{\rho/\gamma}{\diamond} (\rho g)^*\Big)^* = L \overset{\gamma/\rho}{\blacklozenge} (\rho g). \tag{2.13}$$

(viii): By Definition 2.1, Lemma 2.3(iii), and (vi), we get

$$(L \overset{\gamma}{\blacklozenge} g)(\rho\cdot) = (L \overset{1/\gamma}{\diamond} g^*)^*(\rho\cdot) = \Big((L \overset{1/\gamma}{\diamond} g^*)(\cdot/\rho)\Big)^* = \Big(L \overset{\rho^2/\gamma}{\diamond} \big(g(\rho\cdot)\big)^*\Big)^* = L \overset{\gamma/\rho^2}{\blacklozenge} \big(g(\rho\cdot)\big), \tag{2.14}$$

which completes the proof. ◻

Proposition 2.11 *Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$, let $g\colon \mathcal{G} \to]-\infty, +\infty]$ be a proper function such that* $\operatorname{cam} g \neq \varnothing$*, let $\gamma \in]0, +\infty[$, and set $\Phi = \mathscr{Q}_{\mathcal{G}} - \mathscr{Q}_{\mathcal{H}} \circ L^*$. Then the following hold:*

(i) $L \overset{\gamma}{\diamond} g = L^* \triangleright (g^{**} + \Phi/\gamma)$.

(ii) $L \overset{\gamma}{\blacklozenge} g = (g^* + \gamma\Phi)^* \circ L$.

(iii) $\operatorname{dom}(L \overset{\gamma}{\diamond} g) = L^*(\operatorname{dom} g^{**})$.

(iv) *Suppose that one of the following are satisfied:*

  (a) $0 < \|L\| < 1$.

  (b) $\operatorname{dom} g^{**} = \mathcal{G}$.

  Then $\operatorname{dom}(L \overset{\gamma}{\blacklozenge} g) = \mathcal{H}$.

(v) $L \overset{\gamma}{\blacklozenge} g \geqslant {}^{\gamma}(g^{**}) \circ L$.

Proof. By Lemma 2.2(v), $g^* \in \Gamma_0(\mathcal{G})$. Therefore, Lemma 2.5(vi) implies that $\operatorname{dom} {}^{\frac{1}{\gamma}}(g^*) = \mathcal{G}$ and that ${}^{\frac{1}{\gamma}}(g^*) \in \Gamma_0(\mathcal{G})$.

(i): Let $x \in \mathcal{H}$. Because $\operatorname{dom} {}^{\frac{1}{\gamma}}(g^*) - \operatorname{ran} L = \mathcal{G}$, it follows from Definition 2.1 and items (iii) and (i) in Lemma 2.6 that

$$
\begin{aligned}
(L \overset{\gamma}{\lozenge} g)(x) &= \Big(\Big({}^{\frac{1}{\gamma}}(g^*) \circ L \Big)^* - \frac{1}{\gamma} \mathscr{Q}_{\mathcal{H}} \Big)(x) \\
&= \Big(L^* \triangleright \Big({}^{\frac{1}{\gamma}}(g^*) \Big)^* \Big)(x) - \frac{1}{\gamma} \mathscr{Q}_{\mathcal{H}}(x) \\
&= \Big(L^* \triangleright \Big(g^{**} + \frac{1}{\gamma} \mathscr{Q}_{\mathcal{G}} \Big) \Big)(x) - \frac{1}{\gamma} \mathscr{Q}_{\mathcal{H}}(x) \\
&= \min_{\substack{y \in \mathcal{G} \\ L^* y = x}} \Big(g^{**}(y) + \frac{1}{\gamma} \mathscr{Q}_{\mathcal{G}}(y) \Big) - \frac{1}{\gamma} \mathscr{Q}_{\mathcal{H}}(x) \\
&= \min_{\substack{y \in \mathcal{G} \\ L^* y = x}} \Big(g^{**}(y) + \frac{1}{\gamma} \mathscr{Q}_{\mathcal{G}}(y) - \frac{1}{\gamma} \mathscr{Q}_{\mathcal{H}}(L^* y) \Big) \\
&= \min_{\substack{y \in \mathcal{G} \\ L^* y = x}} \Big(g^{**}(y) + \frac{1}{\gamma} \Phi(y) \Big). \qquad (2.15)
\end{aligned}
$$

(ii): By Definition 2.1, (i), and Lemmas 2.2(iii) and 2.6(ii),

$$
L \overset{\gamma}{\blacklozenge} g = \big(L \overset{1/\gamma}{\lozenge} g^* \big)^* = \Big(L^* \triangleright \big(g^{***} + \gamma \Phi \big) \Big)^* = \Big(L^* \triangleright \big(g^* + \gamma \Phi \big) \Big)^* = \big(g^* + \gamma \Phi \big)^* \circ L. \qquad (2.16)
$$

(iii): Since $\operatorname{dom} \Phi = \mathcal{G}$, [2, Proposition 12.36(i)] and (i) imply that $\operatorname{dom}(L \overset{\gamma}{\lozenge} g) = L^*(\operatorname{dom}(g^{**} + \Phi/\gamma)) = L^*(\operatorname{dom} g^{**})$.

(iv): By Lemma 2.8, $\Phi \in \Gamma_0(\mathcal{G})$. Because $\operatorname{dom} \Phi = \mathcal{G}$, the identity $(\gamma\Phi)^* = \Phi^*/\gamma$ and [2, Proposition 15.2] imply that

$$
\big(g^* + \gamma \Phi \big)^* = g^{**} \,\square\, (\gamma\Phi)^* = g^{**} \,\square\, \big(\Phi^*/\gamma \big). \qquad (2.17)
$$

On the other hand, we have $(1 - \|L\|^2) \mathscr{Q}_{\mathcal{G}} \leqslant \Phi$. Hence, in view of property (iv)(a) and Lemma 2.2(ii), we have $\Phi^* \leqslant \mathscr{Q}_{\mathcal{G}}/(1 - \|L\|^2)$, which yields $\operatorname{dom} \Phi^* = \mathcal{G}$. We thus deduce from (2.17) that $\operatorname{dom}(g^* + \gamma\Phi)^* = \operatorname{dom} g^{**} + \operatorname{dom} \Phi^* = \mathcal{G}$ and obtain the assertion via (ii).

(v): Since $\Phi \leqslant \mathscr{Q}_\mathcal{G}$, $g^* + \gamma\Phi \leqslant g^* + \gamma\mathscr{Q}_\mathcal{G}$. In turn, Lemmas 2.5(viii) and 2.2(ii), and (ii) imply that

$$^{\gamma}(g^{**}) \circ L = \big(g^* + \gamma\mathscr{Q}_\mathcal{G}\big)^* \circ L \leqslant \big(g^* + \gamma\Phi\big)^* \circ L = L \overset{\gamma}{\blacklozenge} g, \tag{2.18}$$

which completes the proof. ☐

Remark 2.12 Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ satisfies $\|L\| = 1$, set $\Phi = \mathscr{Q}_\mathcal{G} - \mathscr{Q}_\mathcal{H} \circ L^*$, and set $A = \mathrm{Id}_\mathcal{G} - L \circ L^*$. Then A is monotone and self-adjoint, $\Phi\colon y \mapsto \langle y \mid Ay\rangle_\mathcal{G}/2$, and Lemma 2.9 shows that $\operatorname{dom}\Phi^* = \operatorname{ran} A$ under the assumption that $\operatorname{ran} A$ is closed. In this case, arguing as in (2.17) and using Proposition 2.11(ii), we obtain $\operatorname{dom}(L \overset{\gamma}{\blacklozenge} g) = L^{-1}(\operatorname{dom} g^{**} + \operatorname{ran} A)$.

Proposition 2.13 *Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ be such that $\operatorname{ran} L$ is closed and $\ker L = \{0\}$, let $g\colon \mathcal{G} \to \left]-\infty, +\infty\right]$ be a proper function such that $\operatorname{cam} g \neq \varnothing$, and let $\gamma \in \left]0, +\infty\right[$. Then the following hold:*

(i) *Suppose that g^{**} is coercive. Then $L \overset{\gamma}{\blacklozenge} g$ is coercive.*

(ii) *Suppose that g^{**} is supercoercive. Then $L \overset{\gamma}{\blacklozenge} g$ is supercoercive.*

Proof. It follows from [2, Fact 2.26] that there exists $\alpha \in \left]0, +\infty\right[$ such that $\|L \cdot\|_\mathcal{G} \geqslant \alpha\|\cdot\|_\mathcal{H}$. Thus, $\|Lx\|_\mathcal{G} \to +\infty$ as $\|x\|_\mathcal{H} \to +\infty$. On the other hand, combining Lemmas 2.2(v) and 2.5(ii), we obtain $g^{**} \in \Gamma_0(\mathcal{G})$.

(i): By [2, Corollary 14.18(i)], $^{\gamma}(g^{**})$ is coercive. Therefore, Proposition 2.11(v) implies that $(L \overset{\gamma}{\blacklozenge} g)(x) \geqslant (^{\gamma}(g^{**}))(Lx) \to +\infty$ as $\|x\|_\mathcal{H} \to +\infty$.

(ii): By [2, Corollary 14.18(ii)], $^{\gamma}(g^{**})$ is supercoercive. Hence, Proposition 2.11(v) yields

$$\frac{\big(L \overset{\gamma}{\blacklozenge} g\big)(x)}{\|x\|_\mathcal{H}} \geqslant \frac{^{\gamma}\big(g^{**}\big)(Lx)}{\|x\|_\mathcal{H}} \geqslant \alpha\frac{^{\gamma}\big(g^{**}\big)(Lx)}{\|Lx\|_\mathcal{G}} \to +\infty \quad \text{as} \quad \|x\|_\mathcal{H} \to +\infty, \tag{2.19}$$

which concludes the proof. ☐

The next proposition studies the effect of quadratic perturbations and translations.

Proposition 2.14 *Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$, $g \in \Gamma_0(\mathcal{G})$, $\alpha \in \mathbb{R}$, $\gamma \in \left]0, +\infty\right[$, $\rho \in \left[0, +\infty\right[$, and $u \in \mathcal{H}$. Given $w \in \mathcal{G}$, set $\tau_w g\colon y \mapsto g(y - w)$. Then the following hold:*

(i) *Set $\beta = \gamma/(1 + \rho\gamma)$. Then $L \overset{\gamma}{\diamond} (g + \rho\mathscr{Q}_\mathcal{G} + \langle\cdot \mid Lu\rangle_\mathcal{G} + \alpha) = (L \overset{\beta}{\diamond} g) + \rho\mathscr{Q}_\mathcal{H} + \langle\cdot \mid u\rangle_\mathcal{H} + \alpha$.*

(ii) *$L \overset{\gamma}{\blacklozenge} (\tau_{Lu}\, g + \alpha) = \tau_u(L \overset{\gamma}{\blacklozenge} g) + \alpha$.*

Proof. (i): Let $x \in \mathcal{H}$, set $h = g + \rho\mathscr{Q}_\mathcal{G} + \langle\cdot \mid Lu\rangle_\mathcal{G} + \alpha$, and set $\Phi = \mathscr{Q}_\mathcal{G} - \mathscr{Q}_\mathcal{H} \circ L^*$. Since $g \in \Gamma_0(\mathcal{G})$ and $\rho \geqslant 0$, we have $h \in \Gamma_0(\mathcal{G})$. In turn, Lemma 2.5(ii) yields $h^* \in \Gamma_0(\mathcal{G})$, $h^{**} = h$, and $g^{**} = g$. Therefore, it follows from Proposition 2.11(i) that

$$\big(L \overset{\gamma}{\diamond} h\big)(x) = \min_{\substack{y \in \mathcal{G} \\ L^*y = x}} \left(h(y) + \frac{1}{\gamma}\Phi(y)\right)$$

$$
\begin{aligned}
&= \min_{\substack{y\in\mathcal{G}\\ L^*y=x}} \Big(g(y) + \rho \mathscr{Q}_{\mathcal{G}}(y) + \langle y \mid Lu\rangle_{\mathcal{G}} + \alpha + \frac{1}{\gamma}\Phi(y) \Big) \\
&= \min_{\substack{y\in\mathcal{G}\\ L^*y=x}} \Big(g(y) + \rho\Phi(y) + \rho \mathscr{Q}_{\mathcal{H}}(L^*y) + \langle L^*y \mid u\rangle_{\mathcal{H}} + \frac{1}{\gamma}\Phi(y) \Big) + \alpha \\
&= \min_{\substack{y\in\mathcal{G}\\ L^*y=x}} \Big(g(y) + \Big(\rho + \frac{1}{\gamma}\Big)\Phi(y) \Big) + \rho \mathscr{Q}_{\mathcal{H}}(x) + \langle x \mid u\rangle_{\mathcal{H}} + \alpha \\
&= \min_{\substack{y\in\mathcal{G}\\ L^*y=x}} \Big(g(y) + \frac{1}{\beta}\Phi(y) \Big) + \rho \mathscr{Q}_{\mathcal{H}}(x) + \langle x \mid u\rangle_{\mathcal{H}} + \alpha \\
&= \big(L \overset{\beta}{\diamond} g\big)(x) + \rho \mathscr{Q}_{\mathcal{H}}(x) + \langle x \mid u\rangle_{\mathcal{H}} + \alpha. \qquad (2.20)
\end{aligned}
$$

(ii): Set $h = \tau_{Lu}\, g + \alpha$. We recall from [2, Proposition 13.23(iii)] that $h^* = g^* + \langle\cdot \mid Lu\rangle_{\mathcal{G}} - \alpha$. Hence, using Definition 2.1 and (i), we get

$$
\begin{aligned}
L \overset{\gamma}{\blacklozenge} h &= \Big(L \overset{1/\gamma}{\diamond} \big(g^* + \langle\cdot \mid Lu\rangle_{\mathcal{G}} - \alpha\big) \Big)^* \\
&= \Big(\big(L \overset{1/\gamma}{\diamond} g^*\big) + \langle\cdot \mid u\rangle_{\mathcal{H}} - \alpha \Big)^* \\
&= \tau_u \big(L \overset{1/\gamma}{\diamond} g^*\big)^* + \alpha \\
&= \tau_u \big(L \overset{\gamma}{\blacklozenge} g\big) + \alpha, \qquad (2.21)
\end{aligned}
$$

as claimed. ☐

2.2.3.2 Convex-analytical properties

We first study the convexity, Legendre conjugacy, and differentiability properties of proximal compositions. We then turn our attention to the evaluation of their proximity operators, subdifferentials, Moreau envelopes, recession functions, and perspective functions.

Proposition 2.15 *Suppose that $0 \neq L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$, let $g\colon \mathcal{G} \to \left]-\infty, +\infty\right]$ be a proper function such that* cam $g \neq \varnothing$*, let $\gamma \in \left]0, +\infty\right[$, and let $\alpha \in \left[-1/\gamma, +\infty\right[$. Suppose that $g^{**} - \alpha\mathscr{Q}_{\mathcal{G}}$ is convex and set $\beta = (\alpha + 1/\gamma)/\|L\|^2 - 1/\gamma$. Then $L \overset{\gamma}{\diamond} g - \beta\mathscr{Q}_{\mathcal{H}} \in \Gamma_0(\mathcal{H})$.*

Proof. By Lemma 2.2(v), $g^* \in \Gamma_0(\mathcal{G})$. Thus, Lemma 2.5(vi) implies that $\frac{1}{\gamma}(g^*) \circ L \in \Gamma_0(\mathcal{H})$. In turn, Lemma 2.5(ii) and Definition 2.1 yield $L \overset{\gamma}{\diamond} g + \mathscr{Q}_{\mathcal{H}}/\gamma = (\frac{1}{\gamma}(g^*) \circ L)^* \in \Gamma_0(\mathcal{H})$. Since $(-\beta - 1/\gamma)\mathscr{Q}_{\mathcal{H}}$ is continuous with domain $\mathcal{G}$, by [2, Lemma 1.27], $L \overset{\gamma}{\diamond} g - \beta\mathscr{Q}_{\mathcal{H}} = L \overset{\gamma}{\diamond} g + \mathscr{Q}_{\mathcal{H}}/\gamma + (-\beta - 1/\gamma)\mathscr{Q}_{\mathcal{H}}$ is proper and lower semicontinuous. It remains to show that $L \overset{\gamma}{\diamond} g - \beta\mathscr{Q}_{\mathcal{H}}$ is convex. Let $x \in \mathcal{H}$, set $\psi = \|L\|^2\mathscr{Q}_{\mathcal{G}} - \mathscr{Q}_{\mathcal{H}} \circ L^*$, and set $\Phi = \mathscr{Q}_{\mathcal{G}} - \mathscr{Q}_{\mathcal{H}} \circ L^*$. By

Proposition 2.11(i),

$$
\begin{aligned}
\big(L \overset{\gamma}{\diamond} g\big)(x) - \beta\mathscr{Q}_{\mathcal{H}}(x) &= \min_{\substack{y\in\mathcal{G}\\ L^*y=x}} \Big(g^{**}(y) + \frac{1}{\gamma}\Phi(y)\Big) - \beta\mathscr{Q}_{\mathcal{H}}(x)\\
&= \min_{\substack{y\in\mathcal{G}\\ L^*y=x}} \Big(g^{**}(y) + \frac{1}{\gamma}\Phi(y) - \beta\mathscr{Q}_{\mathcal{H}}(L^*y)\Big)\\
&= \min_{\substack{y\in\mathcal{G}\\ L^*y=x}} \Big(g^{**}(y) + \frac{1}{\gamma}\mathscr{Q}_{\mathcal{G}}(y) - \frac{1}{\|L\|^2}\Big(\alpha + \frac{1}{\gamma}\Big)\mathscr{Q}_{\mathcal{H}}(L^*y)\Big)\\
&= \min_{\substack{y\in\mathcal{G}\\ L^*y=x}} \Big(\big(g^{**}(y) - \alpha\mathscr{Q}_{\mathcal{G}}(y)\big) + \Big(\beta + \frac{1}{\gamma}\Big)\psi(y)\Big). \qquad (2.22)
\end{aligned}
$$

Since $\nabla\psi = \|L\|^2\mathrm{Id}_{\mathcal{G}} - L\circ L^*$, for every $y \in \mathcal{G}$, $\langle \nabla\psi(y) \mid y\rangle_{\mathcal{G}} = \|L\|^2\|y\|_{\mathcal{G}}^2 - \|L^*y\|_{\mathcal{H}}^2 \geqslant 0$. Therefore, we infer from [2, Proposition 17.7] that ψ is convex. Further, since $\alpha + 1/\gamma \geqslant 0$, $(\beta + 1/\gamma)\psi$ is convex with domain $\mathcal{G}$. By assumption, $g^{**} - \alpha\mathscr{Q}_{\mathcal{G}} \in \Gamma_0(\mathcal{G})$. Hence, the function $(g^{**} - \alpha\mathscr{Q}_{\mathcal{G}}) + (\beta + 1/\gamma)\psi$ is proper and convex. Altogether, in view of (2.22) and [2, Proposition 12.36(ii)], we conclude that $L \overset{\gamma}{\diamond} g - \beta\mathscr{Q}_{\mathcal{H}}$ is convex. ☐

Proposition 2.16 *Suppose that* $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ *satisfies* $0 < \|L\| \leqslant 1$, *let* $g\colon \mathcal{G} \to \left]-\infty,+\infty\right]$ *be a proper function such that* $\operatorname{cam} g \neq \varnothing$, *and let* $\gamma \in \left]0,+\infty\right[$. *Then the following hold:*

(i) $L \overset{\gamma}{\diamond} g \in \Gamma_0(\mathcal{H})$ *and* $L \overset{\gamma}{\blacklozenge} g \in \Gamma_0(\mathcal{H})$.

(ii) $(L \overset{\gamma}{\blacklozenge} g)^* = L \overset{1/\gamma}{\diamond} g^*$.

(iii) $L \overset{\gamma}{\diamond} g = (L \overset{1/\gamma}{\blacklozenge} g^*)^*$.

Proof. Recall that Lemmas 2.2(v) and 2.5(i) assert that $g^* \in \Gamma_0(\mathcal{G})$ and $\operatorname{cam} g^* \neq \varnothing$.

(i): Lemma 2.5(ii) yields $g^{**} \in \Gamma_0(\mathcal{G})$. Now set $\beta = (1/\|L\|^2 - 1)/\gamma$. Then $\beta \geqslant 0$ and, by applying Proposition 2.15 with $\alpha = 0$, we see that $L \overset{\gamma}{\diamond} g - \beta\mathscr{Q}_{\mathcal{H}} \in \Gamma_0(\mathcal{H})$ and hence that $L \overset{\gamma}{\diamond} g \in \Gamma_0(\mathcal{H})$. Likewise, applying Proposition 2.15 with $\alpha = 0$ to $g^* \in \Gamma_0(\mathcal{G})$ and using Lemma 2.2(iii) we get $L \overset{1/\gamma}{\diamond} g^* \in \Gamma_0(\mathcal{H})$. In view of Definition 2.1 and Lemma 2.5(ii), we conclude that $L \overset{\gamma}{\blacklozenge} g \in \Gamma_0(\mathcal{H})$.

(ii): We derive from Definition 2.1, (i), and Lemma 2.5(ii) that $(L \overset{\gamma}{\blacklozenge} g)^* = (L \overset{1/\gamma}{\diamond} g^*)^{**} = L \overset{1/\gamma}{\diamond} g^*$.

(iii): By Proposition 2.10(iv), (i), and Lemma 2.5(ii), $(L \overset{1/\gamma}{\blacklozenge} g^*)^* = (L \overset{\gamma}{\diamond} g)^{**} = L \overset{\gamma}{\diamond} g$. ☐

The next result examines differentiability.

Proposition 2.17 *Suppose that* $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ *satisfies* $0 < \|L\| \leqslant 1$, *let* $g\colon \mathcal{G} \to \left]-\infty,+\infty\right]$ *be a proper function such that* $\operatorname{cam} g \neq \varnothing$, *and let* $\gamma \in \left]0,+\infty\right[$. *Then the following hold:*

(i) *Suppose that* $\|L\| < 1$ *and set* $\beta = \gamma(1/\|L\|^2 - 1)$. *Then* $L \overset{\gamma}{\blacklozenge} g$ *is differentiable with a* $(1/\beta)$*-Lipschitzian gradient.*

(ii) *Let $\theta \in \left]0,+\infty\right[$, suppose that g is real-valued, convex, and differentiable with a θ-Lipschitzian gradient, and set $\beta = (1/\theta+\gamma)/\|L\|^2 - \gamma$. Then $L \overset{\gamma}{\blacklozenge} g$ is differentiable with a $(1/\beta)$-Lipschitzian gradient.*

Proof. We recall that a continuous convex function $f\colon \mathcal{H} \to \mathbb{R}$ is differentiable with a $(1/\beta)$-Lipschitzian gradient if and only if $f^* - \beta\mathscr{Q}_{\mathcal{H}}$ is convex [2, Theorem 18.15]. Further, by Proposition 2.16(ii), $(L \overset{\gamma}{\blacklozenge} g)^* = L \overset{1/\gamma}{\diamond} g^*$.

(i): By Proposition 2.11(iv)(a), $\operatorname{dom}(L \overset{\gamma}{\blacklozenge} g) = \mathcal{H}$. Now set $\alpha = 0$. Since $\alpha > -\gamma$, we deduce from Proposition 2.15 that $L \overset{1/\gamma}{\diamond} g^* - \beta\mathscr{Q}_{\mathcal{H}}$ is convex, i.e., that $(L \overset{\gamma}{\blacklozenge} g)^* - \beta\mathscr{Q}_{\mathcal{H}}$ is convex.

(ii): Since $g \in \Gamma_0(\mathcal{G})$, Lemma 2.5(ii) yields $\operatorname{dom} g^{**} = \operatorname{dom} g = \mathcal{G}$. Thus, it results from Proposition 2.11(iv)(b) that $\operatorname{dom}(L \overset{\gamma}{\blacklozenge} g) = \mathcal{H}$. Now set $\alpha = 1/\theta$. Since $g^* - \alpha\mathscr{Q}_{\mathcal{G}}$is convex and $\alpha > -\gamma$, Proposition 2.15 implies that $(L \overset{\gamma}{\blacklozenge} g)^* - \beta\mathscr{Q}_{\mathcal{H}} = L \overset{1/\gamma}{\diamond} g^* - \beta\mathscr{Q}_{\mathcal{H}}$ is convex. ☐

Remark 2.18 Proposition 2.17(i) guarantees the smoothness of the proximal cocomposition when $0 < \|L\| < 1$. Proposition 2.17(ii) shows that the Lipschitz constant of the gradient of the cocomposition is improved when the original function is itself smooth.

The following proposition motivates calling $L \overset{\gamma}{\diamond} g$ a proximal composition.

Proposition 2.19 *Suppose that $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ satisfies $0 < \|L\| \leqslant 1$, let $g\colon \mathcal{G} \to \left]-\infty,+\infty\right]$ be a proper function such that $\operatorname{cam} g \neq \varnothing$, and let $\gamma \in \left]0,+\infty\right[$. Then the following hold:*

(i) $\operatorname{prox}_{\gamma(L \overset{\gamma}{\diamond} g)} = L^* \circ \operatorname{prox}_{\gamma g^{**}} \circ L$.

(ii) $\operatorname{prox}_{\gamma(L \overset{\gamma}{\blacklozenge} g)} = \mathrm{Id}_{\mathcal{H}} - L^* \circ \big(\mathrm{Id}_{\mathcal{G}} - \operatorname{prox}_{\gamma g^{**}}\big) \circ L$.

Proof. As previously noted, $g^* \in \Gamma_0(\mathcal{G})$ and $g^{**} \in \Gamma_0(\mathcal{G})$.

(i): It follows from Proposition 2.10(v) and Definition 2.1 that

$$\mathscr{Q}_{\mathcal{H}} + \gamma\big(L \overset{\gamma}{\diamond} g\big) = \mathscr{Q}_{\mathcal{H}} + L \overset{1}{\diamond} (\gamma g) = \Big({}^{1}\big((\gamma g)^*\big) \circ L\Big)^*. \tag{2.23}$$

Since Proposition 2.16(i) yields $L \overset{\gamma}{\diamond} g \in \Gamma_0(\mathcal{H})$, we deduce from [2, Corollary 16.48(iii)], (2.23), and items (iii) and (vii) in Lemma 2.5 that

$$\begin{aligned}
\mathrm{Id}_{\mathcal{H}} + \gamma\partial\big(L \overset{\gamma}{\diamond} g\big) &= \partial\Big(\mathscr{Q}_{\mathcal{H}} + \gamma\big(L \overset{\gamma}{\diamond} g\big)\Big) \\
&= \bigg(\nabla\Big({}^{1}\big((\gamma g)^*\big) \circ L\Big)\bigg)^{-1} \\
&= \Big(L^* \circ \big(\mathrm{Id}_{\mathcal{G}} - \operatorname{prox}_{(\gamma g)^*}\big) \circ L\Big)^{-1}.
\end{aligned} \tag{2.24}$$

Hence, by [2, Proposition 16.44] and Lemma 2.5(iv),

$$\operatorname{prox}_{\gamma(L \overset{\gamma}{\diamond} g)} = (\mathrm{Id}_{\mathcal{H}} + \gamma\partial(L \overset{\gamma}{\diamond} g))^{-1} = L^* \circ \operatorname{prox}_{(\gamma g)^{**}} \circ L = L^* \circ \operatorname{prox}_{\gamma g^{**}} \circ L. \tag{2.25}$$

(ii): By Proposition 2.10(vii) and Definition 2.1, $\gamma(L \overset{\gamma}{\blacklozenge} g) = L \overset{1}{\blacklozenge} (\gamma g) = (L \overset{1}{\diamond} (\gamma g)^*)^*$. Therefore, Proposition 2.16(i) and Lemma 2.5(ii) entail that $(\gamma(L \overset{\gamma}{\blacklozenge} g))^* = L \overset{1}{\diamond} (\gamma g)^*$. In turn, we deduce from Lemma 2.5(iv) and (i) that $\operatorname{prox}_{\gamma(L \overset{\gamma}{\blacklozenge} g)} = \mathrm{Id}_{\mathcal{H}} - \operatorname{prox}_{L \overset{1}{\diamond} (\gamma g)^*} = \mathrm{Id}_{\mathcal{H}} - L^* \circ (\mathrm{Id}_{\mathcal{G}} - \operatorname{prox}_{\gamma g^{**}}) \circ L$. ☐

Our next result concerns the subdifferential of proximal compositions. We recall that the parallel composition of $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ by $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ is $L \triangleright A = (L \circ A^{-1} \circ L^*)^{-1}$ [2, Section 25.6].

Proposition 2.20 *Suppose that* $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ *satisfies* $0 < \|L\| \leqslant 1$*, let* $g\colon \mathcal{G} \to \left]-\infty,+\infty\right]$ *be a proper function such that* $\operatorname{cam} g \neq \varnothing$*, and let* $\gamma \in \left]0,+\infty\right[$*. Then the following hold:*

(i) $\partial(L \overset{\gamma}{\diamond} g) = L^* \triangleright (\partial g^{**} + (\mathrm{Id}_{\mathcal{G}} - L \circ L^*)/\gamma)$.

(ii) $\partial(L \overset{\gamma}{\blacklozenge} g) = L^* \circ (\partial g^* + \gamma(\mathrm{Id}_{\mathcal{G}} - L \circ L^*))^{-1} \circ L$.

Proof. As seen in Proposition 2.16(i), $L \overset{\gamma}{\diamond} g \in \Gamma_0(\mathcal{H})$ and $L \overset{\gamma}{\blacklozenge} g \in \Gamma_0(\mathcal{H})$. Now set $\Phi = \mathscr{Q}_{\mathcal{G}} - \mathscr{Q}_{\mathcal{H}} \circ L^*$ and $h = g^{**} + \Phi/\gamma$. We deduce from Lemmas 2.2(v), 2.5(ii), and 2.8 that $g^* \in \Gamma_0(\mathcal{G})$, $g^{**} \in \Gamma_0(\mathcal{G})$, and $\Phi \in \Gamma_0(\mathcal{G})$. Therefore, since $\operatorname{dom}\Phi = \mathcal{G}$, we have $h \in \Gamma_0(\mathcal{G})$ and, by Lemma 2.5(ii), $h^{**} = h$. On the other hand, $\operatorname{dom} h^* \cap \operatorname{ran} L \neq \varnothing$ since Propositions 2.11(ii) and 2.16(i) yield $h^* \circ L = L \overset{1/\gamma}{\blacklozenge} g^* \in \Gamma_0(\mathcal{G})$. Upon invoking Propositions 2.11(i) and 2.16(iii), we get

$$L^* \triangleright h = L \overset{\gamma}{\diamond} g = \Big(L \overset{1/\gamma}{\blacklozenge} g^*\Big)^* = \big(h^* \circ L\big)^*. \tag{2.26}$$

Therefore, [2, Proposition 16.42], Lemma 2.5(iii), and [2, Corollary 16.48(iii)] imply that

$$\partial(h^* \circ L) = L^* \circ \partial h^* \circ L = L^* \circ \big(\partial h\big)^{-1} \circ L = L^* \circ \big(\partial g^{**} + \nabla\Phi/\gamma\big)^{-1} \circ L. \tag{2.27}$$

(i): Combining (2.26), Lemma 2.5(iii), and (2.27), we obtain

$$\begin{aligned}\partial\big(L \overset{\gamma}{\diamond} g\big) = \partial\big(h^* \circ L\big)^* &= \big(\partial(h^* \circ L)\big)^{-1}\\ &= \Big(L^* \circ \big(\partial g^{**} + \nabla\Phi/\gamma\big)^{-1} \circ L\Big)^{-1}\\ &= L^* \triangleright \big(\partial g^{**} + \nabla\Phi/\gamma\big).\end{aligned} \tag{2.28}$$

(ii): By Definition 2.1, Lemma 2.5(iii), (i), and Lemma 2.2(iii),

$$\begin{aligned}\partial\big(L \overset{\gamma}{\blacklozenge} g\big) &= \partial\big(L \overset{1/\gamma}{\diamond} g^*\big)^*\\ &= \Big(\partial\big(L \overset{1/\gamma}{\diamond} g^*\big)\Big)^{-1}\\ &= \Big(L^* \triangleright \big(\partial g^{***} + \gamma\nabla\Phi\big)\Big)^{-1}\end{aligned}$$

$$= L^* \circ \big(\partial g^* + \gamma \nabla \Phi\big)^{-1} \circ L, \tag{2.29}$$

which completes the proof. ◻

Corollary 2.21 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ satisfies $0 < \|L\| \leqslant 1$, let $\beta \in \left]0, +\infty\right[$, let $\gamma \in \left]0, +\infty\right[$, and let $g \colon \mathcal{G} \to \mathbb{R}$ be convex and β-Lipschitzian. Then $L \overset{\gamma}{\blacklozenge} g$ is $(\beta\|L\|)$-Lipschitzian.*

Proof. We recall that a lower semicontinuous convex function $f \colon \mathcal{H} \to \mathbb{R}$ is β-Lipschitzian if and only if $\operatorname{ran} \partial f = \operatorname{dom} \partial f^* \subset B(0; \beta)$ [2, Corollary 17.19]. Since $g \in \Gamma_0(\mathcal{G})$, Lemma 2.5(ii) yields $g^* \in \Gamma_0(\mathcal{G})$. We therefore invoke Proposition 2.20(ii) to get

$$\begin{aligned}
\operatorname{ran} \partial\big(L \overset{\gamma}{\blacklozenge} g\big) &\subset L^*\Big(\operatorname{ran}\big(\partial g^* + \gamma(\mathrm{Id}_{\mathcal{G}} - L \circ L^*)\big)^{-1}\Big) \\
&= L^*\Big(\operatorname{dom}\big(\partial g^* + \gamma(\mathrm{Id}_{\mathcal{G}} - L \circ L^*)\big)\Big) \\
&= L^*\big(\operatorname{dom} \partial g^*\big) \\
&\subset L^*\big(B(0; \beta)\big) \\
&\subset B(0; \beta\|L\|),
\end{aligned} \tag{2.30}$$

where $L \overset{\gamma}{\blacklozenge} g \colon \mathcal{H} \to \left]-\infty, +\infty\right]$ is a real-valued lower semicontinuous convex function by Propositions 2.11(iv)(b) and 2.16(i). ◻

Let us now evaluate Moreau envelopes of proximal cocompositions.

Proposition 2.22 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ satisfies $0 < \|L\| \leqslant 1$, let $g \colon \mathcal{G} \to \left]-\infty, +\infty\right]$ be a proper function such that $\operatorname{cam} g \neq \varnothing$, let $\gamma \in \left]0, +\infty\right[$, and let $\rho \in \left]0, +\infty\right[$. Then the following hold:*

(i) ${}^{\rho}(L \overset{\gamma+\rho}{\blacklozenge} g) = L \overset{\gamma}{\blacklozenge} ({}^{\rho}g)$.

(ii) ${}^{\gamma}(L \overset{\gamma}{\blacklozenge} g) = {}^{\gamma}(g^{**}) \circ L$.

Proof. By Lemma 2.2(v) and Proposition 2.16(i), $L \overset{1/\gamma}{\diamond} g^* \in \Gamma_0(\mathcal{H})$. Therefore, Lemma 2.5(viii) and Definition 2.1 yield

$$\Big(\big(L \overset{1/\gamma}{\diamond} g^*\big) + \rho \mathscr{Q}_{\mathcal{H}}\Big)^* = {}^{\rho}\Big(\big(L \overset{1/\gamma}{\diamond} g^*\big)^*\Big) = {}^{\rho}\big(L \overset{\gamma}{\blacklozenge} g\big). \tag{2.31}$$

(i): We combine Definition 2.1, Lemma 2.6(i), Proposition 2.14(i), and (2.31) to arrive at

$$\begin{aligned}
L \overset{\gamma}{\blacklozenge} ({}^{\rho}g) &= \Big(L \overset{1/\gamma}{\diamond} ({}^{\rho}g)^*\Big)^* \\
&= \Big(L \overset{1/\gamma}{\diamond} (g^* + \rho \mathscr{Q}_{\mathcal{G}})\Big)^* \\
&= \Big(\big(L \overset{1/(\gamma+\rho)}{\diamond} g^*\big) + \rho \mathscr{Q}_{\mathcal{H}}\Big)^*
\end{aligned}$$

$$= {}^{\rho}\big(L \overset{\gamma+\rho}{\blacklozenge} g\big). \tag{2.32}$$

(ii): Since $g^* \in \Gamma_0(\mathcal{G})$, items (ii) and (vi) in Lemma 2.5 imply that ${}^{\gamma}(g^{**}) \in \Gamma_0(\mathcal{G})$ and that $\operatorname{dom} {}^{\gamma}(g^{**}) = \mathcal{G}$. Hence, ${}^{\gamma}(g^{**}) \circ L \in \Gamma_0(\mathcal{H})$ and it follows from Lemma 2.5(ii), Definition 2.1, and (2.31) that

$${}^{\gamma}(g^{**}) \circ L = \Big({}^{\gamma}(g^{**}) \circ L\Big)^{**} = \Big(\big(L \overset{1/\gamma}{\lozenge} g^*\big) + \gamma \mathscr{Q}_{\mathcal{G}}\Big)^* = {}^{\gamma}\big(L \overset{\gamma}{\blacklozenge} g\big), \tag{2.33}$$

as announced. ☐

Corollary 2.23 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ satisfies $0 < \|L\| \leqslant 1$, let $g\colon \mathcal{G} \to \left]-\infty, +\infty\right]$ be a proper function such that $\operatorname{cam} g \neq \varnothing$, and let $\gamma \in \left]0, +\infty\right[$. Then $\operatorname{Argmin}(L \overset{\gamma}{\blacklozenge} g) = \operatorname{Argmin}({}^{\gamma}(g^{**}) \circ L)$.*

Proof. Since the set of minimizers of a function in $\Gamma_0(\mathcal{H})$ coincides with that of its Moreau envelope [2, Propositions 17.5], the assertion follows from Proposition 2.22(ii). ☐

Corollary 2.24 *Let $\mathcal{K}$ be a real Hilbert space, suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ and $S \in \mathcal{B}(\mathcal{K}, \mathcal{H})$ satisfy $\|L\| \leqslant 1$, $\|S\| \leqslant 1$, and $L \circ S \neq 0$, let $g\colon \mathcal{G} \to \left]-\infty, +\infty\right]$ be a proper function such that $\operatorname{cam} g \neq \varnothing$, and let $\gamma \in \left]0, +\infty\right[$. Then the following hold:*

(i) $S \overset{\gamma}{\blacklozenge} (L \overset{\gamma}{\blacklozenge} g) = (L \circ S) \overset{\gamma}{\blacklozenge} g$.

(ii) $S \overset{\gamma}{\lozenge} (L \overset{\gamma}{\lozenge} g) = (L \circ S) \overset{\gamma}{\lozenge} g$.

Proof. (i): Set $f = L \overset{\gamma}{\blacklozenge} g$. Since $\|L \circ S\| \leqslant \|L\|\, \|S\| \leqslant 1$, we deduce from Proposition 2.16(i) that $f \in \Gamma_0(\mathcal{H})$, $S \overset{\gamma}{\blacklozenge} f \in \Gamma_0(\mathcal{K})$, and $(L \circ S) \overset{\gamma}{\blacklozenge} g \in \Gamma_0(\mathcal{K})$. By Lemma 2.5(ii), $f^{**} = f$. Hence, Proposition 2.22(ii) yields

$${}^{\gamma}\big(S \overset{\gamma}{\blacklozenge} f\big) = {}^{\gamma}(f^{**}) \circ S = {}^{\gamma}f \circ S = \Big({}^{\gamma}(g^{**}) \circ L\Big) \circ S = {}^{\gamma}\Big(\big(L \circ S\big) \overset{\gamma}{\blacklozenge} g\Big). \tag{2.34}$$

Therefore, the assertion follows from Lemma 2.7.

(ii): By Proposition 2.16(i), $L \overset{\gamma}{\lozenge} g \in \Gamma_0(\mathcal{H})$, $S \overset{\gamma}{\lozenge} (L \overset{\gamma}{\lozenge} g) \in \Gamma_0(\mathcal{K})$, and $(L \circ S) \overset{\gamma}{\lozenge} g \in \Gamma_0(\mathcal{K})$. Therefore, using Propositions 2.16(iii) and 2.10(ii), together with (i), we get

$$\begin{aligned} S \overset{\gamma}{\lozenge} (L \overset{\gamma}{\lozenge} g) &= \Big(S \overset{1/\gamma}{\blacklozenge} (L \overset{\gamma}{\lozenge} g)^*\Big)^* \\ &= \Big(S \overset{1/\gamma}{\blacklozenge} (L \overset{1/\gamma}{\blacklozenge} g^*)\Big)^* \\ &= \Big(\big(L \circ S\big) \overset{1/\gamma}{\blacklozenge} g^*\Big)^* \\ &= (L \circ S) \overset{\gamma}{\lozenge} g, \end{aligned} \tag{2.35}$$

which completes the proof. ☐

Proposition 2.25 *Suppose that $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ satisfies $0 < \|L\| \leqslant 1$, let $g\colon \mathcal{G} \to \left]-\infty,+\infty\right]$ be a proper function such that $\operatorname{cam} g \neq \varnothing$, and let $\gamma \in \left]0,+\infty\right[$. Then*

$$\operatorname{rec}(L \overset{\gamma}{\blacklozenge} g) = (\operatorname{rec}(g^{**})) \circ L. \tag{2.36}$$

Proof. By Lemmas 2.2(v) and 2.5(ii), $g^* \in \Gamma_0(\mathcal{G})$ and $g^{**} \in \Gamma_0(\mathcal{G})$. Therefore, Lemma 2.5(v), Propositions 2.16(ii) and 2.11(iii), and Lemma 2.2(iii) imply that

$$\begin{aligned}
\operatorname{rec}\big(L \overset{\gamma}{\blacklozenge} g\big) &= \sigma_{\operatorname{dom}(L \overset{\gamma}{\blacklozenge} g)^*} \\
&= \sigma_{\operatorname{dom}(L \overset{1/\gamma}{\diamond} g^*)} \\
&= \sigma_{L^*(\operatorname{dom} g^{***})} = \sigma_{\operatorname{dom} g^*} \circ L \\
&= \big(\operatorname{rec}(g^{**})\big) \circ L,
\end{aligned} \tag{2.37}$$

as claimed. ☐

Proposition 2.26 *Suppose that $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ satisfies $0 < \|L\| \leqslant 1$, let $g \in \Gamma_0(\mathcal{G})$, let*

$$\widetilde{g}\colon \mathcal{G} \oplus \mathbb{R} \to \left]-\infty,+\infty\right]\colon (y,\eta) \mapsto \begin{cases} \eta g(y/\eta), & \text{if } \eta > 0;\\ (\operatorname{rec} g)(y), & \text{if } \eta = 0;\\ +\infty, & \text{otherwise} \end{cases} \tag{2.38}$$

be its perspective, let $\gamma \in \left]0,+\infty\right[$, and set $\widetilde{L}\colon \mathcal{H} \oplus \mathbb{R} \to \mathcal{G} \oplus \mathbb{R}\colon (x,\xi) \mapsto (Lx,\xi)$. Then

$$\widetilde{L \overset{\gamma}{\blacklozenge} g}\colon \mathcal{H} \oplus \mathbb{R} \to \left]-\infty,+\infty\right]\colon (x,\xi) \mapsto \begin{cases} \Big(\widetilde{L} \overset{\xi\gamma}{\blacklozenge} \widetilde{g}\Big)(x,\xi), & \text{if } \xi > 0;\\ (\operatorname{rec} g)(Lx), & \text{if } \xi = 0;\\ +\infty, & \text{otherwise.} \end{cases} \tag{2.39}$$

Proof. Let $(x,\xi) \in \mathcal{H} \oplus \mathbb{R}$, set $\Phi = \mathscr{Q}_{\mathcal{G}} - \mathscr{Q}_{\mathcal{H}} \circ L^*$, and set $\Psi = \mathscr{Q}_{\mathcal{G}\oplus\mathbb{R}} - \mathscr{Q}_{\mathcal{H}\oplus\mathbb{R}} \circ \widetilde{L}^*$. We consider two cases.

- $\xi = 0$: It follows from Proposition 2.25 and Lemma 2.5(i)–(ii) that $(\widetilde{L \overset{\gamma}{\blacklozenge} g})(x,0) = (\operatorname{rec}(L \overset{\gamma}{\blacklozenge} g))(x) = (\operatorname{rec} g)(Lx)$.
- $\xi > 0$: Set $C = \big\{(y^*,\eta) \in \mathcal{G} \oplus \mathbb{R} \mid \eta + g^*(y^*) \leqslant 0\big\}$. Then [8, Items (ii) and (iv) in Proposition 2.3] assert that $\widetilde{g} \in \Gamma_0(\mathcal{G} \oplus \mathbb{R})$ and $(\widetilde{g}\,)^* = \iota_C$. Therefore, by Lemma 2.3(ii),

$$\begin{aligned}
(\forall y^* \in \mathcal{G}) \quad \sup_{\eta\in\mathbb{R}}\big(\eta\xi - (\widetilde{g}\,)^*(y^*,\eta)\big) &= \sup_{\eta\in\mathbb{R}}\big(\eta\xi - \iota_C(y^*,\eta)\big) \\
&= \sup_{\eta\in\left]-\infty,-g^*(y^*)\right]} \eta\xi
\end{aligned}$$

$$
\begin{aligned}
&= -\xi g^*(y^*) \\
&= -\big(\xi g(\cdot/\xi)\big)^*(y^*). \qquad (2.40)
\end{aligned}
$$

On the other hand, for every $\eta \in \mathbb{R}$, $\Psi(\cdot,\eta) = \Phi$ and, since $0 < \|L\| \leqslant 1$, we have $0 < \|\widetilde{L}\| \leqslant 1$. Hence, appealing to Proposition 2.11(ii), (2.40), and Proposition 2.10(vii)–(viii),

$$
\begin{aligned}
\Big(\widetilde{L} \overset{\xi\gamma}{\blacklozenge} \widetilde{g}\,\Big)(x,\xi) &= \big((\widetilde{g}\,)^* + \xi\gamma\Psi\big)^*\big(\widetilde{L}(x,\xi)\big) \\
&= \big((\widetilde{g}\,)^* + \xi\gamma\Psi\big)^*(Lx,\xi) \\
&= \sup_{(y^*,\eta)\in\mathcal{G}\oplus\mathbb{R}} \big(\langle (Lx,\xi) \mid (y^*,\eta)\rangle_{\mathcal{G}\oplus\mathbb{R}} - (\widetilde{g}\,)^*(y^*,\eta) - \xi\gamma\Psi(y^*,\eta)\big) \\
&= \sup_{(y^*,\eta)\in\mathcal{G}\oplus\mathbb{R}} \big(\eta\xi + \langle Lx \mid y^*\rangle_{\mathcal{G}} - (\widetilde{g}\,)^*(y^*,\eta) - \xi\gamma\Phi(y^*)\big) \\
&= \sup_{y^*\in\mathcal{G}} \Big(\langle Lx \mid y^*\rangle_{\mathcal{G}} - \xi\gamma\Phi(y^*) + \sup_{\eta\in\mathbb{R}}\big(\eta\xi - (\widetilde{g}\,)^*(y^*,\eta)\big)\Big) \\
&= \sup_{y^*\in\mathcal{G}} \big(\langle Lx \mid y^*\rangle_{\mathcal{G}} - \xi\gamma\Phi(y^*) - \big(\xi g(\cdot/\xi)\big)^*(y^*)\big) \\
&= \Big(\big(\xi g(\cdot/\xi)\big)^* + \xi\gamma\Phi\Big)^*(Lx) \\
&= \Big(L \overset{\xi\gamma}{\blacklozenge} \big(\xi g(\cdot/\xi)\big)\Big)(x) \\
&= \xi\big(L \overset{\gamma}{\blacklozenge} g\big)(x/\xi) \\
&= \Big(\widetilde{L \overset{\gamma}{\blacklozenge} g}\Big)(x,\xi). \qquad (2.41)
\end{aligned}
$$

We have thus proved (2.39). ☐

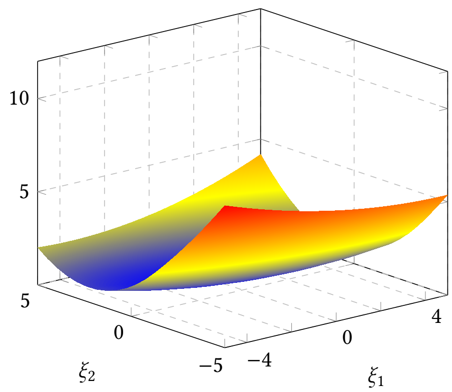
2.2.3.3 Comparison with standard compositions and infimal postcompositions

As mentioned in Section 2.2.1, our discussion involves several ways to compose a function defined on $\mathcal{G}$ with a linear operator from $\mathcal{H}$ to $\mathcal{G}$ in order to obtain a function defined on $\mathcal{H}$: the standard composition (2.3), the infimal postcomposition (2.4), and the proximal composition and cocomposition of Definition 2.1. We saw in Proposition 2.19 that a numerical advantage of the proximal compositions is that their proximity operators are easily decomposable in terms of that of the underlying function. Our purpose here is to compare these various compositions.

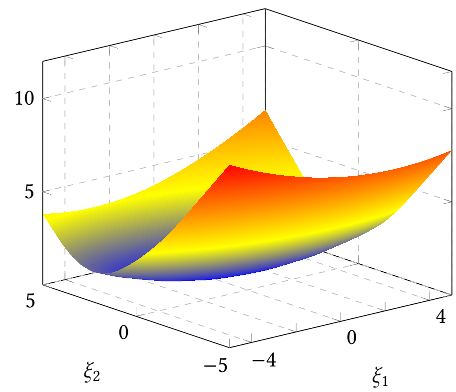
Example 2.27 Let

$$
\begin{cases}
L\colon \mathbb{R}^2 \to \mathbb{R}^5\colon (\xi_1,\xi_2) \mapsto \big(0.5\xi_2, -0.5\xi_1, -0.5\xi_2, 0.3\xi_1 + 0.4\xi_2, 0.1\xi_1 - 0.3\xi_2\big) \\
g\colon \mathbb{R}^5 \to \mathbb{R}\colon (\eta_1,\eta_2,\eta_3,\eta_4,\eta_5) \mapsto \|(\eta_1,\eta_2,\eta_3)\|_1 + \|(\eta_4 - 1, \eta_5 + 2)\|.
\end{cases} \qquad (2.42)
$$

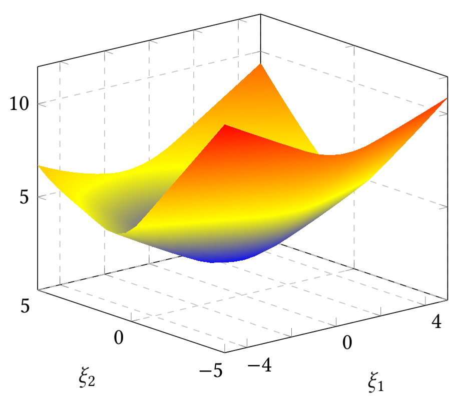
Figure 2.1 shows the graphs of both the standard composition and proximal cocomposition



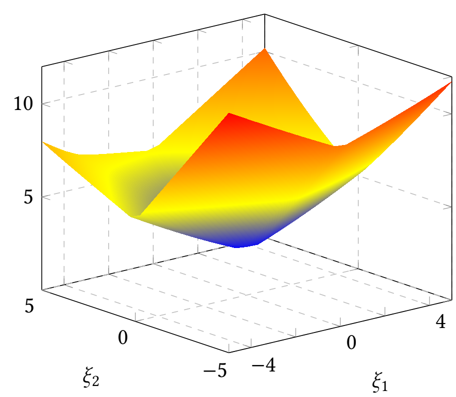
(a) Graph of $L \overset{10}{\blacklozenge} g$.



(b) Graph of $L \overset{5}{\blacklozenge} g$.

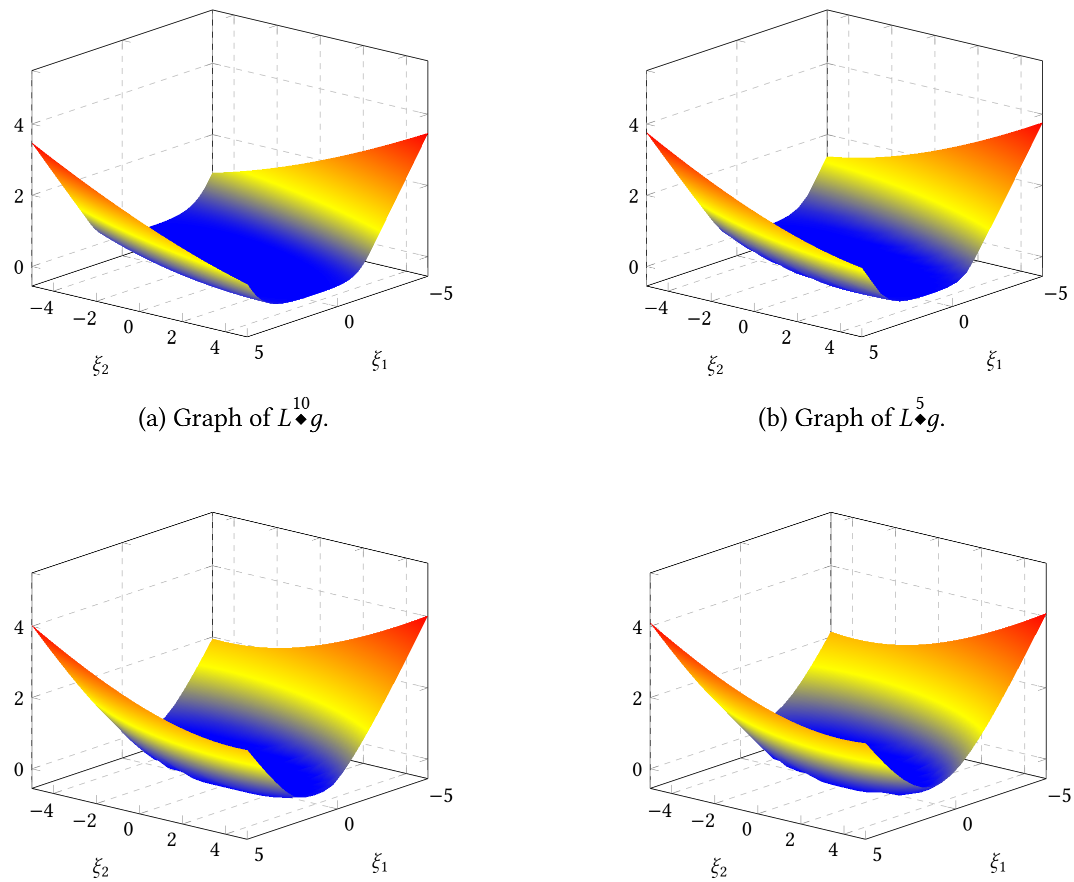


(c) Graph of $L \overset{1}{\blacklozenge} g$.



(d) Graph of $g \circ L$.

Figure 2.1 Graphs of the proximal cocomposition and of the standard composition in Example 2.27.



(a) Graph of $L \overset{10}{\diamond} g$.

(b) Graph of $L \overset{5}{\diamond} g$.

(c) Graph of $L \overset{1}{\diamond} g$.

(d) Graph of $g \circ L$.

Figure 2.2 Graphs of the proximal cocomposition and of the standard composition in Example 2.28.

for various values of γ.

Example 2.28 Let $C = B(0;2)$ and

$$\begin{cases} L\colon \mathbb{R}^2 \to \mathbb{R}^3\colon (\xi_1,\xi_2) \mapsto \big(0.7\xi_1 + 0.1\xi_2, -0.3\xi_1 + 0.4\xi_2, 0.5\xi_1 - 0.3\xi_2\big) \\ g\colon \mathbb{R}^3 \to \mathbb{R}\colon (\eta_1,\eta_2,\eta_3) \mapsto d_C(\eta_1,\eta_2,\eta_3). \end{cases} \tag{2.43}$$

Figure 2.2 shows the graphs of both the standard composition and proximal cocomposition for various values of γ.

As we now show, the pointwise orderings suggested by Figures 2.1 and 2.2 are generally true.

Proposition 2.29 *Suppose that $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ satisfies $0 < \|L\| \leqslant 1$, let $g\colon \mathcal{G} \to \left]-\infty,+\infty\right]$ be a proper function such that $\operatorname{cam} g \neq \varnothing$, and let $\gamma \in \left]0,+\infty\right[$. Then the following hold:*

(i) $L^* \triangleright g^{**} \leqslant L \overset{\gamma}{\diamond} g$.

(ii) ${}^{\gamma}(g^{**}) \circ L \leqslant L \overset{\gamma}{\blacklozenge} g \leqslant g^{**} \circ L$.

(iii) $L \overset{\gamma}{\blacklozenge} g \leqslant L \overset{\gamma}{\diamond} g$.

(iv) *Suppose that L is an isometry. Then $L \overset{\gamma}{\diamond} g = L \overset{\gamma}{\blacklozenge} g$.*

(v) *Suppose that L is a coisometry. Then $L \overset{\gamma}{\diamond} g = L^* \triangleright g^{**}$ and $L \overset{\gamma}{\blacklozenge} g = g^{**} \circ L$.*

(vi) *Suppose that L is invertible with $L^{-1} = L^*$. Then $L \overset{\gamma}{\diamond} g = L^* \triangleright g^{**} = g^{**} \circ L = L \overset{\gamma}{\blacklozenge} g$.*

Proof. Set $\Phi = \mathscr{Q}_{\mathcal{G}} - \mathscr{Q}_{\mathcal{H}} \circ L^*$ and observe that $0 \leqslant \Phi \leqslant \mathscr{Q}_{\mathcal{G}}$.

(i): Let $x \in \mathcal{H}$. By Proposition 2.11(i),

$$\big(L \overset{\gamma}{\diamond} g\big)(x) = \min_{\substack{y \in \mathcal{G} \\ L^*y = x}} \left(g^{**}(y) + \frac{1}{\gamma}\Phi(y)\right) \geqslant \inf_{\substack{y \in \mathcal{G} \\ L^*y = x}} g^{**}(y) = \big(L^* \triangleright g^{**}\big)(x). \tag{2.44}$$

(ii): The leftmost inequality is established in Proposition 2.11(v). Let us prove rightmost inequality. By Lemma 2.2(ii) and (i), $(L \overset{1/\gamma}{\diamond} g^*)^* \leqslant (L^* \triangleright g^{***})^*$. It therefore follows from Definition 2.1 and Lemmas 2.2(iii) and 2.6(ii) that

$$L \overset{\gamma}{\blacklozenge} g = \big(L \overset{1/\gamma}{\diamond} g^*\big)^* \leqslant (L^* \triangleright g^*)^* = g^{**} \circ L. \tag{2.45}$$

(iii): Set $f = {}^{1}(g^{**}) \circ L$. Since $\|L\| \leqslant 1$, $\mathscr{Q}_{\mathcal{G}} \circ L \leqslant \mathscr{Q}_{\mathcal{H}}$, and we deduce from Lemma 2.2(ii) that $(\mathscr{Q}_{\mathcal{H}} - f)^* \leqslant (\mathscr{Q}_{\mathcal{G}} \circ L - f)^*$. However, it results from Lemma 2.5(iv) that $\mathscr{Q}_{\mathcal{G}} \circ L - f = (\mathscr{Q}_{\mathcal{G}} - {}^{1}(g^{**})) \circ L = {}^{1}(g^*) \circ L$. Altogether, it follows from Definition 2.1 and [2, Proposition 13.29] that

$$L \overset{1}{\blacklozenge} g = \big(f^* - \mathscr{Q}_{\mathcal{H}}\big)^* = \big(\mathscr{Q}_{\mathcal{H}} - f\big)^* - \mathscr{Q}_{\mathcal{H}} \leqslant \big({}^{1}(g^*) \circ L\big)^* - \mathscr{Q}_{\mathcal{H}} = L \overset{1}{\diamond} g. \tag{2.46}$$

Hence, by Proposition 2.10(vii), (2.46), and Proposition 2.10(v), we get

$$L \overset{\gamma}{\blacklozenge} g = \frac{1}{\gamma}\big(L \overset{1}{\blacklozenge} (\gamma g)\big) \leqslant \frac{1}{\gamma}\big(L \overset{1}{\diamond} (\gamma g)\big) = L \overset{\gamma}{\diamond} g. \tag{2.47}$$

(iv): Here $\mathscr{Q}_{\mathcal{H}} = \mathscr{Q}_{\mathcal{G}} \circ L$ and therefore the inequalities in the proof of (iii) can be replaced with equalities.

(v): Here $\mathscr{Q}_{\mathcal{G}} = \mathscr{Q}_{\mathcal{H}} \circ L^*$ and thus $\Phi = 0$. Therefore, the result follows from Proposition 2.11(i)–(ii).

(vi): A consequence of (iv) and (v). ☐

Remark 2.30 Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ is an isometry, let $g\colon \mathcal{G} \to \left]-\infty, +\infty\right]$ be a proper function such that $\operatorname{cam} g \neq \varnothing$, and let $\gamma \in \left]0, +\infty\right[$. Then we recover from [2, Proposition 13.24(v)] as well as items (i), (iv), and (ii) in Proposition 2.29 the inequalities

$$(g^* \circ L)^* \leqslant L^* \rhd g^{**} \leqslant L \overset{\gamma}{\diamond} g = L \overset{\gamma}{\blacklozenge} g \leqslant g^{**} \circ L, \tag{2.48}$$

which appear in [9, Proposition 5.4] in the special case in which $\gamma = 1$.

Remark 2.31 Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ is a coisometry, let $g\colon \mathcal{G} \to \left]-\infty, +\infty\right]$ be a proper function such that $\operatorname{cam} g \neq \varnothing$, and let $\gamma \in \left]0, +\infty\right[$. Then Propositions 2.29(v) and 2.22(ii) imply that

$${}^{\gamma}\big((g^{**}) \circ L\big) = {}^{\gamma}\big(L \overset{\gamma}{\blacklozenge} g\big) = {}^{\gamma}(g^{**}) \circ L. \tag{2.49}$$

In particular, when $g \in \Gamma_0(\mathcal{G})$, we recover the fact that ${}^{\gamma}(g \circ L) = {}^{\gamma}g \circ L$ (see [21, Lemma 3]).

Proposition 2.32 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ satisfies $0 < \|L\| \leqslant 1$, let $g\colon \mathcal{G} \to \left]-\infty, +\infty\right]$ be a proper function such that $\operatorname{cam} g \neq \varnothing$, let $\gamma \in \left]0, +\infty\right[$, let $x \in \mathcal{H}$, and set $\Phi = \mathscr{Q}_{\mathcal{G}} - \mathscr{Q}_{\mathcal{H}} \circ L^*$. Then the following hold:*

(i) *Suppose that $y^* \in \partial g(Lx)$. Then $0 \leqslant g(Lx) - (L \overset{\gamma}{\blacklozenge} g)(x) \leqslant \gamma\Phi(y^*)$.*

(ii) *Suppose that $0 \in (\mathrm{Id}_{\mathcal{G}} - L \circ L^*)(\partial g(Lx))$. Then $(L \overset{\gamma}{\blacklozenge} g)(x) = g(Lx)$.*

Proof. (i): By [2, Proposition 16.10], $g(Lx) + g^*(y^*) = \langle Lx \mid y^* \rangle_{\mathcal{G}}$. Further, [2, Proposition 16.5] yields $g^{**}(Lx) = g(Lx) \in \mathbb{R}$. Therefore, we deduce from Propositions 2.29(ii) and 2.11(ii) that $(L \overset{\gamma}{\blacklozenge} g)(x) \in \mathbb{R}$ and that

$$\begin{aligned}
0 &\leqslant g(Lx) - \big(L \overset{\gamma}{\blacklozenge} g\big)(x)\\
&= g(Lx) - \big(g^* + \gamma\Phi\big)^*(Lx)\\
&= g(Lx) - \sup_{y \in \mathcal{G}}\big(\langle Lx \mid y\rangle_{\mathcal{G}} - g^*(y) - \gamma\Phi(y)\big)\\
&\leqslant g(Lx) - \big(\langle Lx \mid y^*\rangle_{\mathcal{G}} - g^*(y^*) - \gamma\Phi(y^*)\big)\\
&= \gamma\Phi(y^*).
\end{aligned} \tag{2.50}$$

(ii): There exists $y^* \in \partial g(Lx)$ such that $L(L^*y^*) = y^*$. Therefore, $\Phi(y^*) = 0$ and the conclusion follows from (i). ☐

Proposition 2.33 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ satisfies $0 < \|L\| \leqslant 1$, let $\beta \in \left]0, +\infty\right[$, let $\gamma \in \left]0, +\infty\right[$, and let $g\colon \mathcal{G} \to \mathbb{R}$ be convex and β-Lipschitzian. Then the following hold:*

(i) $0 \leqslant g \circ L - L \overset{\gamma}{\blacklozenge} g \leqslant \gamma\beta^2/2$.

(ii) $L^* \triangleright g^* \leqslant L \overset{1/\gamma}{\diamond} g^* \leqslant (L^* \triangleright g^*) + \gamma\beta^2/2$.

(iii) *Let $x \in \mathcal{H}$. Then* $\| \operatorname{prox}_{\gamma g \circ L} x - \operatorname{prox}_{\gamma(L \overset{\gamma}{\blacklozenge} g)} x \| \leqslant \gamma\beta$.

Proof. We recall that a lower semicontinuous convex function $f\colon \mathcal{H} \to \mathbb{R}$ is β-Lipschitzian if and only if $\operatorname{ran} \partial f = \operatorname{dom} \partial f^* \subset B(0;\beta)$ [2, Corollary 17.19]. Moreover, since $\operatorname{dom} g = \mathcal{G}$, we have $\operatorname{dom} \partial g = \mathcal{G}$ [2, Proposition 16.27].

(i): Let $x \in \mathcal{H}$ and set $\Phi = \mathscr{Q}_{\mathcal{G}} - \mathscr{Q}_{\mathcal{H}} \circ L^*$. Since $\operatorname{dom} \partial g = \mathcal{G}$, there exists $y^* \in \partial g(Lx) \subset \operatorname{ran} \partial g \subset B(0;\beta)$. Thus, $\Phi(y^*) \leqslant \mathscr{Q}_{\mathcal{G}}(y^*) \leqslant \beta^2/2$ and the result follows from Proposition 2.32(i).

(ii): The leftmost inequality follows from Proposition 2.29(i) and Lemma 2.2(iii). On the other hand, Proposition 2.16(i) implies that $L \overset{1/\gamma}{\diamond} g^* \in \Gamma_0(\mathcal{H})$. Additionally, in view of Lemma 2.2(ii) and (i), $(L \overset{\gamma}{\blacklozenge} g)^* \leqslant (g \circ L - \gamma\beta^2/2)^*$. Finally, we deduce from Proposition 2.16(ii) and [2, Proposition 13.24(v)] that

$$L \overset{1/\gamma}{\diamond} g^* = \big(L \overset{\gamma}{\blacklozenge} g\big)^* \leqslant \left(g \circ L - \frac{\gamma\beta^2}{2}\right)^* = \big(g \circ L\big)^* + \frac{\gamma\beta^2}{2} \leqslant \big(L^* \triangleright g^*\big) + \frac{\gamma\beta^2}{2}. \tag{2.51}$$

(iii): Set $\Psi = \operatorname{Id}_{\mathcal{G}} - L \circ L^*$, $p_\gamma = \operatorname{prox}_{L \blacklozenge (\gamma g)} x$, and $p = \operatorname{prox}_{\gamma g \circ L} x$. We note that, since $\|L\| \leqslant 1$,

$$\Psi \text{ is monotone and } \|\Psi\| \leqslant 1. \tag{2.52}$$

Next, we deduce from [2, Proposition 16.44] and Proposition 2.20(ii) that there exists $y_\gamma \in (\partial(\gamma g)^* + \Psi)^{-1}(Lp_\gamma)$ such that $L^*y_\gamma = x - p_\gamma$. Thus, $y_\gamma \in \gamma\partial g(Lp_\gamma - \Psi y_\gamma)$. On the other hand, by [2, Proposition 16.44 and Corollary 16.53(i)], there exists $y \in \gamma\partial g(Lp)$ such that $L^*y = x - p$. Therefore, the monotonicity of $\gamma\partial g$ [2, Theorem 20.25] entails that

$$\langle (Lp_\gamma - \Psi y_\gamma) - Lp \mid y_\gamma - y \rangle_{\mathcal{G}} \geqslant 0 \tag{2.53}$$

However, by (2.52), the Cauchy–Schwarz inequality, and the fact that $\{y_\gamma, y\} \subset \gamma \operatorname{ran} \partial g \subset B(0;\gamma\beta)$ we derive that

$$\begin{aligned}
\langle (Lp_\gamma - \Psi y_\gamma) - Lp \mid y_\gamma - y \rangle_{\mathcal{G}} \geqslant 0 &\Leftrightarrow \langle p_\gamma - p \mid L^*(y_\gamma - y) \rangle_{\mathcal{H}} - \langle \Psi y_\gamma \mid y_\gamma - y \rangle_{\mathcal{G}} \geqslant 0 \\
&\Leftrightarrow \|p - p_\gamma\|_{\mathcal{H}}^2 \leqslant \langle \Psi y_\gamma \mid y \rangle_{\mathcal{G}} - \langle \Psi y_\gamma \mid y_\gamma \rangle_{\mathcal{G}} \\
&\Rightarrow \|p - p_\gamma\|_{\mathcal{H}}^2 \leqslant \langle \Psi y_\gamma \mid y \rangle_{\mathcal{G}}
\end{aligned}$$

$$
\begin{aligned}
&\Rightarrow \|p-p_\gamma\|_{\mathcal{H}}^2 \leqslant \|\Psi\|\,\|y_\gamma\|_{\mathcal{G}}\,\|y\|_{\mathcal{G}}\\
&\Rightarrow \|p-p_\gamma\|_{\mathcal{H}}^2 \leqslant (\gamma\beta)^2\\
&\Leftrightarrow \|p-p_\gamma\|_{\mathcal{H}} \leqslant \gamma\beta.
\end{aligned}
\tag{2.54}
$$

Since, in view of Proposition 2.10(vii), $\operatorname{prox}_{\gamma(L\overset{\gamma}{\blacklozenge}g)} x = p_\gamma$, the proof is complete. ☐

Example 2.34 Let $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$, let $g \in \Gamma_0(\mathcal{G})$, let $\gamma \in \left]0,+\infty\right[$, and let $\rho \in \left]0,+\infty\right[$. Suppose that $L \circ L^* = \rho \mathrm{Id}_{\mathcal{G}}$. Then the following hold:

(i) Set $h = g(\sqrt{\rho}\cdot)$ and $S = L/\sqrt{\rho}$. Then $g \circ L = S \overset{\gamma}{\blacklozenge} h$.

(ii) $\operatorname{prox}_{\gamma g\circ L} = \mathrm{Id}_{\mathcal{H}} + \rho^{-1}L^* \circ (\operatorname{prox}_{\gamma\rho g} - \mathrm{Id}_{\mathcal{G}}) \circ L$.

Proof. (i): Since $L \circ L^* = \rho \mathrm{Id}_{\mathcal{G}}$, S is a coisometry, and we deduce from Proposition 2.29(v) and Lemma 2.5(ii) that $S \overset{\gamma}{\blacklozenge} h = h \circ S = g \circ L$.

(ii): This follows from (i) and Proposition 2.19(ii) (see also [2, Proposition 24.14]). ☐

Example 2.35 Let V be a closed vector subspace of $\mathcal{H}$ and $\gamma \in \left]0,+\infty\right[$. Then the following hold:

(i) $\operatorname{proj}_V \overset{\gamma}{\diamond} \|\cdot\| = \iota_V + \|\cdot\|$.

(ii) $\operatorname{proj}_V \overset{\gamma}{\blacklozenge} \|\cdot\| = \|\cdot\| \circ \operatorname{proj}_V$.

Proof. Set $\Phi = \mathscr{Q}_{\mathcal{H}} - \mathscr{Q}_{\mathcal{H}} \circ \operatorname{proj}_V$ and let $x \in \mathcal{H}$.

(i): It follows from Proposition 2.11(i), Lemma 2.5(ii), and the identity $\Phi = \mathscr{Q}_{\mathcal{H}} \circ \operatorname{proj}_{V^\perp}$ that

$$
\Big(\operatorname{proj}_V \overset{\gamma}{\diamond} \|\cdot\|\Big)(x) = \inf_{\substack{y\in\mathcal{H}\\ \operatorname{proj}_V y = x}} \left(\|y\| + \frac{1}{2\gamma}\|x-y\|^2\right) = \begin{cases} \|x\|, & \text{if } x \in V\\ +\infty, & \text{if } x \notin V\end{cases} = \iota_V(x) + \|x\|.
\tag{2.55}
$$

(ii): We recall that $\partial\|\cdot\|(x) = \{x/\|x\|\}$ if $x \neq 0$ and that $\partial\|\cdot\|(0) = B(0;1)$ [2, Example 16.32]. Hence,

$$
\begin{aligned}
\operatorname{proj}_{V^\perp}\big(\partial\|\cdot\|(\operatorname{proj}_V x)\big) &= \begin{cases} \big\{\operatorname{proj}_{V^\perp}\big(\operatorname{proj}_V x/\|\operatorname{proj}_V x\|\big)\big\}, & \text{if } \operatorname{proj}_V x \neq 0;\\ \operatorname{proj}_{V^\perp}\big(B(0;1)\big), & \text{if } \operatorname{proj}_V x = 0\end{cases}\\
&= \begin{cases} \{0\}, & \text{if } x \notin V^\perp;\\ \operatorname{proj}_{V^\perp}\big(B(0;1)\big), & \text{if } x \in V^\perp\end{cases}\\
&\ni 0.
\end{aligned}
\tag{2.56}
$$

However, $\mathrm{Id} - \operatorname{proj}_V \circ \operatorname{proj}_V^* = \operatorname{proj}_{V^\perp}$. Therefore, in view of Proposition 2.32(ii), this confirms that $\operatorname{proj}_V \overset{\gamma}{\blacklozenge} \|\cdot\| = \|\cdot\| \circ \operatorname{proj}_V$. ☐

Remark 2.36 In contrast with Proposition 2.29(v), Example 2.35(ii) shows an instance in which the proximal cocomposition coincides with the standard composition for a linear operator which is not a coisometry.

2.2.3.4 Asymptotic properties

We investigate the asymptotic properties of the families $(L \overset{\gamma}{\diamond} g)_{\gamma\in]0,+\infty[}$ and $(L \overset{\gamma}{\blacklozenge} g)_{\gamma\in]0,+\infty[}$ as γ varies. These results provide further connections between the compositions (2.3), (2.4), and the proximal compositions of Definition 2.1.

Proposition 2.37 *Suppose that $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ satisfies $0 < \|L\| \leqslant 1$ and let $g\colon \mathcal{G} \to]-\infty,+\infty]$ be a proper function such that* $\operatorname{cam} g \neq \varnothing$. *Suppose that $x \in L^{-1}(\operatorname{dom} g^{**})$ and set, for every $\gamma \in]0,+\infty[$, $x_\gamma = \operatorname{prox}_{\gamma(L \overset{\gamma}{\blacklozenge} g)} x$. Then*

$$\lim_{0<\gamma\to 0} \big(L \overset{\gamma}{\blacklozenge} g\big)(x_\gamma) = g^{**}(Lx). \tag{2.57}$$

Proof. We first observe that, by virtue of Proposition 2.16(i), $(x_\gamma)_{\gamma\in]0,+\infty[}$ is well defined. Appealing to Proposition 2.22(ii), we get

$$\big(L \overset{\gamma}{\blacklozenge} g\big)(x_\gamma) + \frac{1}{\gamma}\mathscr{Q}_{\mathcal{H}}(x - x_\gamma) = {}^{\gamma}\big(L \overset{\gamma}{\blacklozenge} g\big)(x) = {}^{\gamma}\big(g^{**}\big)(Lx). \tag{2.58}$$

On the other hand, by Proposition 2.19(ii),

$$\frac{1}{\gamma}\mathscr{Q}_{\mathcal{H}}(x - x_\gamma) = \frac{1}{\gamma}\mathscr{Q}_{\mathcal{H}}\Big(L^*\big(Lx - \operatorname{prox}_{\gamma g^{**}}(Lx)\big)\Big) \leqslant \frac{1}{\gamma}\|L\|^2\mathscr{Q}_{\mathcal{G}}\big(Lx - \operatorname{prox}_{\gamma g^{**}}(Lx)\big). \tag{2.59}$$

Therefore, since $Lx \in \operatorname{dom} g^{**}$, [2, Proposition 12.33(iii)] implies that $\gamma^{-1}\mathscr{Q}_{\mathcal{H}}(x - x_\gamma) \to 0$ as $\gamma \to 0$. Finally, by (2.58) and [2, Proposition 12.33(ii)],

$$\lim_{0<\gamma\to 0} \big(L \overset{\gamma}{\blacklozenge} g\big)(x_\gamma) = \lim_{0<\gamma\to 0} {}^{\gamma}\big(g^{**}\big)(Lx) = \big(g^{**}\big)(Lx), \tag{2.60}$$

as claimed. ☐

Theorem 2.38 *Suppose that $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ satisfies $0 < \|L\| \leqslant 1$, let $g\colon \mathcal{G} \to]-\infty,+\infty]$ be a proper function such that* $\operatorname{cam} g \neq \varnothing$, *and let $x \in \mathcal{H}$. Then the following hold:*

(i) *The function $]0,+\infty[\to]-\infty,+\infty]\colon \gamma \mapsto (L \overset{\gamma}{\diamond} g)(x)$ is decreasing.*

(ii) *The function $]0,+\infty[\to]-\infty,+\infty]\colon \gamma \mapsto (L \overset{\gamma}{\blacklozenge} g)(x)$ is decreasing.*

(iii) $\displaystyle\lim_{\gamma\to+\infty} (L \overset{\gamma}{\diamond} g)(x) = (L^* \triangleright g^{**})(x)$.

(iv) $\displaystyle\lim_{0<\gamma\to 0} (L \overset{\gamma}{\blacklozenge} g)(x) = g^{**}(Lx)$.

(v) *Suppose that $\|L\| < 1$. Then* $\displaystyle\lim_{\gamma\to+\infty} (L \overset{\gamma}{\blacklozenge} g)(x) = \inf_{y\in\mathcal{G}} g^{**}(y)$.

(vi) *Suppose that* $\|L\| = 1$ *and that* $V = \operatorname{ran}(\mathrm{Id}_{\mathcal{G}} - L \circ L^*)$ *is closed. Then* $\lim_{\gamma \to +\infty} (L \overset{\gamma}{\blacklozenge} g)(x) = \inf_{y \in Lx - V} g^{**}(y)$.

Proof. Set $\Phi = \mathscr{Q}_{\mathcal{G}} - \mathscr{Q}_{\mathcal{H}} \circ L^*$.

(i): Fix $\gamma_1 \in \left]0, +\infty\right[$ and $\gamma_2 \in \left]0, +\infty\right[$ such that $\gamma_1 \leqslant \gamma_2$. Then we deduce from Proposition 2.11(i) that

$$L \overset{\gamma_2}{\diamond} g = L^* \triangleright \big(g^{**} + \Phi/\gamma_2\big) \leqslant L^* \triangleright \big(g^{**} + \Phi/\gamma_1\big) = L \overset{\gamma_1}{\diamond} g. \tag{2.61}$$

(ii): Fix $\gamma_1 \in \left]0, +\infty\right[$ and $\gamma_2 \in \left]0, +\infty\right[$ such that $\gamma_1 \leqslant \gamma_2$. By (i), $L \overset{1/\gamma_1}{\diamond} g^* \leqslant L \overset{1/\gamma_2}{\diamond} g^*$. Therefore, appealing to Definition 2.1 and Lemma 2.2(ii), we get

$$L \overset{\gamma_2}{\blacklozenge} g = \big(L \overset{1/\gamma_2}{\diamond} g^*\big)^* \leqslant \big(L \overset{1/\gamma_1}{\diamond} g^*\big)^* = L \overset{\gamma_1}{\blacklozenge} g. \tag{2.62}$$

(iii): Since $\Phi \geqslant 0$, it follows from (i) and Proposition 2.11(i) that

$$\begin{aligned}
\lim_{\gamma \to +\infty} \big(L \overset{\gamma}{\diamond} g\big)(x) &= \inf_{\gamma \in]0, +\infty[} \left(L^* \triangleright \left(g^{**} + \frac{1}{\gamma}\Phi\right)\right)(x) \\
&= \inf_{\gamma \in]0, +\infty[} \left(\inf_{\substack{y \in \mathcal{G} \\ L^* y = x}} \left(g^{**}(y) + \frac{1}{\gamma}\Phi(y)\right)\right) \\
&= \inf_{\substack{y \in \mathcal{G} \\ L^* y = x}} \left(\inf_{\gamma \in]0, +\infty[} \left(g^{**}(y) + \frac{1}{\gamma}\Phi(y)\right)\right) \\
&= \inf_{\substack{y \in \mathcal{G} \\ L^* y = x}} g^{**}(y) \\
&= \big(L^* \triangleright g^{**}\big)(x).
\end{aligned} \tag{2.63}$$

(iv): By [2, Proposition 12.33(ii)], ${}^{\gamma}\big(g^{**}\big) \to g^{**}$ as $0 < \gamma \to 0$. The claim therefore follows from Proposition 2.29(ii).

(v)–(vi): As in the proof of Proposition 2.11(iv), $(g^* + \gamma\Phi)^* = g^{**} \,\square\, (\Phi^*/\gamma)$. Thus, it follows from Proposition 2.11(ii) that

$$L \overset{\gamma}{\blacklozenge} g = \Big(g^{**} \,\square\, \big(\Phi^*/\gamma\big)\Big) \circ L. \tag{2.64}$$

Moreover, since $\Phi \leqslant \mathscr{Q}_{\mathcal{G}}$, Lemma 2.2(ii) yields $\mathscr{Q}_{\mathcal{G}} \leqslant \Phi^*$. Altogether, using (ii) and (2.64), we obtain

$$\lim_{\gamma \to +\infty} \big(L \overset{\gamma}{\blacklozenge} g\big)(x) = \inf_{\gamma \in]0, +\infty[} \left(g^{**} \,\square\, \frac{\Phi^*}{\gamma}\right)(Lx)$$

$$= \inf_{\gamma\in]0,+\infty[} \left(\inf_{y\in\mathcal{G}} \left(g^{**}(y) + \frac{1}{\gamma}\Phi^*(Lx-y) \right) \right)$$
$$= \inf_{y\in Lx-\operatorname{dom}\Phi^*} \left(\inf_{\gamma\in]0,+\infty[} \left(g^{**}(y) + \frac{1}{\gamma}\Phi^*(Lx-y) \right) \right)$$
$$= \inf_{y\in Lx-\operatorname{dom}\Phi^*} g^{**}(y). \tag{2.65}$$

We set $A = \mathrm{Id}_{\mathcal{G}} - L\circ L^*$ and observe that $\Phi\colon y\mapsto \langle y \mid Ay\rangle_{\mathcal{G}}/2$. In case (v), since $\|L\| < 1$, A is invertible and Lemma 2.9 asserts that $\operatorname{dom}\Phi^* = \operatorname{ran} A = \mathcal{G}$ in (2.65). Finally, case (vi) follows from Lemma 2.9 and (2.65). ☐

Corollary 2.39 *Suppose that $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ is an isometry, let $g\in\Gamma_0(\mathcal{G})$, and let $x\in\mathcal{H}$. Then the following hold:*

(i) $\displaystyle\lim_{\gamma\to+\infty} (L \overset{\gamma}{\diamond} g)(x) = (L^* \triangleright g)(x)$.

(ii) $\displaystyle\lim_{0<\gamma\to 0} (L \overset{\gamma}{\diamond} g)(x) = g(Lx)$.

Proof. By Proposition 2.29(iv), $L \overset{\gamma}{\blacklozenge} g = L \overset{\gamma}{\diamond} g$, whereas Lemma 2.5(ii) yields $g^{**} = g$.

(i): A consequence of Theorem 2.38(iii).

(ii): A consequence of Theorem 2.38(iv). ☐

Example 2.40 Let $V \neq \{0\}$ be a closed vector subspace of $\mathcal{G}$, let $g\in\Gamma_0(\mathcal{G})$, and let $x\in\mathcal{G}$. Then

$$\lim_{\gamma\to+\infty} \big(\mathrm{proj}_V \overset{\gamma}{\blacklozenge} g\big)(x) = \inf_{v\in V^\perp} g(x+v). \tag{2.66}$$

Proof. Since $\|\mathrm{proj}_V\| = 1$ and $\operatorname{ran}(\mathrm{Id}_{\mathcal{G}} - \mathrm{proj}_V\circ \mathrm{proj}_V^*) = V^\perp$, it follows from Theorem 2.38(vi) and Lemma 2.5(ii) that

$$\lim_{\gamma\to+\infty} \big(\mathrm{proj}_V \overset{\gamma}{\blacklozenge} g\big)(x) = \inf_{y\in \mathrm{proj}_V x - V^\perp} g(y) = \inf_{y\in x+V^\perp} g(y) = \inf_{v\in V^\perp} g(x+v), \tag{2.67}$$

as announced. ☐

We now turn our attention to epi-convergence. As discussed in [1], this notion plays a central role in the approximation of variational problems. It will allow us to connect asymptotically the proximal composition to the infimal postcomposition, and the proximal cocomposition to the standard composition as γ evolves.

Definition 2.41 ([1, Chapter 1], [17, Chapter 7]) Suppose that $\mathcal{H}$ is finite-dimensional, and let $(f_n)_{n\in\mathbb{N}}$ and f be functions from $\mathcal{H}$ to $[-\infty,+\infty]$. We say that $(f_n)_{n\in\mathbb{N}}$ *epi-converges* to f, in symbols $f_n \overset{e}{\to} f$, if the following hold for every $x\in\mathcal{H}$:

(i) For every sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathcal{H}$ such that $x_n\to x$, $f(x) \leqslant \underline{\lim}\, f_n(x_n)$.

(ii) There exists a sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathcal{H}$ such that $x_n\to x$ and $\overline{\lim}\, f_n(x_n) \leqslant f(x)$.

The *epi-topology* is the topology induced by epi-convergence.

Lemma 2.42 *Suppose that $\mathcal{H}$ and $\mathcal{G}$ are finite-dimensional, let $(L_n)_{n\in\mathbb{N}}$ and L be operators in $\mathcal{B}(\mathcal{H},\mathcal{G})$, let $(g_n)_{n\in\mathbb{N}}$ and g be functions in $\Gamma_0(\mathcal{G})$, and let $(\gamma_n)_{n\in\mathbb{N}}$ and γ be reals in $]0,+\infty[$. Suppose that $L_n \to L$, $g_n \overset{e}{\to} g$, and $\gamma_n \to \gamma$. Then the following hold:*

(i) $\gamma_n g_n \overset{e}{\to} \gamma g$.

(ii) $g_n^* \overset{e}{\to} g^*$.

(iii) *Suppose that $h\colon \mathcal{G} \to \mathbb{R}$ is continuous. Then $g_n + \gamma_n h \overset{e}{\to} g + \gamma h$.*

(iv) *Suppose that $0 \in \operatorname{int}(\operatorname{dom} g - \operatorname{ran} L)$. Then $g_n \circ L_n \overset{e}{\to} g \circ L$.*

Proof. (i): [17, Exercise 7.8(d)].

(ii): [17, Theorem 11.34].

(iii): It follows from (i) and [17, Exercise 7.8(a)] that $g_n/\gamma_n + h \overset{e}{\to} g/\gamma + h$. Invoking (i) once more, we obtain $g_n + \gamma_n h = \gamma_n(g_n/\gamma_n + h) \overset{e}{\to} \gamma(g/\gamma + h) = g + \gamma h$.

(iv): [17, Exercise 7.47(a)]. ☐

Theorem 2.43 *Suppose that $\mathcal{H}$ and $\mathcal{G}$ are finite-dimensional, let $(L_n)_{n\in\mathbb{N}}$ and L be operators in $\mathcal{B}(\mathcal{H},\mathcal{G})$, let $(g_n)_{n\in\mathbb{N}}$ and g be functions in $\Gamma_0(\mathcal{G})$, and let $(\gamma_n)_{n\in\mathbb{N}}$ and γ be reals in $]0,+\infty[$. Then the following hold:*

(i) *Suppose that $L_n \to L$, $g_n \overset{e}{\to} g$, and $\gamma_n \to \gamma$. Then the following are satisfied:*

  (a) $L_n \overset{\gamma_n}{\diamond} g_n \overset{e}{\to} L \overset{\gamma}{\diamond} g$.

  (b) $L_n \overset{\gamma_n}{\blacklozenge} g_n \overset{e}{\to} L \overset{\gamma}{\blacklozenge} g$.

(ii) *Suppose that $0 < \|L\| \leqslant 1$. Then the following are satisfied:*

  (a) *Suppose that $\gamma_n \uparrow +\infty$. Then $L \overset{\gamma_n}{\diamond} g \overset{e}{\to} (L^* \triangleright g)^{\smallsmile}$.*

  (b) *Suppose that $\gamma_n \downarrow 0$. Then $L \overset{\gamma_n}{\blacklozenge} g \overset{e}{\to} g \circ L$.*

Proof. (i)(a): It follows from Lemmas 2.5(viii) and 2.42(ii)–(iii) that

$$
{}^{\frac{1}{\gamma_n}}(g_n^*) = \Big(g_n + \frac{1}{\gamma_n}\mathcal{Q}_{\mathcal{G}}\Big)^* \overset{e}{\to} \Big(g + \frac{1}{\gamma}\mathcal{Q}_{\mathcal{G}}\Big)^* = {}^{\frac{1}{\gamma}}(g^*). \tag{2.68}
$$

Since Lemmas 2.2(v) and 2.5(vi) yield $\operatorname{dom}{}^{\frac{1}{\gamma}}(g^*) = \mathcal{G}$, Lemma 2.42(iv) and (2.68) imply that ${}^{\frac{1}{\gamma_n}}(g_n^*) \circ L_n \overset{e}{\to} {}^{\frac{1}{\gamma}}(g^*) \circ L$. Finally, appealing to Definition 2.1 and Lemma 2.42(ii)–(iii), we conclude that

$$
L_n \overset{\gamma_n}{\diamond} g_n = \Big({}^{\frac{1}{\gamma_n}}(g_n^*) \circ L_n\Big)^* - \frac{1}{\gamma_n}\mathcal{Q}_{\mathcal{H}} \overset{e}{\to} \Big({}^{\frac{1}{\gamma}}(g^*) \circ L\Big)^* - \frac{1}{\gamma}\mathcal{Q}_{\mathcal{H}} = L \overset{\gamma}{\diamond} g. \tag{2.69}
$$

(i)(b): By Lemma 2.42(ii), $g_n^* \overset{e}{\to} g^*$. Therefore, upon combining (i)(a) and Lemma 2.42(ii), we obtain

$$
L_n \overset{\gamma_n}{\blacklozenge} g_n = \big(L_n \overset{1/\gamma_n}{\diamond} g_n^*\big)^* \overset{e}{\to} \big(L \overset{1/\gamma}{\diamond} g^*\big)^* = L \overset{\gamma}{\blacklozenge} g. \tag{2.70}
$$

(ii)(a): Set $f = L^* \triangleright g$ and $(\forall n \in \mathbb{N})$ $f_n = L \overset{\gamma_n}{\diamond} g$. It follows from items (i) and (iii) in Theorem 2.38, as well as Lemma 2.5(ii), that $(f_n)_{n\in\mathbb{N}}$ is decreasing and pointwise convergent to f as $n \to +\infty$. Further, since f is convex by [2, Proposition 12.36(ii)], we deduce from [17, Proposition 7.4(c)] and [2, Corollary 9.10] that

$$f_n \xrightarrow{e} \overline{\inf_{n\in\mathbb{N}} f_n} = \overline{f} = \breve{f}. \tag{2.71}$$

(ii)(b): Set $f = g \circ L$ and $(\forall n \in \mathbb{N})$ $f_n = L \overset{\gamma_n}{\blacklozenge} g$. Since $(\gamma_n)_{n\in\mathbb{N}}$ is decreasing, $(f_n)_{n\in\mathbb{N}}$ is increasing by Theorem 2.38(ii). Further, Theorem 2.38(iv) and Lemma 2.5(ii) imply that $(f_n)_{n\in\mathbb{N}}$ converges pointwise to f as $n \to +\infty$. On the other hand, Proposition 2.16(i) implies that $(\forall n \in \mathbb{N})$ $\overline{f_n} = f_n$. Therefore, by virtue of [17, Proposition 7.4(d)],

$$f_n \xrightarrow{e} \sup_{n\in\mathbb{N}} \overline{f_n} = \sup_{n\in\mathbb{N}} f_n = f, \tag{2.72}$$

which concludes the proof. ☐

Corollary 2.44 *Suppose that $\mathcal{H}$ and $\mathcal{G}$ are finite-dimensional, let $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$, let $g \in \Gamma_0(\mathcal{G})$, and let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $]0,+\infty[$. Suppose that L is an isometry and that $(\operatorname{ri}\operatorname{dom} g^*) \cap (\operatorname{ran} L) \neq \varnothing$. Then the following hold:*

(i) *Suppose that $\gamma_n \uparrow +\infty$. Then $L \overset{\gamma_n}{\diamond} g \xrightarrow{e} L^* \triangleright g$.*

(ii) *Suppose that $\gamma_n \downarrow 0$. Then $L \overset{\gamma_n}{\diamond} g \xrightarrow{e} g \circ L$.*

(iii) *For every $t \in [0,1]$, set $\gamma_t = \tan(\pi t/2)$. Then the operator*

$$T\colon [0,1] \to \Gamma_0(\mathcal{H})\colon t \to \begin{cases} g \circ L, & \text{if } t = 0; \\ L \overset{\gamma_t}{\diamond} g, & \text{if } 0 < t < 1; \\ L^* \triangleright g, & \text{if } t = 1 \end{cases} \tag{2.73}$$

is continuous with respect to the epi-topology.

Proof. Proposition 2.29(iv) yields $(\forall \gamma \in]0,+\infty[)$ $L \overset{\gamma}{\blacklozenge} g = L \overset{\gamma}{\diamond} g$. Further, [2, Proposition 6.19(x)] implies that $0 \in \operatorname{sri}(\operatorname{dom} g^* - \operatorname{ran} L)$. Therefore, by virtue of Lemmas 2.6(iii) and 2.5(ii), we get $L^* \triangleright g \in \Gamma_0(\mathcal{H})$.

(i): A consequence of Theorem 2.43(ii)(a).

(ii): See Theorem 2.43(ii)(b).

(iii): Theorem 2.43(i)(a) guarantees the epi-continuity of T on $]0,1[$. Finally, (i) and (ii) imply that $\lim_{0<t\to 0} T(t) = T(0)$ and $\lim_{1>t\to 1} T(t) = T(1)$, respectively. ☐

Remark 2.45 Suppose that $\mathcal{H}$ and $\mathcal{G}$ are finite-dimensional and that $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ satisfies $0 < \|L\| \leqslant 1$, let $g \in \Gamma_0(\mathcal{G})$, and let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $]0,+\infty[$. Under a qualification condition (see Lemma 2.6(iii)), $L^* \triangleright g \in \Gamma_0(\mathcal{H})$ and, consequently, $L^* \triangleright g = (L^* \triangleright g)^{\smile}$. In this case, Theorem 2.38(iii) and Theorem 2.43(ii)(a) show that the proximal composition converges pointwise and epi-converges to the infimal postcomposition as $\gamma_n \uparrow +\infty$. On the other hand, Theorem 2.38(iv) and Theorem 2.43(ii)(b) show that the proximal cocomposition converges pointwise and epi-converges to the standard composition. Further, in the particular case in which $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ is an isometry, Corollary 2.44(iii) asserts that $g \circ L$ and $L^* \triangleright g$ are homotopic via the proximal composition with respect to the epi-topology.

Proposition 2.46 *Suppose that $\mathcal{H}$ and $\mathcal{G}$ are finite-dimensional and that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ satisfies $0 < \|L\| \leqslant 1$, let $g \in \Gamma_0(\mathcal{G})$, and let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $]0, +\infty[$ such that $\gamma_n \downarrow 0$. Suppose that $\operatorname{dom} g \cap \operatorname{ran} L \neq \varnothing$ and that $g \circ L$ is coercive. Then the following hold:*

(i) $\inf_{x\in\mathcal{H}}(L \overset{\gamma_n}{\blacklozenge} g)(x) \to \min_{x\in\mathcal{H}} g(Lx)$.

(ii) *There exists $N \subset \mathbb{N}$ such that $\mathbb{N} \smallsetminus N$ is finite and $(\forall n \in N)$ $\operatorname{Argmin}(L \overset{\gamma_n}{\blacklozenge} g) \neq \varnothing$. Further,*

$$\overline{\lim}\operatorname{Argmin}\big(L \overset{\gamma_n}{\blacklozenge} g\big) \subset \operatorname{Argmin}\big(g \circ L\big). \tag{2.74}$$

Proof. Set $f = g \circ L$ and $(\forall n \in \mathbb{N})$ $f_n = L \overset{\gamma_n}{\blacklozenge} g$. Since $\operatorname{dom} g \cap \operatorname{ran} L \neq \varnothing$, $f \in \Gamma_0(\mathcal{H})$. Thus, by [2, Proposition 11.15(i)], f has a minimizer over $\mathcal{H}$. Further, by Proposition 2.16(i), for every $n \in \mathbb{N}$, $f_n \in \Gamma_0(\mathcal{H})$ and, by Theorem 2.43(ii)(b), $f_n \overset{e}{\to} f$. On the other hand, [2, Proposition 11.12] asserts that the lower level sets $(\operatorname{lev}_{\leqslant\xi} f)_{\xi\in\mathbb{R}}$ are bounded. Altogether, by virtue of [17, Exercise 7.32(c)], for every $\xi \in \mathbb{R}$, there exists $N_\xi \in \mathbb{N}$ such that $\bigcup_{n\geqslant N_\xi} \operatorname{lev}_{\leqslant\xi} f_n$ is bounded.

(i)–(ii): A consequence of [17, Theorem 7.33]. ☐

2.2.4 Integral proximal mixtures

2.2.4.1 Definition and mathematical setting

Integral proximal mixtures were introduced in [7] as a tool to combine arbitrary families of convex functions and linear operators in such a way that the proximity operator of the mixture can be expressed explicitly in terms of the individual proximity operators. They extend the proximal mixtures of [9], which were designed for finite families. In this section, we use the results of Section 2.2.3 to study their variational properties. This investigation is carried out in the same framework as in [7], which hinges on the following assumptions. Henceforth, we adopt the customary convention that the integral of an $\mathcal{F}$-measurable function $\vartheta\colon \Omega \to [-\infty, +\infty]$ is the usual Lebesgue integral $\int_\Omega \vartheta d\mu$, except when the Lebesgue integral $\int_\Omega \max\{\vartheta, 0\} d\mu$ is $+\infty$, in which case $\int_\Omega \vartheta d\mu = +\infty$.

Assumption 2.47 *Let $(\Omega, \mathcal{F}, \mu)$ be a complete σ-finite measure space, let $(\mathsf{G}_\omega)_{\omega\in\Omega}$ be a family of real Hilbert spaces, and let $\prod_{\omega\in\Omega} \mathsf{G}_\omega$ be the usual real vector space of mappings x defined on Ω*

such that $(\forall\omega\in\Omega)\ x(\omega)\in\mathsf{G}_\omega$. *Let* $((\mathsf{G}_\omega)_{\omega\in\Omega},\mathfrak{G})$ *be an* $\mathcal{F}$*-measurable vector field of real Hilbert spaces, that is,* $\mathfrak{G}$ *is a vector subspace of* $\prod_{\omega\in\Omega}\mathsf{G}_\omega$ *which satisfies the following:*

[A] *For every* $x\in\mathfrak{G}$*, the function* $\Omega\to\mathbb{R}\colon\omega\mapsto\|x(\omega)\|_{\mathsf{G}_\omega}$ *is* $\mathcal{F}$*-measurable.*

[B] *For every* $x\in\prod_{\omega\in\Omega}\mathsf{G}_\omega$*,*

$$\big[\ (\forall y\in\mathfrak{G})\ \ \Omega\to\mathbb{R}\colon\omega\mapsto\langle x(\omega)\mid y(\omega)\rangle_{\mathsf{G}_\omega}\ \text{is } \mathcal{F}\text{-measurable}\ \big]\quad\Rightarrow\quad x\in\mathfrak{G}. \tag{2.75}$$

[C] *There exists a sequence* $(e_n)_{n\in\mathbb{N}}$ *in* $\mathfrak{G}$ *such that* $(\forall\omega\in\Omega)\ \overline{\operatorname{span}}\{e_n(\omega)\}_{n\in\mathbb{N}}=\mathsf{G}_\omega$.

Set $\mathfrak{H}=\big\{x\in\mathfrak{G}\mid\int_\Omega\|x(\omega)\|^2_{\mathsf{G}_\omega}\mu(d\omega)<+\infty\big\}$*, and let* $\mathcal{G}$ *be the real Hilbert space of equivalence classes of* μ-a.e. *equal mappings in* $\mathfrak{H}$ *equipped with the scalar product*

$$\langle\cdot\mid\cdot\rangle_{\mathcal{G}}\colon\mathcal{G}\times\mathcal{G}\to\mathbb{R}\colon(x,y)\mapsto\int_\Omega\langle x(\omega)\mid y(\omega)\rangle_{\mathsf{G}_\omega}\mu(d\omega), \tag{2.76}$$

where we adopt the common practice of designating by x *both an equivalence class in* $\mathcal{G}$ *and a representative of it in* $\mathfrak{H}$*. We write*

$$\mathcal{G}={}^{\mathfrak{G}}\!\!\int_\Omega^{\oplus}\mathsf{G}_\omega\mu(d\omega) \tag{2.77}$$

and call $\mathcal{G}$ *the Hilbert direct integral of* $((\mathsf{G}_\omega)_{\omega\in\Omega},\mathfrak{G})$ *[13].*

Assumption 2.48 *Assumption 2.47 and the following are in force:*

[A] H *is a separable real Hilbert space.*

[B] *For every* $\omega\in\Omega$*,* $\mathsf{L}_\omega\in\mathcal{B}(\mathsf{H},\mathsf{G}_\omega)$.

[C] *For every* $\mathsf{x}\in\mathsf{H}$*, the mapping* $\mathfrak{e}_{\mathsf{L}}\mathsf{x}\colon\omega\mapsto\mathsf{L}_\omega\mathsf{x}$ *lies in* $\mathfrak{G}$.

[D] $0<\int_\Omega\|\mathsf{L}_\omega\|^2\mu(d\omega)\leqslant 1$.

Given a complete σ-finite measure space $(\Omega,\mathcal{F},\mu)$, a separable real Hilbert space H with Borel σ-algebra $\mathcal{B}_{\mathsf{H}}$, and $p\in[1,+\infty[$, we set

$$\begin{aligned}&\mathscr{L}^p\big(\Omega,\mathcal{F},\mu;\mathsf{H}\big)\\&=\left\{x\colon\Omega\to\mathsf{H}\;\middle|\;x\text{ is }(\mathcal{F},\mathcal{B}_{\mathsf{H}})\text{-measurable and }\int_\Omega\|x(\omega)\|^p_{\mathsf{H}}\,\mu(d\omega)<+\infty\right\}.\end{aligned} \tag{2.78}$$

The Lebesgue integral (also known as the Bochner integral) of $x\in\mathscr{L}^1\big(\Omega,\mathcal{F},\mu;\mathsf{H}\big)$ is denoted by $\int_\Omega x(\omega)\mu(d\omega)$. The space of equivalence classes of μ-a.e. equal mappings in $\mathscr{L}^p(\Omega,\mathcal{F},\mu;\mathsf{H})$ is denoted by $L^p(\Omega,\mathcal{F},\mu;\mathsf{H})$.

Assumption 2.49 *Assumption 2.47 and the following are in force:*

[A] *For every* $\omega\in\Omega$*,* $\mathsf{g}_\omega\colon\mathsf{G}_\omega\to\,]-\infty,+\infty]$ *satisfies* $\operatorname{cam}\mathsf{g}_\omega\neq\varnothing$.

[B] *For every* $x^*\in\mathfrak{H}$*, the mapping* $\omega\mapsto\operatorname{prox}_{\mathsf{g}^*_\omega}x^*(\omega)$ *lies in* $\mathfrak{G}$.

[C] *There exists* $r \in \mathfrak{H}$ *such that the function* $\omega \mapsto \mathsf{g}_\omega(r(\omega))$ *lies in* $\mathscr{L}^1(\Omega, \mathcal{F}, \mu; \mathbb{R})$.

[D] *There exists* $r^* \in \mathfrak{H}$ *such that the function* $\omega \mapsto \mathsf{g}_\omega^*(r^*(\omega))$ *lies in* $\mathscr{L}^1(\Omega, \mathcal{F}, \mu; \mathbb{R})$.

The following construct will also be required.

Definition 2.50 ([6, Definition 1.4]) Suppose that Assumption 2.47 is in force and, for every $\omega \in \Omega$, let $\mathsf{g}_\omega \colon \mathsf{G}_\omega \to [-\infty, +\infty]$. Suppose that, for every $x \in \mathfrak{H}$, the function $\Omega \to [-\infty, +\infty] \colon \omega \mapsto \mathsf{g}_\omega(x(\omega))$ is $\mathcal{F}$-measurable. The *Hilbert direct integral* of the functions $(\mathsf{g}_\omega)_{\omega\in\Omega}$ relative to $\mathfrak{G}$ is

$$\prescript{\mathfrak{G}}{}{\int_\Omega^\oplus} \mathsf{g}_\omega \mu(d\omega) \colon \mathcal{G} \to [-\infty, +\infty] \colon x \mapsto \int_\Omega \mathsf{g}_\omega\big(x(\omega)\big)\mu(d\omega). \tag{2.79}$$

We introduce below parametrized versions of the integral proximal mixtures of [7, Definition 4.2].

Definition 2.51 Suppose that Assumptions 2.48 and 2.49 are in force, and let $\gamma \in \left]0, +\infty\right[$. The *integral proximal mixture* of $(\mathsf{g}_\omega)_{\omega\in\Omega}$ and $(\mathsf{L}_\omega)_{\omega\in\Omega}$ with parameter γ is

$$\overset{\diamond}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega} = \mathsf{h}^* - \frac{1}{\gamma}\mathscr{Q}_\mathsf{H}, \quad \text{where} \quad (\forall \mathsf{x} \in \mathsf{H}) \quad \mathsf{h}(\mathsf{x}) = \int_\Omega \prescript{\frac{1}{\gamma}}{}{(\mathsf{g}_\omega^*)}(\mathsf{L}_\omega \mathsf{x})\mu(d\omega), \tag{2.80}$$

and the *integral proximal comixture* of $(\mathsf{g}_\omega)_{\omega\in\Omega}$ and $(\mathsf{L}_\omega)_{\omega\in\Omega}$ with parameter γ is

$$\overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega} = \left(\overset{\diamond}{\mathsf{M}}_{1/\gamma}\big(\mathsf{L}_\omega, \mathsf{g}_\omega^*\big)_{\omega\in\Omega}\right)^*. \tag{2.81}$$

2.2.4.2 Properties

The following proposition adopts the pattern of [7, Theorem 4.3] by connecting integral proximal mixtures to proximal compositions in the more general context of Definitions 2.1 and 2.51.

Proposition 2.52 *Suppose that Assumptions* 2.48 *and* 2.49 *are in force, and let* $\gamma \in \left]0, +\infty\right[$. *Define*

$$L \colon \mathsf{H} \to \mathcal{G} \colon \mathsf{x} \mapsto \mathfrak{e}_\mathsf{L}\mathsf{x} \tag{2.82}$$

and

$$g = \prescript{\mathfrak{G}}{}{\int_\Omega^\oplus} \mathsf{g}_\omega^{**}\mu(d\omega). \tag{2.83}$$

Then the following hold:

(i) $L \in \mathcal{B}(\mathsf{H}, \mathcal{G})$ *and* $0 < \|L\| \leqslant 1$.

(ii) $L^* \colon \mathcal{G} \to \mathsf{H} \colon x^* \mapsto \int_\Omega \mathsf{L}_\omega^*(x^*(\omega))\mu(d\omega)$.

(iii) $g \in \Gamma_0(\mathcal{G})$.

(iv) $\overset{\diamond}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega} = L \overset{\gamma}{\diamond} g$.

(v) $\overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega} = L \overset{\gamma}{\blacklozenge} g$.

Proof. (i): We deduce from [6, Proposition 3.12(ii)] and Assumption 2.48[D] that $L \in \mathcal{B}(\mathsf{H}, \mathcal{G})$ and that $0 < \|L\|^2 \leqslant \int_\Omega \|\mathsf{L}_\omega\|^2 \mu(d\omega) \leqslant 1$.

(ii): See [6, Proposition 3.12(v)].

To establish (iii)–(v), set $\vartheta\colon \Omega \to \mathbb{R}\colon \omega \mapsto -\mathsf{g}_\omega^{**}(r(\omega))$ and $(\forall \omega \in \Omega)$ $\mathsf{f}_\omega = \mathsf{g}_\omega^*$. Let us show that $(\mathsf{f}_\omega)_{\omega\in\Omega}$ satisfies the following:

[A]' For every $\omega \in \Omega$, $\mathsf{f}_\omega \in \Gamma_0(\mathsf{G}_\omega)$.

[B]' For every $x \in \mathfrak{H}$, the mapping $\omega \mapsto \mathrm{prox}_{\mathsf{f}_\omega}(x(\omega))$ lies in $\mathfrak{G}$.

[C]' The function $\omega \mapsto \mathsf{f}_\omega(r^*(\omega))$ lies in $\mathscr{L}^1(\Omega, \mathcal{F}, \mu; \mathbb{R})$.

[D]' $\vartheta \in \mathscr{L}^1(\Omega, \mathcal{F}, \mu; \mathbb{R})$ and, for every $\omega \in \Omega$, $\mathsf{f}_\omega \geqslant \langle r(\omega) \mid \cdot \rangle_{\mathsf{G}_\omega} + \vartheta(\omega)$.

This will confirm that $(\mathsf{f}_\omega)_{\omega\in\Omega}$ satisfies the properties of [6, Assumption 4.6]. First, it follows from items [A] and [C] in Assumption 2.49 and from Lemma 2.2(v) that [A]' holds. Second, Assumption 2.49[B] implies that [B]' holds, while Assumption 2.49[D] implies that [C]' holds. Let us now show that $\vartheta \in \mathscr{L}^1(\Omega, \mathcal{F}, \mu; \mathbb{R})$. As in the proof of [6, Theorem 4.7(ix)], $-\vartheta$ is $\mathcal{F}$-measurable. Further, by (2.1) and Lemma 2.2(i),

$$(\forall \omega \in \Omega) \quad \langle \cdot \mid r^*(\omega) \rangle_{\mathsf{G}_\omega} - \mathsf{g}_\omega^*\big(r^*(\omega)\big) \leqslant \mathsf{g}_\omega^{**} \leqslant \mathsf{g}_\omega. \tag{2.84}$$

Thus, we infer from Assumption 2.49[C]–[D] that g_ω^{**} is bounded by integrable functions, which shows that

$$\vartheta \in \mathscr{L}^1(\Omega, \mathcal{F}, \mu; \mathbb{R}). \tag{2.85}$$

On the other hand, it follows from Lemma 2.2(iii) and (2.1) that, for every $\omega \in \Omega$, $\mathsf{f}_\omega = \mathsf{g}_\omega^{***} \geqslant \langle r(\omega) \mid \cdot \rangle_{\mathsf{G}_\omega} - \mathsf{g}_\omega^{**}(r(\omega)) = \langle r(\omega) \mid \cdot \rangle_{\mathsf{G}_\omega} + \vartheta(\omega)$, which provides [D]'. Therefore $(\mathsf{f}_\omega)_{\omega\in\Omega}$ satisfies the conclusions of [6, Theorem 4.7]. In particular, [6, Theorem 4.7(i)–(ii)] entail that

$$f = {}^{\mathfrak{G}}\!\!\int_\Omega^{\oplus} \mathsf{f}_\omega \mu(d\omega) \tag{2.86}$$

is a well-defined function in $\Gamma_0(\mathcal{G})$ and from [6, Theorem 4.7(ix)] and Lemma 2.5(ii) that

$$g = f^* \in \Gamma_0(\mathcal{G}). \tag{2.87}$$

(iii): See (2.87).

(iv): By [6, Theorem 4.7(viii)],

$$\tfrac{1}{\gamma} f = {}^{\mathfrak{G}}\!\!\int_\Omega^{\oplus} \tfrac{1}{\gamma} \mathsf{f}_\omega \mu(d\omega). \tag{2.88}$$

Further, by (iii) and Lemma 2.5(ii), $g^* = f$. In turn, (2.82) and (2.88) imply that

$$\tfrac{1}{\gamma}(g^*) \circ L \colon \mathsf{H} \to \mathbb{R} \colon \mathsf{x} \mapsto \int_\Omega \tfrac{1}{\gamma}(\mathsf{g}_\omega^*)(\mathsf{L}_\omega \mathsf{x})\mu(d\omega). \tag{2.89}$$

In view of Definitions 2.1 and 2.51, the assertion is proved.

(v): Let us show that $(\mathsf{f}_\omega)_{\omega\in\Omega}$ fulfills the properties of Assumption 2.49 by showing that the following hold:

[A]” For every $\omega \in \Omega$, $\mathsf{f}_\omega \colon \mathsf{G}_\omega \to \left]-\infty,+\infty\right]$ satisfies $\operatorname{cam} \mathsf{f}_\omega \neq \varnothing$.

[B]” For every $x^* \in \mathfrak{H}$, the mapping $\omega \mapsto \operatorname{prox}_{\mathsf{f}_\omega^*} x^*(\omega)$ lies in $\mathfrak{G}$.

[C]” The function $\omega \mapsto \mathsf{f}_\omega(r^*(\omega))$ lies in $\mathscr{L}^1(\Omega, \mathcal{F}, \mu; \mathbb{R})$.

[D]” The function $\omega \mapsto \mathsf{f}_\omega^*(r(\omega))$ lies in $\mathscr{L}^1(\Omega, \mathcal{F}, \mu; \mathbb{R})$.

We first note that [A]’ and Lemma 2.5(i) imply that [A]” holds, and that [C]’⇔[C]”. Additionally, it follows from (2.85) that [D]” holds. It remains to establish [B]”. Assumption 2.49[B] asserts that, for every $x^* \in \mathfrak{H}$, the mapping $\omega \mapsto \operatorname{prox}_{\mathsf{f}_\omega} x^*(\omega)$ lies in $\mathfrak{G}$. Therefore, the inclusion $\mathfrak{H} \subset \mathfrak{G}$, Lemma 2.5(iv), and the fact the $\mathfrak{G}$ is a vector space imply that, for every $x^* \in \mathfrak{H}$, the mapping $\omega \mapsto \operatorname{prox}_{\mathsf{f}_\omega^*} x^*(\omega) = x^*(\omega) - \operatorname{prox}_{\mathsf{f}_\omega} x^*(\omega)$ lies in $\mathfrak{G}$, which provides [B]”. Hence, we combine Definition 2.51, the application of (iv) to $(\mathsf{f}_\omega)_{\omega\in\Omega}$, (2.87), Lemma 2.5(ii), and Definition 2.1, to obtain

$$\overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega} = \left(\overset{\lozenge}{\mathsf{M}}_{1/\gamma}(\mathsf{L}_\omega, \mathsf{f}_\omega)_{\omega\in\Omega}\right)^* = \left(L \overset{1/\gamma}{\lozenge} f\right)^* = \left(L \overset{1/\gamma}{\lozenge} g^*\right)^* = L \overset{\gamma}{\blacklozenge} g, \tag{2.90}$$

which completes the proof. ☐

Our main results on integral proximal mixtures are the following.

Theorem 2.53 *Suppose that Assumptions 2.48 and 2.49 are in force, and let $\gamma \in \left]0,+\infty\right[$. Then the following hold:*

(i) $\overset{\lozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega} \in \Gamma_0(\mathsf{H})$.

(ii) $\overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega} \in \Gamma_0(\mathsf{H})$.

(iii) $(\overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega})^* = \overset{\lozenge}{\mathsf{M}}_{1/\gamma}(\mathsf{L}_\omega, \mathsf{g}_\omega^*)_{\omega\in\Omega}$.

(iv) $\overset{\lozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega} = (\overset{\blacklozenge}{\mathsf{M}}_{1/\gamma}(\mathsf{L}_\omega, \mathsf{g}_\omega^*)_{\omega\in\Omega})^*$.

(v) *Let* $\mathsf{x} \in \mathsf{H}$. *Then* $\operatorname{prox}_{\gamma \overset{\lozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega}} \mathsf{x} = \int_\Omega \mathsf{L}_\omega^*(\operatorname{prox}_{\gamma \mathsf{g}_\omega^{**}}(\mathsf{L}_\omega \mathsf{x}))\,\mu(d\omega)$.

(vi) *Let* $\mathsf{x} \in \mathsf{H}$. *Then* $\operatorname{prox}_{\gamma \overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega}} \mathsf{x} = \mathsf{x} - \int_\Omega \mathsf{L}_\omega^*(\mathsf{L}_\omega \mathsf{x} - \operatorname{prox}_{\gamma \mathsf{g}_\omega^{**}}(\mathsf{L}_\omega \mathsf{x}))\,\mu(d\omega)$.

(vii) *Define g as in (2.83) and L as in (2.82). Then the following are satisfied:*

(a) $\partial(\overset{\lozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega}) = L^* \rhd (\partial g + (\mathrm{Id}_{\mathcal{G}} - L \circ L^*)/\gamma)$.

(b) $\partial(\overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega}) = L^* \circ (\partial g^* + \gamma(\mathrm{Id}_{\mathcal{G}} - L \circ L^*))^{-1} \circ L$.

(viii) *Let* $\mathsf{x} \in \mathsf{H}$. *Then* ${}^{\gamma}(\overset{\blacklozenge}{\mathsf{M}}_{\gamma}(\mathsf{L}_{\omega}, \mathsf{g}_{\omega})_{\omega\in\Omega})(\mathsf{x}) = \int_{\Omega} {}^{\gamma}(\mathsf{g}_{\omega}^{**})(\mathsf{L}_{\omega}\mathsf{x})\,\mu(d\omega)$.

(ix) $\operatorname{Argmin}_{\mathsf{x}\in\mathsf{H}}(\overset{\blacklozenge}{\mathsf{M}}_{\gamma}(\mathsf{L}_{\omega}, \mathsf{g}_{\omega})_{\omega\in\Omega})(\mathsf{x}) = \operatorname{Argmin}_{\mathsf{x}\in\mathsf{H}} \int_{\Omega} {}^{\gamma}(\mathsf{g}_{\omega}^{**})(\mathsf{L}_{\omega}\mathsf{x})\,\mu(d\omega)$.

(x) *Let* $\mathsf{x} \in \mathsf{H}$. *Then* $(\operatorname{rec}\overset{\blacklozenge}{\mathsf{M}}_{\gamma}(\mathsf{L}_{\omega}, \mathsf{g}_{\omega})_{\omega\in\Omega})(\mathsf{x}) = \int_{\Omega}(\operatorname{rec}(\mathsf{g}_{\omega}^{**}))(\mathsf{L}_{\omega}\mathsf{x})\,\mu(d\omega)$.

(xi) *Suppose that* μ *is a probability measure and that there exists* $\beta \in \left]0, +\infty\right[$ *such that, for every* $\omega \in \Omega$, $\mathsf{g}_{\omega}\colon \mathsf{G}_{\omega} \to \mathbb{R}$ *is convex and* β*-Lipschitzian. Then* $\overset{\blacklozenge}{\mathsf{M}}_{\gamma}(\mathsf{L}_{\omega}, \mathsf{g}_{\omega})_{\omega\in\Omega}$ *is* β*-Lipschitzian.*

Proof. Define L as in (2.82) and g as in (2.83). Recall from items (i) and (iii) in Proposition 2.52 that $L \in \mathcal{B}(\mathsf{H}, \mathcal{G})$, $0 < \|L\| \leqslant 1$, and $g \in \Gamma_0(\mathcal{G})$. Additionally, by Proposition 2.52(iv)–(v),

$$\overset{\diamond}{\mathsf{M}}_{\gamma}(\mathsf{L}_{\omega}, \mathsf{g}_{\omega})_{\omega\in\Omega} = L \overset{\gamma}{\diamond} g \quad \text{and} \quad \overset{\blacklozenge}{\mathsf{M}}_{\gamma}(\mathsf{L}_{\omega}, \mathsf{g}_{\omega})_{\omega\in\Omega} = L \overset{\gamma}{\blacklozenge} g. \tag{2.91}$$

Also, proceeding as in the proof of Proposition 2.52, it can be shown that

$$\big(\mathsf{g}_{\omega}^{**}\big)_{\omega\in\Omega} \text{ satisfies the properties of [6, Assumption 4.6].} \tag{2.92}$$

Thus, by [6, Theorem 4.7(iv)],

$$(\forall x \in \mathcal{G}) \quad \big(\operatorname{prox}_{\gamma g} x\big)(\omega) = \operatorname{prox}_{\gamma \mathsf{g}_{\omega}^{**}}\big(x(\omega)\big) \ \text{ for } \mu\text{-almost every } \omega \in \Omega. \tag{2.93}$$

(i)–(iv): These are consequences of (2.91) and Proposition 2.16.

(v): It follows from (2.91), Propositions 2.19(i) and 2.52(ii), and (2.93) that

$$\operatorname{prox}_{\gamma\overset{\diamond}{\mathsf{M}}_{\gamma}(\mathsf{L}_{\omega}, \mathsf{g}_{\omega})_{\omega\in\Omega}} \mathsf{x} = L^{*}\big(\operatorname{prox}_{\gamma g}(L\mathsf{x})\big) = \int_{\Omega} \mathsf{L}_{\omega}^{*}\Big(\operatorname{prox}_{\gamma \mathsf{g}_{\omega}^{**}}(\mathsf{L}_{\omega}\mathsf{x})\Big)\,\mu(d\omega). \tag{2.94}$$

(vi): It follows from (2.91), Propositions 2.19(ii) and 2.52(ii), and (2.93) that

$$\begin{aligned} \operatorname{prox}_{\gamma\overset{\blacklozenge}{\mathsf{M}}_{\gamma}(\mathsf{L}_{\omega}, \mathsf{g}_{\omega})_{\omega\in\Omega}} \mathsf{x} &= \mathsf{x} - L^{*}\big(L\mathsf{x} - \operatorname{prox}_{\gamma g}(L\mathsf{x})\big) \\ &= \mathsf{x} - \int_{\Omega} \mathsf{L}_{\omega}^{*}\Big(\mathsf{L}_{\omega}\mathsf{x} - \operatorname{prox}_{\gamma \mathsf{g}_{\omega}^{**}}(\mathsf{L}_{\omega}\mathsf{x})\Big)\mu(d\omega). \end{aligned} \tag{2.95}$$

(vii): A consequence of (2.91) and Proposition 2.20.

(viii): By (2.92) and [6, Theorem 4.7(viii)],

$$^{\gamma}g = {}^{\mathfrak{G}}\!\!\int_{\Omega}^{\oplus} {}^{\gamma}\big(\mathsf{g}_{\omega}^{**}\big)\mu(d\omega). \tag{2.96}$$

However, by Lemma 2.5(ii), $g = g^{**}$. Therefore, (2.91), Proposition 2.22(ii) and (2.96) yield

$${}^{\gamma}\Big(\overset{\blacklozenge}{\mathsf{M}}_{\gamma}(\mathsf{L}_{\omega}, \mathsf{g}_{\omega})_{\omega\in\Omega}\Big)(\mathsf{x}) = {}^{\gamma}(L \overset{\gamma}{\blacklozenge} g)(\mathsf{x}) = {}^{\gamma}g(L\mathsf{x}) = \int_{\Omega} {}^{\gamma}\big(\mathsf{g}_{\omega}^{**}\big)(\mathsf{L}_{\omega}\mathsf{x})\,\mu(d\omega). \tag{2.97}$$

(ix): The assertion is obtained by using successively (2.91), Corollary 2.23, and (viii).

(x): By (2.92) and [6, Theorem 4.7(x)],

$$\operatorname{rec} g = \int_{\Omega}^{\mathfrak{G}\,\oplus} \operatorname{rec}\big(\mathsf{g}_\omega^{**}\big)\mu(d\omega) \tag{2.98}$$

However, by Lemma 2.5(ii), $g = g^{**}$. Hence, it results from (2.91), Proposition 2.25, and (2.98) that

$$\Big(\operatorname{rec} \overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega}\Big)(\mathsf{x}) = \Big(\operatorname{rec}\big(L \overset{\gamma}{\blacklozenge} g\big)\Big)(\mathsf{x}) = \big(\operatorname{rec} g\big)(L\mathsf{x}) = \int_\Omega \big(\operatorname{rec}(\mathsf{g}_\omega^{**})\big)(\mathsf{L}_\omega \mathsf{x})\,\mu(d\omega). \tag{2.99}$$

(xi): It follows from (2.83), Lemma 2.5(ii), and Jensen's inequality ([2, Proposition 9.24]) that

$$\begin{aligned}(\forall x \in \mathcal{G})(\forall y \in \mathcal{G}) \quad |g(x) - g(y)|^2 &= \left|\int_\Omega \Big(\mathsf{g}_\omega\big(x(\omega)\big) - \mathsf{g}_\omega\big(y(\omega)\big)\Big)\mu(d\omega)\right|^2 \\ &\leqslant \int_\Omega \big|\mathsf{g}_\omega\big(x(\omega)\big) - \mathsf{g}_\omega\big(y(\omega)\big)\big|^2 \mu(d\omega) \\ &\leqslant \beta^2 \int_\Omega \|x(\omega) - y(\omega)\|^2_{\mathsf{G}_\omega} \mu(d\omega) \\ &= \beta^2 \|x - y\|^2_{\mathcal{G}}.\end{aligned} \tag{2.100}$$

Therefore, g is β-Lipschitzian, and the conclusion follows from (2.91) and Corollary 2.21. ☐

Our second batch of results focuses on approximation properties.

Theorem 2.54 *Suppose that Assumptions 2.48 and 2.49 are in force. For every* $\mathsf{x} \in \mathsf{H}$*, define*

$$\Big(\overset{\triangleright}{\mathsf{M}}(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega}\Big)(\mathsf{x}) = \inf\left\{\int_\Omega \mathsf{g}_\omega^{**}(x(\omega))\mu(d\omega) \;\middle|\; x \in \mathcal{G} \text{ and } \int_\Omega \mathsf{L}_\omega^{*}(x(\omega))\mu(d\omega) = \mathsf{x}\right\} \tag{2.101}$$

and write $(\overset{\triangleright}{\mathsf{M}}(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega})(\mathsf{x}) = (\overset{\triangleright}{\mathsf{M}}(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\omega})(\mathsf{x})$ *if the infimum is attained. Then the following hold:*

(i) *Let* $\gamma \in \,]0, +\infty[$*. Then* $\overset{\diamond}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega} \geqslant \overset{\triangleright}{\mathsf{M}}(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega}$*.*

(ii) *Let* $\gamma \in \,]0, +\infty[$ *and* $\mathsf{x} \in \mathsf{H}$*. Then*

$$\int_\Omega {}^{\gamma}(\mathsf{g}_\omega^{**})(\mathsf{L}_\omega \mathsf{x})\,\mu(d\omega) \leqslant \Big(\overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega}\Big)(\mathsf{x}) \leqslant \int_\Omega \mathsf{g}_\omega^{**}(\mathsf{L}_\omega \mathsf{x})\,\mu(d\omega). \tag{2.102}$$

(iii) *Let* $\gamma \in \,]0, +\infty[$*. Then* $\overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega} \leqslant \overset{\diamond}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega}$*.*

(iv) *Let* $\gamma \in \,]0, +\infty[$ *and suppose that* μ *is a probability measure and that, for every* $\omega \in \Omega$*,* L_ω *is an isometry. Then* $\overset{\diamond}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega} = \overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega}$*.*

(v) *Let* $\gamma \in \left]0,+\infty\right[$ *and suppose that* L *in* (2.82) *is a coisometry. Then the following are satisfied:*

(a) $\overset{\diamond}{\mathsf{M}}_\gamma(\mathsf{L}_\omega,\mathsf{g}_\omega)_{\omega\in\Omega} = \overset{\triangleright}{\mathsf{M}}(\mathsf{L}_\omega,\mathsf{g}_\omega)_{\omega\in\Omega}$.

(b) *Let* $\mathsf{x} \in \mathsf{H}$. *Then* $(\overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega,\mathsf{g}_\omega)_{\omega\in\Omega})(\mathsf{x}) = \int_\Omega \mathsf{g}_\omega^{**}(\mathsf{L}_\omega\mathsf{x})\,\mu(d\omega)$.

(vi) *Let* $\mathsf{x} \in \mathsf{H}$. *Then the following are satisfied:*

(a) $\lim_{\gamma\to+\infty}(\overset{\diamond}{\mathsf{M}}_\gamma(\mathsf{L}_\omega,\mathsf{g}_\omega)_{\omega\in\Omega})(\mathsf{x}) = (\overset{\triangleright}{\mathsf{M}}(\mathsf{L}_\omega,\mathsf{g}_\omega)_{\omega\in\Omega})(\mathsf{x})$.

(b) $\lim_{0<\gamma\to 0}(\overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega,\mathsf{g}_\omega)_{\omega\in\Omega})(\mathsf{x}) = \int_\Omega \mathsf{g}_\omega^{**}(\mathsf{L}_\omega\mathsf{x})\,\mu(d\omega)$.

(vii) *Suppose that* H *and* $\mathcal{G}$ *are finite-dimensional, and let* $(\gamma_n)_{n\in\mathbb{N}}$ *be a sequence in* $\left]0,+\infty\right[$. *Then the following are satisfied:*

(a) *Suppose that* $\gamma_n \uparrow +\infty$. *Then* $\overset{\diamond}{\mathsf{M}}_{\gamma_n}(\mathsf{L}_\omega,\mathsf{g}_\omega)_{\omega\in\Omega} \xrightarrow{e} \Big(\overset{\triangleright}{\mathsf{M}}(\mathsf{L}_\omega,\mathsf{g}_\omega)_{\omega\in\Omega}\Big)^{\breve{}}$.

(b) *Suppose that* $\gamma_n \downarrow 0$. *Then* $\overset{\blacklozenge}{\mathsf{M}}_{\gamma_n}(\mathsf{L}_\omega,\mathsf{g}_\omega)_{\omega\in\Omega} \xrightarrow{e} \mathsf{f}$, *where* $(\forall \mathsf{x} \in \mathsf{H})\ \mathsf{f}(\mathsf{x}) = \int_\Omega \mathsf{g}_\omega^{**}(\mathsf{L}_\omega\mathsf{x})\,\mu(d\omega)$.

(c) *Suppose that* $\gamma_n \downarrow 0$ *and that the function* $\mathsf{x} \mapsto \int_\Omega \mathsf{g}_\omega^{**}(\mathsf{L}_\omega\mathsf{x})\mu(d\omega)$ *is proper and coercive. Then*

$$\inf_{\mathsf{x}\in\mathsf{H}}(\overset{\blacklozenge}{\mathsf{M}}_{\gamma_n}(\mathsf{L}_\omega,\mathsf{g}_\omega)_{\omega\in\Omega})(\mathsf{x}) \to \min_{\mathsf{x}\in\mathsf{H}}\int_\Omega \mathsf{g}_\omega^{**}(\mathsf{L}_\omega\mathsf{x})\,\mu(d\omega). \tag{2.103}$$

Proof. Define L as in (2.82) and g as in (2.83), and recall from items (i) and (iii) of Proposition 2.52 that $L \in \mathcal{B}(\mathsf{H},\mathcal{G})$, $0 < \|L\| \leqslant 1$, and $g \in \Gamma_0(\mathcal{G})$. Further, by Proposition 2.52(iv)–(v),

$$\overset{\diamond}{\mathsf{M}}_\gamma(\mathsf{L}_\omega,\mathsf{g}_\omega)_{\omega\in\Omega} = L \overset{\gamma}{\diamond} g, \quad \text{and} \quad \overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega,\mathsf{g}_\omega)_{\omega\in\Omega} = L \overset{\gamma}{\blacklozenge} g. \tag{2.104}$$

Additionally, Proposition 2.52(ii) yields

$$(\forall \mathsf{x} \in \mathsf{H}) \quad \big(L^* \triangleright g\big)(\mathsf{x}) = \inf_{\substack{x\in\mathcal{G}\\ L^*x=\mathsf{x}}} g(x) = \Big(\overset{\triangleright}{\mathsf{M}}(\mathsf{L}_\omega,\mathsf{g}_\omega)_{\omega\in\Omega}\Big)(\mathsf{x}). \tag{2.105}$$

On the other hand,

$$(\forall \mathsf{x} \in \mathsf{H}) \quad g(L\mathsf{x}) = \int_\Omega \mathsf{g}_\omega^{**}\big((\mathfrak{e}_{\mathsf{L}}\mathsf{x})(\omega)\big)\mu(d\omega) = \int_\Omega \mathsf{g}_\omega^{**}\big(\mathsf{L}_\omega\mathsf{x}\big)\mu(d\omega). \tag{2.106}$$

(i): The assertion follows from (2.104), (2.105), and Proposition 2.29(i).

(ii): Combine (2.104), (2.106), and Proposition 2.29(ii).

(iii): This is a consequence of (2.104) and Proposition 2.29(iii).

(iv): We have

$$(\forall \mathsf{x} \in \mathsf{H}) \quad \|L\mathsf{x}\|_{\mathcal{G}}^2 = \int_\Omega \|\mathsf{L}_\omega \mathsf{x}\|_{\mathsf{G}_\omega}^2 \mu(d\omega) = \int_\Omega \|\mathsf{x}\|_{\mathsf{H}}^2 \mu(d\omega) = \mu(\Omega)\|\mathsf{x}\|_{\mathsf{H}}^2 = \|\mathsf{x}\|_{\mathsf{H}}^2. \tag{2.107}$$

Therefore, L is an isometry and the assertion follows from (2.104) and Proposition 2.29(iv).

(v)(a): This follows from (2.104), (2.105), and Proposition 2.29(v).

(v)(b): This follows from (2.104), (2.106), and Proposition 2.29(v).

(vi)(a): This follows from (2.104), (2.105), and Theorem 2.38(iii).

(vi)(b): This follows from (2.104), (2.106), and Theorem 2.38(iv).

(vii)(a): This follows from (2.104), (2.105), and Theorem 2.43(ii)(a).

(vii)(b): This follows from (2.104), (2.106), and Theorem 2.43(ii)(b).

(vii)(c): This follows from (2.104), (2.106), and Proposition 2.46(i). ☐

Example 2.55 Let $p \in \mathbb{N} \smallsetminus \{0\}$, let $(\alpha_k)_{1\leqslant k\leqslant p}$ be a family in $]0,+\infty[$, let H and $(\mathsf{G}_k)_{1\leqslant k\leqslant p}$ be separable real Hilbert spaces, let $\mathfrak{G} = \mathsf{G}_1 \times \cdots \times \mathsf{G}_p$ be the usual Cartesian product vector space, with generic element $x = (\mathsf{x}_k)_{1\leqslant k\leqslant p}$, and, for every $k \in \{1,\ldots,p\}$, let $\mathsf{L}_k \in \mathcal{B}(\mathsf{H},\mathsf{G}_k)$ and let $\mathsf{g}_k \in \Gamma_0(\mathsf{G}_k)$. Suppose that $0 < \sum_{k=1}^p \alpha_k \|\mathsf{L}_k\|^2 \leqslant 1$ and set

$$\Omega = \{1,\ldots,p\}, \quad \mathcal{F} = 2^{\{1,\ldots,p\}}, \quad \text{and} \quad (\forall k \in \{1,\ldots,p\}) \;\; \mu(\{k\}) = \alpha_k. \tag{2.108}$$

Then $((\mathsf{G}_k)_{1\leqslant k\leqslant p}, \mathfrak{G})$ is an $\mathcal{F}$-measurable vector field of real Hilbert spaces and the space ${}^{\mathfrak{G}}\!\int_\Omega^\oplus \mathsf{G}_\omega \mu(d\omega)$ is the weighted Hilbert direct sum of $(\mathsf{G}_k)_{1\leqslant k\leqslant p}$, namely the Hilbert space obtained by equipping $\mathfrak{G}$ with the scalar product $(x,y) \mapsto \sum_{k=1}^p \alpha_k \langle \mathsf{x}_k \mid \mathsf{y}_k\rangle_{\mathsf{G}_k}$. Further, $\int_\Omega \|\mathsf{L}_\omega\|^2 \mu(d\omega) = \sum_{k=1}^p \alpha_k \|\mathsf{L}_k\|^2 \in \,]0,1]$. Therefore, Assumptions 2.48 and 2.49 are satisfied, and (2.80) becomes a parametrized version of the *proximal mixture* of [9, Example 5.9], namely,

$$\overset{\diamond}{\mathsf{M}}_\gamma(\mathsf{L}_k,\mathsf{g}_k)_{1\leqslant k\leqslant p} = \left(\sum_{k=1}^p \alpha_k \, {}^{\frac{1}{\gamma}}(\mathsf{g}_k^*) \circ \mathsf{L}_k\right)^* - \frac{1}{\gamma}\mathcal{Q}_\mathsf{H}, \tag{2.109}$$

while (2.81) becomes a parametrized version of the *proximal comixture*

$$\overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{L}_k,\mathsf{g}_k)_{1\leqslant k\leqslant p} = \left(\left(\sum_{k=1}^p \alpha_k \, {}^{\gamma}(\mathsf{g}_k^{**}) \circ \mathsf{L}_k\right)^* - \gamma\mathcal{Q}_\mathsf{H}\right)^*. \tag{2.110}$$

In particular, for every $\mathsf{x} \in \mathsf{H}$, we derive from Theorem 2.54(vi) the following new facts:

(i) $\displaystyle \lim_{\gamma\to+\infty} \Big(\overset{\diamond}{\mathsf{M}}_\gamma(\mathsf{L}_k,\mathsf{g}_k)_{1\leqslant k\leqslant p}\Big)(\mathsf{x}) = \Big(\overset{\triangleright}{\mathsf{M}}(\mathsf{L}_k,\mathsf{g}_k)_{1\leqslant k\leqslant p}\Big)(\mathsf{x}) = \inf_{\substack{\mathsf{y}_1\in\mathsf{G}_1,\ldots,\mathsf{y}_p\in\mathsf{G}_p\\ \sum_{k=1}^p \alpha_k \mathsf{L}_k^*\mathsf{y}_k = \mathsf{x}}} \left(\sum_{k=1}^p \alpha_k \mathsf{g}_k^{**}(\mathsf{y}_k)\right).$

(ii) $\displaystyle \lim_{0<\gamma\to 0} \Big(\overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{L}_k,\mathsf{g}_k)_{1\leqslant k\leqslant p}\Big)(\mathsf{x}) = \sum_{k=1}^p \alpha_k \mathsf{g}_k^{**}(\mathsf{L}_k\mathsf{x}).$

Example 2.56 In the context of Example 2.55, suppose that H is finite-dimensional and that, for every $k \in \{1, \dots, p\}$, G_k is finite-dimensional and $\mathsf{g}_k \in \Gamma_0(\mathsf{G}_k)$. Let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $]0, +\infty[$. Then we obtain the following new results on proximal mixtures and comixtures.

(i) Suppose that $\gamma_n \uparrow +\infty$. Then Theorem 2.54(vii)(a) implies that

$$\overset{\diamond}{\mathsf{M}}_{\gamma_n}(\mathsf{L}_k, \mathsf{g}_k)_{1\leqslant k\leqslant p} \overset{e}{\to} \Big(\overset{\triangleright}{\mathsf{M}}(\mathsf{L}_k, \mathsf{g}_k)_{1\leqslant k\leqslant p}\Big)^{\breve{}}. \tag{2.111}$$

(ii) Suppose that $\gamma_n \downarrow 0$. Then Theorem 2.54(vii)(b) implies that

$$\overset{\blacklozenge}{\mathsf{M}}_{\gamma_n}(\mathsf{L}_k, \mathsf{g}_k)_{1\leqslant k\leqslant p} \overset{e}{\to} \sum_{k=1}^{p} \alpha_k \mathsf{g}_k \circ \mathsf{L}_k. \tag{2.112}$$

(iii) Suppose that $\gamma_n \downarrow 0$ and that the function $\sum_{k=1}^{p} \alpha_k \mathsf{g}_k \circ \mathsf{L}_k$ is proper and coercive. Then Theorem 2.54(vii)(c) implies that

$$\inf_{\mathsf{x}\in\mathsf{H}} \Big(\overset{\blacklozenge}{\mathsf{M}}_{\gamma_n}(\mathsf{L}_k, \mathsf{g}_k)_{1\leqslant k\leqslant p}\Big)(\mathsf{x}) \to \min_{\mathsf{x}\in\mathsf{H}} \sum_{k=1}^{p} \alpha_k \mathsf{g}_k(\mathsf{L}_k \mathsf{x}). \tag{2.113}$$

Remark 2.57 In connection with Example 2.56, it was empirically argued in [11] (see also [14, 15, 18, 20] for the special cases of proximal averages) that, in variational formulations arising in image recovery and machine learning, combining linear operators $(\mathsf{L}_k)_{1\leqslant k\leqslant p}$ and convex functions $(\mathsf{g}_k)_{1\leqslant k\leqslant p}$ by means of the proximal comixture (2.110) instead of the standard averaging operation $\sum_{k=1}^{p} \alpha_k \mathsf{g}_k \circ \mathsf{L}_k$ had modeling and numerical advantages. For instance, the proximity of the former is intractable in general [12], while that of the latter is explicitly given by Theorem 2.53(vi) to be $\mathsf{Id}_{\mathsf{H}} - \sum_{k=1}^{p} \alpha_k (\mathsf{L}_k^* \circ (\mathsf{Id}_{\mathsf{G}_k} - \mathrm{prox}_{\gamma \mathsf{g}_k}) \circ \mathsf{L}_k)$, which makes the implementation of first-order optimization algorithms [10] straightforward. The results of Example 2.56 provide a theoretical context that sheds more light on such an approximation.

2.2.4.3 Proximal expectations

We specialize the results of Section 2.2.4.2 to the proximal expectation. This operation, introduced in [7, Definition 4.6] as an extension of the proximal average for finite families, performs a nonlinear averaging of an arbitrary family of functions. We study here the following extension of it which incorporates a parameter.

Definition 2.58 Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a complete probability space, let H be a separable real Hilbert space, let $(\mathsf{f}_\omega)_{\omega\in\Omega}$ be a family of functions in $\Gamma_0(\mathsf{H})$ such that the function

$$\Omega \times \mathsf{H} \to]-\infty, +\infty] \colon (\omega, \mathsf{x}) \mapsto \mathsf{f}_\omega(\mathsf{x}) \tag{2.114}$$

is $\mathcal{F}\otimes\mathcal{B}_{\mathsf{H}}$-measurable. Suppose that there exist vectors $r\in\mathscr{L}^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ and $r^*\in\mathscr{L}^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ such that the functions $\omega\mapsto\mathsf{f}_\omega(r(\omega))$ and $\omega\mapsto\mathsf{f}_\omega^*(r^*(\omega))$ belong to $\mathscr{L}^1(\Omega,\mathcal{F},\mathsf{P};\mathbb{R})$. The *proximal expectation* of $(\mathsf{f}_\omega)_{\omega\in\Omega}$ with parameter $\gamma\in\left]0,+\infty\right[$ is

$$\overset{\diamond}{\mathsf{E}}_\gamma(\mathsf{f}_\omega)_{\omega\in\Omega}=\mathsf{h}^*-\frac{1}{\gamma}\mathscr{Q}_{\mathsf{H}},\quad\text{where}\quad(\forall\mathsf{x}\in\mathsf{H})\quad\mathsf{h}(\mathsf{x})=\int_\Omega{}^{\frac{1}{\gamma}}(\mathsf{f}_\omega^*)(\mathsf{x})\mathsf{P}(d\omega).\tag{2.115}$$

An inspection of Definition 2.51 suggests that the proximal expectation can be viewed as the instance of the integral proximal mixture in which $(\forall\omega\in\Omega)$ $\mathsf{G}_\omega=\mathsf{H}$ and $\mathsf{L}_\omega=\mathrm{Id}_{\mathsf{H}}$. This fact opens the possibility of specializing the results of Section 2.2.4.2 to obtain properties of the proximal expectation. Let us formalize these ideas.

Proposition 2.59 *Consider the setting of Definition* 2.58 *and let* $\gamma\in\left]0,+\infty\right[$. *Then the following hold:*

(i) $\overset{\diamond}{\mathsf{E}}_\gamma(\mathsf{f}_\omega)_{\omega\in\Omega}=\overset{\diamond}{\mathsf{M}}_\gamma\big(\mathrm{Id}_{\mathsf{H}},\mathsf{f}_\omega\big)_{\omega\in\Omega}=\overset{\blacklozenge}{\mathsf{M}}_\gamma\big(\mathrm{Id}_{\mathsf{H}},\mathsf{f}_\omega\big)_{\omega\in\Omega}$.

(ii) $\overset{\diamond}{\mathsf{E}}_\gamma(\mathsf{f}_\omega)_{\omega\in\Omega}\in\Gamma_0(\mathsf{H})$.

(iii) $(\overset{\diamond}{\mathsf{E}}_\gamma(\mathsf{f}_\omega)_{\omega\in\Omega})^*=\overset{\diamond}{\mathsf{E}}_{1/\gamma}(\mathsf{f}_\omega^*)_{\omega\in\Omega}$.

(iv) *Let* $\mathsf{x}\in\mathsf{H}$. *Then* $\mathrm{prox}_{\gamma\overset{\diamond}{\mathsf{E}}_\gamma(\mathsf{f}_\omega)_{\omega\in\Omega}}\mathsf{x}=\displaystyle\int_\Omega\mathrm{prox}_{\gamma\mathsf{f}_\omega}\mathsf{x}\,\mathsf{P}(d\omega)$.

(v) *Let* $\mathsf{x}\in\mathsf{H}$. *Then* ${}^{\gamma}(\overset{\diamond}{\mathsf{E}}_\gamma(\mathsf{f}_\omega)_{\omega\in\Omega})(\mathsf{x})=\displaystyle\int_\Omega{}^{\gamma}\mathsf{f}_\omega(\mathsf{x})\,\mathsf{P}(d\omega)$.

(vi) $\mathrm{Argmin}_{\mathsf{x}\in\mathsf{H}}(\overset{\diamond}{\mathsf{E}}_\gamma(\mathsf{f}_\omega)_{\omega\in\Omega})(\mathsf{x})=\mathrm{Argmin}_{\mathsf{x}\in\mathsf{H}}\displaystyle\int_\Omega{}^{\gamma}\mathsf{f}_\omega(\mathsf{x})\,\mathsf{P}(d\omega)$.

(vii) *Let* $\mathsf{x}\in\mathsf{H}$. *Then* $(\mathrm{rec}\,\overset{\diamond}{\mathsf{E}}_\gamma(\mathsf{f}_\omega)_{\omega\in\Omega})(\mathsf{x})=\displaystyle\int_\Omega(\mathrm{rec}\,\mathsf{f}_\omega)(\mathsf{x})\,\mathsf{P}(d\omega)$.

(viii) *Suppose that there exists* $\beta\in\left]0,+\infty\right[$ *such that, for every* $\omega\in\Omega$, $\mathsf{f}_\omega\colon\mathsf{H}\to\mathbb{R}$ *is* β-*Lipschitzian. Then* $\overset{\diamond}{\mathsf{E}}_\gamma(\mathsf{f}_\omega)_{\omega\in\Omega}$ *is* β-*Lipschitzian.*

Proof. (i): As in the proof of [7, Proposition 4.7], the family $(\mathsf{f}_\omega)_{\omega\in\Omega}$ fulfills the properties of Assumption 2.49. Therefore, the conclusion follows from (2.115), (2.80), and Theorem 2.54(iv).

(ii)–(viii): Combine (i) and Theorem 2.53. ☐

Remark 2.60 Item (iv) in Proposition 2.59 justifies calling $\overset{\diamond}{\mathsf{E}}_\gamma(\mathsf{f}_\omega)_{\omega\in\Omega}$ the proximal expectation of $(\mathsf{f}_\omega)_{\omega\in\Omega}$: its proximity operator is the expected value of the individual ones.

Proposition 2.61 *Consider the setting of Definition* 2.58. *For every* $\mathsf{x}\in\mathsf{H}$, *define*

$$\begin{aligned}&\Big(\overset{\triangleright}{\mathsf{E}}(\mathsf{f}_\omega)_{\omega\in\Omega}\Big)(\mathsf{x})\\&=\inf\bigg\{\int_\Omega\mathsf{f}_\omega(x(\omega))\mathsf{P}(d\omega)\;\bigg|\;x\in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})\ \text{and}\ \int_\Omega x(\omega)\mathsf{P}(d\omega)=\mathsf{x}\bigg\}.\end{aligned}\tag{2.116}$$

Then the following hold:

(i) *Let* $\gamma \in \,]0,+\infty[$ *and* $\mathsf{x} \in \mathsf{H}$. *Then* $(\overset{\diamond}{\mathsf{E}}_\gamma(\mathsf{f}_\omega)_{\omega\in\Omega})(\mathsf{x}) \geqslant \int_\Omega {}^\gamma\mathsf{f}_\omega(\mathsf{x})\,\mathsf{P}(d\omega)$.

(ii) *Let* $\gamma \in \,]0,+\infty[$ *and* $\mathsf{x} \in \mathsf{H}$. *Then*

$$\Big(\overset{\triangleright}{\mathsf{E}}(\mathsf{f}_\omega)_{\omega\in\Omega}\Big)(\mathsf{x}) \leqslant \Big(\overset{\diamond}{\mathsf{E}}_\gamma(\mathsf{f}_\omega)_{\omega\in\Omega}\Big)(\mathsf{x}) \leqslant \int_\Omega \mathsf{f}_\omega(\mathsf{x})\,\mathsf{P}(d\omega). \tag{2.117}$$

(iii) *Let* $\mathsf{x} \in \mathsf{H}$. *Then the following are satisfied:*

(a) $\lim_{\gamma\to+\infty}(\overset{\diamond}{\mathsf{E}}_\gamma(\mathsf{f}_\omega)_{\omega\in\Omega})(\mathsf{x}) = (\overset{\triangleright}{\mathsf{E}}(\mathsf{f}_\omega)_{\omega\in\Omega})(\mathsf{x})$.

(b) $\lim_{0<\gamma\to 0}(\overset{\diamond}{\mathsf{E}}_\gamma(\mathsf{f}_\omega)_{\omega\in\Omega})(\mathsf{x}) = \int_\Omega \mathsf{f}_\omega(\mathsf{x})\,\mathsf{P}(d\omega)$.

(iv) *Suppose that* H *and* $\mathcal{G}$ *are finite-dimensional, and let* $(\gamma_n)_{n\in\mathbb{N}}$ *be a sequence in* $]0,+\infty[$. *Then the following are satisfied:*

(a) *Suppose that* $\gamma_n \uparrow +\infty$. *Then* $\overset{\diamond}{\mathsf{E}}_{\gamma_n}(\mathsf{f}_\omega)_{\omega\in\Omega} \xrightarrow{e} \big(\overset{\triangleright}{\mathsf{E}}(\mathsf{f}_\omega)_{\omega\in\Omega}\big)^{\breve{}}$.

(b) *Suppose that* $\gamma_n \downarrow 0$. *Then* $\overset{\diamond}{\mathsf{E}}_{\gamma_n}(\mathsf{f}_\omega)_{\omega\in\Omega} \xrightarrow{e} \mathsf{f}$, *where* $\mathsf{f}\colon \mathsf{x} \mapsto \int_\Omega \mathsf{f}_\omega(\mathsf{x})\,\mathsf{P}(d\omega)$.

(c) *Suppose that* $\gamma_n \downarrow 0$ *and that the function* $\mathsf{x} \mapsto \int_\Omega \mathsf{f}_\omega(\mathsf{x})\mathsf{P}(d\omega)$ *is proper and coercive. Then* $\inf_{\mathsf{x}\in\mathsf{H}}(\overset{\diamond}{\mathsf{E}}_{\gamma_n}(\mathsf{f}_\omega)_{\omega\in\Omega})(\mathsf{x}) \to \min_{\mathsf{x}\in\mathsf{H}} \int_\Omega \mathsf{f}_\omega(\mathsf{x})\mathsf{P}(d\omega)$.

Proof. Combine Proposition 2.59(i) and Theorem 2.54. ☐

Remark 2.62 Suppose that $(\mathsf{f}_k)_{1\leqslant k\leqslant p}$ is a finite family of functions in $\Gamma_0(\mathsf{H})$ and define P as in (2.108), with the additional assumption that $\sum_{k=1}^p \alpha_k = 1$. Then $\overset{\diamond}{\mathsf{E}}(\mathsf{f}_k)_{1\leqslant k\leqslant p}$ is the *proximal average* of $(\mathsf{f}_k)_{1\leqslant k\leqslant p}$, studied in [3] (see also [9, Example 5.9]), namely,

$$\overset{\diamond}{\mathsf{E}}_\gamma(\mathsf{f}_k)_{1\leqslant k\leqslant p} = \left(\sum_{k=1}^p \alpha_k \,{}^{\frac{1}{\gamma}}(\mathsf{f}_k^*)\right)^* - \frac{1}{\gamma}\mathcal{Q}_\mathsf{H} = \mathsf{pav}_\gamma(\mathsf{f}_k)_{1\leqslant k\leqslant p}. \tag{2.118}$$

In this context, Propositions 2.59(i)–(vi) and 2.61 recover properties presented in [3]. On the other hand, Proposition 2.59(vii)–(viii) yields the following new properties of the proximal average:

(i) $\mathrm{rec}(\mathsf{pav}_\gamma(\mathsf{f}_k)_{1\leqslant k\leqslant p}) = \sum_{k=1}^p \alpha_k \,\mathrm{rec}\,\mathsf{f}_k$.

(ii) Suppose that there exists $\beta \in \,]0,+\infty[$ such that, for every $k \in \{1,\ldots,p\}$, $\mathsf{f}_k\colon \mathsf{H} \to \mathbb{R}$ is β-Lipschitzian. Then $\mathsf{pav}_\gamma(\mathsf{f}_k)_{1\leqslant k\leqslant p}$ is β-Lipschitzian.

We conclude by making a connection between proximal expectations and integral proximal comixtures that extends Proposition 2.59(i).

Proposition 2.63 *Let* $(\Omega,\mathcal{F},\mathsf{P})$ *be a complete probability space, suppose that Assumptions* 2.48 *and* 2.49 *are in force, and let* $\gamma \in \,]0,+\infty[$. *Further, for every* $\omega \in \Omega$, *suppose that* $0 < \|\mathsf{L}_\omega\| \leqslant 1$

and set $\mathsf{f}_\omega = \mathsf{L}_\omega \overset{\gamma}{\blacklozenge} \mathsf{g}_\omega$*. Suppose that the function*

$$\Omega \times \mathsf{H} \to \left]-\infty, +\infty\right] \colon (\omega, \mathsf{x}) \mapsto \mathsf{f}_\omega(\mathsf{x}) \tag{2.119}$$

is $\mathcal{F} \otimes \mathcal{B}_{\mathsf{H}}$*-measurable and that there exist* $s \in \mathscr{L}^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ *and* $s^* \in \mathscr{L}^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ *such that the functions* $\omega \mapsto \mathsf{f}_\omega(s(\omega))$ *and* $\omega \mapsto \mathsf{f}^*_\omega(s^*(\omega))$ *lie in* $\mathscr{L}^1(\Omega, \mathcal{F}, \mathsf{P}; \mathbb{R})$*. Then*

$$\overset{\diamond}{\mathsf{E}}_\gamma\Big(\mathsf{L}_\omega \overset{\gamma}{\blacklozenge} \mathsf{g}_\omega\Big)_{\omega\in\Omega} = \overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega}. \tag{2.120}$$

Proof. As in the proof of [7, Proposition 4.7], the family $(\mathsf{f}_\omega)_{\omega\in\Omega}$ fulfills the properties of Assumption 2.49. On the other hand, Proposition 2.59(ii) and Theorem 2.53(ii) assert that $\overset{\diamond}{\mathsf{E}}_\gamma(\mathsf{f}_\omega)_{\omega\in\Omega}$ and $\overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega}$ are in $\Gamma_0(\mathsf{H})$. Further, Propositions 2.59(v) and 2.22(ii), together with Theorem 2.53(viii) yield

$$\begin{aligned}
(\forall \mathsf{x} \in \mathsf{H}) \quad {}^{\gamma}\Big(\overset{\diamond}{\mathsf{E}}_\gamma(\mathsf{f}_\omega)_{\omega\in\Omega}\Big)(\mathsf{x}) &= \int_\Omega {}^{\gamma}\mathsf{f}_\omega(\mathsf{x})\, \mathsf{P}(d\omega) \\
&= \int_\Omega {}^{\gamma}\big(\mathsf{L}_\omega \overset{\gamma}{\blacklozenge} \mathsf{g}_\omega\big)(\mathsf{x})\, \mathsf{P}(d\omega) \\
&= \int_\Omega {}^{\gamma}\big(\mathsf{g}^{**}_\omega\big)(\mathsf{L}_\omega \mathsf{x})\, \mathsf{P}(d\omega) \\
&= {}^{\gamma}\Big(\overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{L}_\omega, \mathsf{g}_\omega)_{\omega\in\Omega}\Big)(\mathsf{x}),
\end{aligned} \tag{2.121}$$

and the assertion therefore follows from Lemma 2.7. ☐

Acknowledgement

This work was supported by the National Science Foundation under grant CCF-2211123.

References

[1] H. Attouch, *Variational Convergence for Functions and Operators.* Pitman, Boston, MA, 1984.

[2] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces,* 2nd ed. Springer, New York, 2017.

[3] H. H. Bauschke, R. Goebel, Y. Lucet, and X. Wang, The proximal average: Basic theory, *SIAM J. Optim.*, vol. 19, pp. 766–785, 2008.

[4] S. R. Becker and P. L. Combettes, An algorithm for splitting parallel sums of linearly composed monotone operators, with applications to signal recovery, *J. Nonlinear Convex Anal.*, vol. 15, pp. 137–159, 2014.

[5] L. M. Briceño-Arias and N. Pustelnik, Infimal post-composition approach for composite convex optimization applied to image restoration, *Signal Process.*, vol. 223, art. 109549, 2024.

[6] M. N. Bùi and P. L. Combettes, Hilbert direct integrals of monotone operators, *Canad. J. Math.*, vol. 77, pp. 1425–1456, 2025.

[7] M. N. Bùi and P. L. Combettes, Integral resolvent and proximal mixtures, *J. Optim. Theory Appl.*, vol. 203, pp. 2328–2353, 2024.

[8] P. L. Combettes, Perspective functions: properties, constructions, and examples, *Set-Valued Var. Anal.*, vol. 26, pp. 247–264, 2018.

[9] P. L. Combettes, Resolvent and proximal compositions, *Set-Valued Var. Anal.*, vol. 31, art. 22, 29 pp., 2023.

[10] P. L. Combettes, The geometry of monotone operator splitting methods, *Acta Numer.*, vol. 33, pp. 487–632, 2024.

[11] P. L. Combettes and D. J. Cornejo, Signal recovery with proximal comixtures, *Proc. Europ. Signal Process. Conf.*, pp. 2637–2641, 2024.

[12] 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.

[13] J. Dixmier, *Les Algèbres d'Opérateurs dans l'Espace Hilbertien (Algèbres de von Neumann)*, 2e éd. Gauthier–Villars, Paris, 1969.

[14] Z. Hu, K. Shaloudegi, G. Zhang, and Y. Yu, Federated learning meets multi-objective optimization, *IEEE Trans. Network Sci. Eng.*, vol. 9, pp. 2039–2051, 2022.

[15] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, Federated optimization in heterogeneous networks, *Proc. Machine Learn. Syst.*, vol. 2, pp. 429–450, 2020.

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

[17] R. T. Rockafellar and R. J. B. Wets, *Variational Analysis*, 3rd printing. Springer, New York, 2009.

[18] L. Shen, W. Liu, J. Huang, Y.-G. Jiang, and S. Ma, Adaptive proximal average approximation for composite convex minimization, *Proc. 31st AAAI Conf. Artificial Intell.*, pp. 2513–2519, 2017.

[19] F. Xue, The operator splitting schemes revisited: Primal-dual gap and degeneracy reduction by a unified analysis, *Optim. Lett.*, vol. 18, pp. 155–194, 2024.

[20] Y.-L. Yu, Better approximation and faster algorithm using the proximal average, *Proc. Conf. Adv. Neural Inform. Process. Syst.*, pp. 458–466, 2013.

[21] M. Yukawa, H. Kaneko, K. Suzuki, and I. Yamada, Linearly-involved Moreau-enhanced-over-subspace model: Debiased sparse modeling and stable outlier-robust regression, *IEEE Trans. Signal Process.*, vol. 71, pp. 1232–1247, 2023.

Chapter 3

PARAMETRIZED FAMILIES OF RESOLVENT COMPOSITIONS

3.1 Introduction and context

We devote this chapter to question **(Q2)** of Chapter 1. We provide an in-depth analysis of the parametrized resolvent compositions and derive asymptotic results concerning operator convergence.

This chapter presents the following journal article:

D. J. Cornejo, Parametrized families of resolvent compositions, *Set-Valued and Variational Analysis*, vol. 33, art. 6, 24 pp., 2025.

3.1.1 Article: Parametrized families of resolvent compositions

Abstract. This paper presents an in-depth analysis of a parametrized version of the resolvent composition, an operation that combines a set-valued operator and a linear operator. We provide new properties and examples, and show that resolvent compositions can be interpreted as parallel compositions of perturbed operators. Additionally, we establish new monotonicity results, even in cases when the initial operator is not monotone. Finally, we derive asymptotic results regarding operator convergence, specifically focusing on graph-convergence and the ρ-Hausdorff distance.

3.1.2 Introduction

Throughout, $\mathcal{H}$ is a real Hilbert space with power set $2^{\mathcal{H}}$, identity operator $\mathrm{Id}_{\mathcal{H}}$, scalar product $\langle \cdot \mid \cdot \rangle_{\mathcal{H}}$, and associated norm $\|\cdot\|_{\mathcal{H}}$. In addition, $\mathcal{G}$ is a real Hilbert space, the space of bounded linear operators from $\mathcal{H}$ to $\mathcal{G}$ is denoted by $\mathcal{B}(\mathcal{H},\mathcal{G})$, and $\mathcal{B}(\mathcal{H}) = B(\mathcal{H},\mathcal{H})$. Let $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$. The adjoint of L is denoted by L^*, and the parallel composition of a set-valued operator $B\colon \mathcal{G} \to 2^{\mathcal{G}}$ by L^* is the operator from $\mathcal{H}$ to $2^{\mathcal{H}}$ given by

$$L^* \triangleright B = \left(L^* \circ B^{-1} \circ L\right)^{-1}. \tag{3.1}$$

We focus our attention on new methods to combine a set-valued operator with a linear operator, which have recently been introduced in [11], where they have been studied only for the case $\gamma = 1$.

Definition 3.1 Let $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$, let $B\colon \mathcal{G} \to 2^{\mathcal{G}}$, and let $\gamma \in \left]0,+\infty\right[$. The *resolvent composition* of B and L with parameter γ is the operator $L \overset{\gamma}{\diamond} B\colon \mathcal{H} \to 2^{\mathcal{H}}$ given by

$$L \overset{\gamma}{\diamond} B = L^* \triangleright (B + \gamma^{-1}\mathrm{Id}_{\mathcal{G}}) - \gamma^{-1}\mathrm{Id}_{\mathcal{H}} \tag{3.2}$$

and the *resolvent cocomposition* of B and L with parameter γ is the operator $L \overset{\gamma}{\blacklozenge} B\colon \mathcal{H} \to 2^{\mathcal{H}}$ given by $L \overset{\gamma}{\blacklozenge} B = (L \overset{1/\gamma}{\diamond} B^{-1})^{-1}$. Further, $L \diamond B = L \overset{1}{\diamond} B$ and $L \blacklozenge B = L \overset{1}{\blacklozenge} B$.

Resolvents of set-valued operators are essential in the numerical solution of monotone inclusion problems [12,16,18,19,21,22]. A motivation for studying the resolvent compositions of Definition 3.1 stems from the fact that their resolvent can be computed explicitly, unlike those of the standard composite operators $L^* \circ B \circ L$ and $L^* \triangleright B$, for which the resolvent is typically intractable and requires dedicated numerical methods [1,10,15]. Resolvent compositions also show up in relaxations of inconsistent inclusion problems [8,11]. For instance, these new composite operators can be utilized to model relaxations of convex feasibility and nonlinear reconstruction problems [17]. Furthermore, the resolvent composition of the subdifferential of a proper lower semicontinuous convex function is the subdifferential of a function called the *proximal composition* (see [8,11,14]), which has been used in image recovery and machine learning applications [13].

The goal of this paper is to present an in-depth analysis of the parametrized compositions of Definition 3.1. We provide various new properties and examples, as well as connections with connection with the parallel composition $L^* \triangleright B$ and the standard composition $L^* \circ B \circ L$. Additionally, we establish new monotonicity results, including the preservation of monotonicity, strongly monotonicity, and maximally monotonicity, and examine the Fitzpatrick function of the resolvent composition. Finally, we investigate the convergence of the operators $L \overset{\gamma}{\diamond} B$ and $L \overset{\gamma}{\blacklozenge} B$ as γ varies, by examining the graph-convergence and the ρ-Hausdorff distance convergence.

The remainder of the paper is organized as follows. In Section 3.1.3, we provide our notation and necessary mathematical background. In Section 3.1.4, we investigate new properties of the parametrized resolvent compositions and present several examples. Section 3.1.5 is devoted to study the monotonicity of resolvent compositions. Finally, Section 3.1.6 provides convergence results for parametrized resolvent compositions as the parameter varies.

3.1.3 Notation and background

We first present our notation, which follows [7].

Let $A\colon \mathcal{H}\to 2^{\mathcal{H}}$ be a set-valued operator. The domain of A is $\operatorname{dom} A=\big\{x\in\mathcal{H}\mid Ax\neq\varnothing\big\}$, the range of A is $\operatorname{ran} A=\big\{x^*\in\mathcal{H}\mid(\exists\,x\in\mathcal{H})\,x^*\in Ax\big\}$, the graph of A is $\operatorname{gra} A=\big\{(x,x^*)\in\mathcal{H}\times\mathcal{H}\mid x^*\in Ax\big\}$, the zeros of A is $\operatorname{zer} A=\big\{x\in\mathcal{H}\mid 0\in Ax\big\}$, the inverse of A is the set-valued operator A^{-1} with graph $\big\{(x^*,x)\in\mathcal{H}\times\mathcal{H}\mid x^*\in Ax\big\}$. The resolvent of A is

$$J_A=(\mathrm{Id}_{\mathcal{H}}+A)^{-1} \tag{3.3}$$

and the Yosida approximation of A of index $\gamma\in\left]0,+\infty\right[$ is

$${}^{\gamma}A=\gamma^{-1}(\mathrm{Id}_{\mathcal{H}}-J_{\gamma A})=(A^{-1}+\gamma\mathrm{Id}_{\mathcal{H}})^{-1}. \tag{3.4}$$

In particular, when $\gamma=1$,

$$\mathrm{Id}_{\mathcal{H}}-J_A=J_{A^{-1}}. \tag{3.5}$$

The operator A is monotone if

$$\big(\forall(x_1,x_1^*)\in\operatorname{gra} A\big)\big(\forall(x_2,x_2^*)\in\operatorname{gra} A\big)\;\langle x_1-x_2\mid x_1^*-x_2^*\rangle_{\mathcal{H}}\geqslant 0, \tag{3.6}$$

α-strongly monotone for some $\alpha\in\left]0,+\infty\right[$ if $A-\alpha\mathrm{Id}_{\mathcal{H}}$ is monotone, and maximally monotone if it is monotone and there exists no monotone operator $B\colon\mathcal{H}\to 2^{\mathcal{H}}$ such that $\operatorname{gra} B$ properly contains $\operatorname{gra} A$.

Let $L\in\mathcal{B}(\mathcal{H},\mathcal{G})$. If $\operatorname{ran} L$ is closed, the generalized (or Moore–Penrose) inverse of L is denoted by $L^{\dagger}$. Further, L is an isometry if $L^*\circ L=\mathrm{Id}_{\mathcal{H}}$ and a coisometry if $L\circ L^*=\mathrm{Id}_{\mathcal{G}}$.

Let D be a nonempty subset of $\mathcal{H}$ and let $T\colon D\to\mathcal{H}$. Then T is nonexpansive if

$$(\forall x\in D)(\forall y\in D)\quad\|Tx-Ty\|_{\mathcal{H}}\leqslant\|x-y\|_{\mathcal{H}}, \tag{3.7}$$

and firmly nonexpansive if $2T-\mathrm{Id}_{\mathcal{H}}$ is nonexpansive.

The Legendre conjugate of $f\colon\mathcal{H}\to[-\infty,+\infty]$ is the function

$$f^*\colon\mathcal{H}\to[-\infty,+\infty]\colon x^*\mapsto\sup_{x\in\mathcal{H}}\big(\langle x\mid x^*\rangle_{\mathcal{H}}-f(x)\big) \tag{3.8}$$

and the Moreau envelope of $f\colon\mathcal{H}\to\left]-\infty,+\infty\right]$ of parameter $\gamma\in\left]0,+\infty\right[$ is

$${}^{\gamma}f\colon\mathcal{H}\to[-\infty,+\infty]\colon x\mapsto\inf_{y\in\mathcal{H}}\Big(f(y)+\frac{1}{2\gamma}\|x-y\|_{\mathcal{H}}^2\Big). \tag{3.9}$$

A function $f\colon \mathcal{H} \to]-\infty,+\infty]$ is proper if $\operatorname{dom} f = \{x \in \mathcal{H} \mid f(x) < +\infty\} \neq \varnothing$. The set of proper lower semicontinuous convex functions from $\mathcal{H}$ to $]-\infty,+\infty]$ is denoted by $\Gamma_0(\mathcal{H})$. The subdifferential of a function $f \in \Gamma_0(\mathcal{H})$ is

$$\partial f\colon \mathcal{H} \to 2^{\mathcal{H}}\colon x \mapsto \big\{x^* \in \mathcal{H} \mid (\forall z \in \mathcal{H})\ \langle z - x \mid x^*\rangle_{\mathcal{H}} + f(x) \leqslant f(z)\big\}, \tag{3.10}$$

and its inverse is

$$(\partial f)^{-1} = \partial f^*. \tag{3.11}$$

Let C be a nonempty convex subset of $\mathcal{H}$. The normal cone of C is denoted by N_C and the strong relative interior of C is denoted by $\operatorname{sri} C$. Additionally, if C is closed, the projection operator onto C is denoted by proj_C. Finally, the closed ball with center $x \in \mathcal{H}$ and radius $\rho \in]0,+\infty[$ is denoted by $B(x;\rho)$.

The following facts will be used in the paper.

Lemma 3.2 *Let $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$, let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$, let $U\colon \mathcal{G} \to \mathcal{G}$, and let $\rho \in]0,+\infty[$. Then the following hold:*

(i) $\operatorname{dom}(L \triangleright A) \subset L(\operatorname{dom} A)$.

(ii) *Let $\mathcal{K}$ be a real Hilbert space and let $S \in \mathcal{B}(\mathcal{G},\mathcal{K})$. Then $S \triangleright (L \triangleright A) = (S \circ L) \triangleright A$.*

(iii) $\rho(L \triangleright A) = L \triangleright (\rho A)$.

(iv) $(L \triangleright A)(\rho \mathrm{Id}_{\mathcal{G}}) = L \triangleright (A(\rho \mathrm{Id}_{\mathcal{H}}))$.

(v) $L \triangleright A + U = L \triangleright (A + L^* \circ U \circ L)$.

Proof. (i): By (3.1), $\operatorname{dom}(L \triangleright A) = \operatorname{ran}(L \circ A^{-1} \circ L^*) \subset L(\operatorname{ran} A^{-1}) = L(\operatorname{dom} A)$.

(ii): [7, Proposition 25.42(ii)]

(iii): Since $A^{-1}(\rho^{-1}\mathrm{Id}_{\mathcal{H}}) = (\rho A)^{-1}$, it follows from (3.1) that $\rho(L \triangleright A) = (L \circ A^{-1} \circ L^*(\rho^{-1}\mathrm{Id}_{\mathcal{G}}))^{-1} = (L \circ (A^{-1}(\rho^{-1}\mathrm{Id}_{\mathcal{H}})) \circ L^*)^{-1} = (L \circ (\rho A)^{-1} \circ L^*)^{-1} = L \triangleright (\rho A)$.

(iv): Since $\rho^{-1}A^{-1} = (A(\rho\mathrm{Id}_{\mathcal{H}}))^{-1}$, it follows from (3.1) that $(L \triangleright A)(\rho\mathrm{Id}_{\mathcal{G}}) = (\rho^{-1} L \circ A^{-1} \circ L^*)^{-1} = (L \circ (\rho^{-1}A^{-1}) \circ L^*)^{-1} = (L \circ (A(\rho\mathrm{Id}_{\mathcal{H}}))^{-1} \circ L^*)^{-1} = L \triangleright (A(\rho\mathrm{Id}_{\mathcal{H}}))$.

(v): Let $x \in \mathcal{H}$ and set $M = L \triangleright A + U$. It follows from (3.1) that

$$\begin{aligned}
x^* \in Mx &\Leftrightarrow x^* - Ux \in \big(L \circ A^{-1} \circ L^*\big)^{-1} x \\
&\Leftrightarrow x \in L\Big(A^{-1}\big(L^*x^* - L^*(Ux)\big)\Big) \\
&\Leftrightarrow (\exists\, y \in \mathcal{G}) \quad y \in A^{-1}\big(L^*x^* - L^*(Ux)\big) \quad \text{and} \quad x = Ly \\
&\Leftrightarrow (\exists\, y \in \mathcal{G}) \quad L^*x^* \in Ay + L^*\big(U(Ly)\big) \quad \text{and} \quad x = Ly \\
&\Leftrightarrow (\exists\, y \in \mathcal{G}) \quad y \in \big(A + L^* \circ U \circ L\big)^{-1}(L^*x^*) \quad \text{and} \quad x = Ly
\end{aligned}$$

$$\Leftrightarrow x \in L\Big(\big(A + L^* \circ U \circ L\big)^{-1}(L^* x^*)\Big)$$

$$\Leftrightarrow x^* \in \big(L \triangleright (A + L^* \circ U \circ L)\big)x, \tag{3.12}$$

which completes the proof. ☐

3.1.4 Resolvent compositions

This section outlines the general properties of the parametrized compositions of Definition 3.1, which will be utilized subsequently.

Proposition 3.3 *Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$, let $B\colon \mathcal{G} \to 2^{\mathcal{G}}$, let $\gamma \in \left]0, +\infty\right[$, and let $\rho \in \left]0, +\infty\right[$. Then the following hold:*

(i) $\operatorname{dom}(L \overset{\gamma}{\diamond} B) \subset L^*(\operatorname{dom} B)$.

(ii) $\operatorname{ran}(L \overset{\gamma}{\blacklozenge} B) \subset L^*(\operatorname{ran} B)$.

(iii) $L \overset{\gamma}{\diamond} B = (L \overset{1/\gamma}{\blacklozenge} B^{-1})^{-1}$.

(iv) $\rho(L \overset{\gamma}{\diamond} B) = L \overset{\gamma/\rho}{\diamond} (\rho B)$.

(v) $(L \overset{\gamma}{\diamond} B)(\rho \mathrm{Id}_{\mathcal{H}}) = L \overset{\gamma/\rho}{\diamond} (B(\rho \mathrm{Id}_{\mathcal{G}}))$.

(vi) $\rho(L \overset{\gamma}{\blacklozenge} B) = L \overset{\gamma/\rho}{\blacklozenge} (\rho B)$.

(vii) $(L \overset{\gamma}{\blacklozenge} B)(\rho \mathrm{Id}_{\mathcal{H}}) = L \overset{\gamma/\rho}{\blacklozenge} (B(\rho \mathrm{Id}_{\mathcal{G}}))$.

(viii) *Let $z \in \mathcal{H}$, let $w \in \mathcal{G}$, and set $\tau_w B\colon x \mapsto B(x - w)$. Then the following hold:*

  (a) $(L \overset{\gamma}{\diamond} B) - z = L \overset{\gamma}{\diamond} (B - Lz)$.

  (b) $\tau_z(L \overset{\gamma}{\blacklozenge} B) = L \overset{\gamma}{\blacklozenge} (\tau_{Lz} B)$.

(ix) *Let $\mathcal{K}$ be a real Hilbert space and $S \in \mathcal{B}(\mathcal{K}, \mathcal{H})$. Then the following hold:*

  (a) $S \overset{\gamma}{\diamond} (L \overset{\gamma}{\diamond} B) = (L \circ S) \overset{\gamma}{\diamond} B$.

  (b) $S \overset{\gamma}{\blacklozenge} (L \overset{\gamma}{\blacklozenge} B) = (L \circ S) \overset{\gamma}{\blacklozenge} B$.

(x) *Set $\beta = \gamma/(1 + \rho\gamma)$. Then $L \overset{\gamma}{\diamond} (B + \rho \mathrm{Id}_{\mathcal{G}}) = (L \overset{\beta}{\diamond} B) + \rho \mathrm{Id}_{\mathcal{H}}$.*

(xi) $^{\rho}(L \overset{\gamma+\rho}{\blacklozenge} B) = L \overset{\gamma}{\blacklozenge} ({}^{\rho}B)$.

(xii) $^{\gamma}(L \overset{\gamma}{\blacklozenge} B) = L^* \circ ({}^{\gamma}B) \circ L$.

(xiii) $\operatorname{zer}(L \overset{\gamma}{\blacklozenge} B) = \operatorname{zer}(L^* \circ ({}^{\gamma}B) \circ L)$.

Proof. (i): By Definition 3.1 and Lemma 3.2(i), $\operatorname{dom}(L \overset{\gamma}{\diamond} B) \subset L^*(\operatorname{dom}(B + \gamma \mathrm{Id}_{\mathcal{G}})) = L^*(\operatorname{dom} B)$.

(ii): By Definition 3.1 and (i), $\operatorname{ran}(L \overset{\gamma}{\blacklozenge} B) = \operatorname{dom}(L \overset{1/\gamma}{\diamond} B^{-1}) \subset L^*(\operatorname{dom} B^{-1}) = L^*(\operatorname{ran} B)$.

(iii): This follows from Definition 3.1.

(iv): It follows from Definition 3.1 and Lemma 3.2(iii) that

$$\begin{aligned}\rho\big(L \overset{\gamma}{\diamond} B\big) &= \rho\big(L^* \rhd (B + \gamma^{-1}\mathrm{Id}_{\mathcal{G}})\big) - \rho\gamma^{-1}\mathrm{Id}_{\mathcal{H}} \\ &= L^* \rhd (\rho B + \rho\gamma^{-1}\mathrm{Id}_{\mathcal{G}}) - \rho\gamma^{-1}\mathrm{Id}_{\mathcal{H}} \\ &= L \overset{\gamma/\rho}{\diamond} (\rho B).\end{aligned} \tag{3.13}$$

(v): By Definition 3.1 and Lemma 3.2(iv), we obtain

$$\begin{aligned}\big(L \overset{\gamma}{\diamond} B\big)(\rho\mathrm{Id}_{\mathcal{H}}) &= \big(L^* \rhd (B + \gamma^{-1}\mathrm{Id}_{\mathcal{G}})\big)(\rho\mathrm{Id}_{\mathcal{H}}) - \rho\gamma^{-1}\mathrm{Id}_{\mathcal{H}} \\ &= L^* \rhd \big(B(\rho\mathrm{Id}_{\mathcal{G}}) + \rho\gamma^{-1}\mathrm{Id}_{\mathcal{G}}\big) - \rho\gamma^{-1}\mathrm{Id}_{\mathcal{H}} \\ &= L \overset{\gamma/\rho}{\diamond} \big(B(\rho\mathrm{Id}_{\mathcal{G}})\big).\end{aligned} \tag{3.14}$$

(vi): By Definition 3.1 and (v),

$$\begin{aligned}\rho\big(L \overset{\gamma}{\blacklozenge} B\big) = \rho\big(L \overset{1/\gamma}{\diamond} B^{-1}\big)^{-1} &= \Big(\big(L \overset{1/\gamma}{\diamond} B^{-1}\big)(\mathrm{Id}_{\mathcal{H}}/\rho)\Big)^{-1} \\ &= \big(L \overset{\rho/\gamma}{\diamond} (\rho B)^{-1}\big)^{-1} = L \overset{\gamma/\rho}{\blacklozenge} (\rho B).\end{aligned} \tag{3.15}$$

(vii): By Definition 3.1 and (iv),

$$\begin{aligned}\big(L \overset{\gamma}{\blacklozenge} B\big)(\rho\mathrm{Id}_{\mathcal{H}}) = \big(L \overset{1/\gamma}{\diamond} B^{-1}\big)^{-1}(\rho\mathrm{Id}_{\mathcal{H}}) &= \Big(\rho^{-1}\big(L \overset{1/\gamma}{\diamond} B^{-1}\big)\Big)^{-1} \\ &= \Big(L \overset{\rho/\gamma}{\diamond} \big(B(\rho\mathrm{Id}_{\mathcal{G}})\big)^{-1}\Big)^{-1} = L \overset{\gamma/\rho}{\blacklozenge} \big(B(\rho\mathrm{Id}_{\mathcal{G}})\big).\end{aligned} \tag{3.16}$$

(viii)(a): Set $U\colon \mathcal{H} \to \mathcal{H}\colon x \mapsto z$. By Definition 3.1 and Lemma 3.2(v),

$$\begin{aligned}\big(L \overset{\gamma}{\diamond} B\big) - z &= L^* \rhd (B + \gamma^{-1}\mathrm{Id}_{\mathcal{G}} - L \circ U \circ L^*) - \gamma^{-1}\mathrm{Id}_{\mathcal{H}} \\ &= L^* \rhd (B - Lz + \gamma^{-1}\mathrm{Id}_{\mathcal{G}}) - \gamma^{-1}\mathrm{Id}_{\mathcal{H}} \\ &= L \overset{\gamma}{\diamond} (B - Lz).\end{aligned} \tag{3.17}$$

(viii)(b): Since $\tau_w B = (B^{-1} + w)^{-1}$, we combine Definition 3.1 and (viii)(a) to derive

$$\tau_z\big(L \overset{\gamma}{\blacklozenge} B\big) = \Big(\big(L \overset{1/\gamma}{\diamond} B^{-1}\big) + z\Big)^{-1} = \big(L \overset{1/\gamma}{\diamond} (B^{-1} + Lz)\big)^{-1} = L \overset{\gamma}{\blacklozenge} (\tau_{Lz} B). \tag{3.18}$$

(ix)(a): By Definition 3.1 and Lemma 3.2(ii),

$$
\begin{aligned}
(L \circ S) \overset{\gamma}{\diamond} B &= (L \circ S)^* \triangleright (B + \gamma^{-1}\mathrm{Id}_{\mathcal{G}}) - \gamma^{-1}\mathrm{Id}_{\mathcal{K}} \\
&= S^* \triangleright \big(L^* \triangleright (B + \gamma^{-1}\mathrm{Id}_{\mathcal{G}})\big) - \gamma^{-1}\mathrm{Id}_{\mathcal{K}} \\
&= S^* \triangleright \big((L \overset{\gamma}{\diamond} B) + \gamma^{-1}\mathrm{Id}_{\mathcal{H}}\big) - \gamma^{-1}\mathrm{Id}_{\mathcal{K}} \\
&= S \overset{\gamma}{\diamond} \big(L \overset{\gamma}{\diamond} B\big).
\end{aligned} \tag{3.19}
$$

(ix)(b): We combine Definition 3.1 and (ix)(a) to obtain

$$
S \overset{\gamma}{\blacklozenge} \big(L \overset{\gamma}{\blacklozenge} B\big) = \Big(S \overset{1/\gamma}{\diamond} \big(L \overset{1/\gamma}{\diamond} B^{-1}\big)\Big)^{-1} = \Big((L \circ S) \overset{1/\gamma}{\diamond} B^{-1}\Big)^{-1} = (L \circ S) \overset{\gamma}{\blacklozenge} B. \tag{3.20}
$$

(x): Since $\beta^{-1} = \gamma^{-1} + \rho$, we deduce from Definition 3.1 that

$$
\begin{aligned}
L \overset{\gamma}{\diamond} (B + \rho\mathrm{Id}_{\mathcal{G}}) &= L^* \triangleright (B + \rho\mathrm{Id}_{\mathcal{G}} + \gamma^{-1}\mathrm{Id}_{\mathcal{G}}) - \gamma^{-1}\mathrm{Id}_{\mathcal{H}} \\
&= L^* \triangleright (B + \beta^{-1}\mathrm{Id}_{\mathcal{G}}) - \beta^{-1}\mathrm{Id}_{\mathcal{H}} + \rho\mathrm{Id}_{\mathcal{H}} \\
&= \big(L \overset{\beta}{\diamond} B\big) + \rho\mathrm{Id}_{\mathcal{H}}.
\end{aligned} \tag{3.21}
$$

(xi): By (3.4), Definition 3.1, and (x),

$$
\begin{aligned}
{}^{\rho}\big(L \overset{\gamma+\rho}{\blacklozenge} B\big) &= \Big(\big(L \overset{1/(\gamma+\rho)}{\diamond} B^{-1}\big) + \rho\mathrm{Id}_{\mathcal{H}}\Big)^{-1} \\
&= \big(L \overset{1/\gamma}{\diamond} (B^{-1} + \rho\mathrm{Id}_{\mathcal{G}})\big)^{-1} \\
&= L \overset{\gamma}{\blacklozenge} (B^{-1} + \rho\mathrm{Id}_{\mathcal{G}})^{-1} \\
&= L \overset{\gamma}{\blacklozenge} ({}^{\rho}B).
\end{aligned} \tag{3.22}
$$

(xii): It follows from (3.4) and Definition 3.1 that

$$
{}^{\gamma}\big(L \overset{\gamma}{\blacklozenge} B\big) = \Big(\big(L \overset{1/\gamma}{\diamond} B^{-1}\big) + \gamma\mathrm{Id}_{\mathcal{H}}\Big)^{-1} = \big(L^* \triangleright (B^{-1} + \gamma\mathrm{Id}_{\mathcal{G}})\big)^{-1} = L^* \circ ({}^{\gamma}B) \circ L. \tag{3.23}
$$

(xiii): Set $A = L \overset{\gamma}{\blacklozenge} B$. It follows from (xii), (3.4), and (3.3) that $0 \in \mathrm{zer}(L^* \circ ({}^{\gamma}B) \circ L) \Leftrightarrow 0 \in \mathrm{zer}({}^{\gamma}A) \Leftrightarrow x \in J_{\gamma A}x \Leftrightarrow 0 \in Ax \Leftrightarrow x \in \mathrm{zer}\, A$. ☐

The following proposition shows that the resolvent of the operators of Definition 3.1 can be computed explicitly.

Proposition 3.4 *Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$, let $B\colon \mathcal{G} \to 2^{\mathcal{G}}$, and let $\gamma \in \left]0, +\infty\right[$. Then the following hold:*

(i) $J_{\gamma(L\overset{\gamma}{\diamond}B)} = L^* \circ J_{\gamma B} \circ L$.

(ii) $J_{\gamma(L\overset{\gamma}{\blacklozenge}B)} = \mathrm{Id}_{\mathcal{H}} - L^* \circ (\mathrm{Id}_{\mathcal{G}} - J_{\gamma B}) \circ L$.

Proof. (i): By (3.3), Proposition 3.3(iv), and Definition 3.1,

$$J_{\gamma(L\overset{\gamma}{\diamond}B)} = \Big(\mathrm{Id}_{\mathcal{H}} + \gamma\big(L \overset{\gamma}{\diamond} B\big)\Big)^{-1} = \Big(\mathrm{Id}_{\mathcal{H}} + \big(L \diamond (\gamma B)\big)\Big)^{-1}$$
$$= \big(L^* \rhd (\gamma B + \mathrm{Id}_{\mathcal{G}})\big)^{-1} = L^* \circ J_{\gamma B} \circ L. \tag{3.24}$$

(ii): By Proposition 3.3(vi), Definition 3.1, (3.5), and (i), we obtain

$$J_{\gamma(L\overset{\gamma}{\blacklozenge}B)} = J_{L\blacklozenge(\gamma B)} = J_{\left(L\diamond(\gamma B)^{-1}\right)^{-1}} = \mathrm{Id}_{\mathcal{H}} - J_{L\diamond(\gamma B)^{-1}} = \mathrm{Id}_{\mathcal{H}} - L^* \circ (\mathrm{Id}_{\mathcal{G}} - J_{\gamma B}) \circ L, \tag{3.25}$$

as claimed. ☐

The next result interprets resolvent compositions as parallel compositions of perturbed operators.

Proposition 3.5 *Let $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$, let $B\colon \mathcal{G} \to 2^{\mathcal{G}}$, let $\gamma \in \left]0,+\infty\right[$, and set $\Psi = \mathrm{Id}_{\mathcal{G}} - L \circ L^*$. Then the following hold:*

(i) $L \overset{\gamma}{\diamond} B = L^* \rhd (B + \gamma^{-1}\Psi)$.

(ii) $L \overset{\gamma}{\blacklozenge} B = L^* \circ (B^{-1} + \gamma\Psi)^{-1} \circ L$.

Proof. (i): Combining Definition 3.1 and Lemma 3.2(v), we deduce that

$$L \overset{\gamma}{\diamond} B = L^* \rhd \big(B + \gamma^{-1}\mathrm{Id}_{\mathcal{G}}\big) - \gamma^{-1}\mathrm{Id}_{\mathcal{H}}$$
$$= L^* \rhd \big(B + \gamma^{-1}\mathrm{Id}_{\mathcal{G}} + L \circ (-\gamma^{-1}\mathrm{Id}_{\mathcal{H}}) \circ L^*\big)$$
$$= L^* \rhd \big(B + \gamma^{-1}\Psi\big). \tag{3.26}$$

(ii): It follows from Definition 3.1 and (i) that

$$L \overset{\gamma}{\blacklozenge} B = \big(L \overset{1/\gamma}{\diamond} B^{-1}\big)^{-1} = \big(L^* \rhd (B^{-1} + \gamma\Psi)\big)^{-1} = L^* \circ (B^{-1} + \gamma\Psi)^{-1} \circ L, \tag{3.27}$$

as announced. ☐

We proceed to provide particular instances in which the standard, parallel, and resolvent compositions coincide.

Proposition 3.6 *Let $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$, let $B\colon \mathcal{G} \to 2^{\mathcal{G}}$, and let $\gamma \in \left]0,+\infty\right[$. Then the following hold:*

(i) *Suppose that L is an isometry. Then $L \overset{\gamma}{\diamond} B = L \overset{\gamma}{\blacklozenge} B$.*

(ii) *Suppose that L is a coisometry. Then $L \overset{\gamma}{\diamond} B = L^* \triangleright B$ and $L \overset{\gamma}{\blacklozenge} B = L^* \circ B \circ L$.*

(iii) *Suppose that L is invertible with $L^{-1} = L^*$. Then $L \overset{\gamma}{\diamond} B = L^* \triangleright B = L^* \circ B \circ L = L \overset{\gamma}{\blacklozenge} B$.*

Proof. (i): Since $L^* \circ L = \mathrm{Id}_{\mathcal{H}}$, we deduce from Proposition 3.4 and (3.3) that

$$\gamma(L \overset{\gamma}{\blacklozenge} B) = \Big(J_{\gamma(L \overset{\gamma}{\blacklozenge} B)}\Big)^{-1} - \mathrm{Id}_{\mathcal{H}} = \big(L^* \circ J_{\gamma B} \circ L\big)^{-1} - \mathrm{Id}_{\mathcal{H}} = \gamma(L \overset{\gamma}{\diamond} B). \tag{3.28}$$

(ii): Since L is a coisometry, $\Psi = \mathrm{Id}_{\mathcal{G}} - L \circ L^* = 0$. Therefore, we derive from Proposition 3.5 that $L \overset{\gamma}{\diamond} B = L^* \triangleright B$ and $L \overset{\gamma}{\blacklozenge} B = L^* \circ B \circ L$.

(iii): This follows from (i) and (ii). ☐

Corollary 3.7 *Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$, let $B \colon \mathcal{G} \to 2^{\mathcal{G}}$, and let $\gamma \in \,]0, +\infty[$. Let $\mathcal{K}$ be a real Hilbert space and suppose that $M \in \mathcal{B}(\mathcal{K}, \mathcal{H})$ and $S \in \mathcal{B}(\mathcal{G})$ are coisometries. Then the following hold:*

(i) *$M^* \triangleright (L \overset{\gamma}{\diamond} B) = (L \circ M) \overset{\gamma}{\diamond} B$.*

(ii) *$M^* \circ (L \overset{\gamma}{\blacklozenge} B) \circ M = (L \circ M) \overset{\gamma}{\blacklozenge} B$.*

(iii) *$L \overset{\gamma}{\diamond} (S^* \triangleright B) = (S \circ L) \overset{\gamma}{\diamond} B$.*

(iv) *$L \overset{\gamma}{\blacklozenge} (S^* \circ B \circ S) = (S \circ L) \overset{\gamma}{\blacklozenge} B$*

Proof. Combine Proposition 3.6(ii) and Proposition 3.3(ix). ☐

Now, we present several examples of resolvent compositions and cocompositions, starting with the representation of the Yosida approximation as one such composition.

Example 3.8 Let $A \colon \mathcal{H} \to 2^{\mathcal{H}}$ and $\gamma \in \,]0, +\infty[$. Set $L = \mathrm{Id}_{\mathcal{H}}/2$ and $B = 2A(2\mathrm{Id}_{\mathcal{H}})$. Then $L \overset{\gamma/3}{\blacklozenge} B = {}^{\gamma}\!A$.

Proof. It follows from Proposition 3.5(ii) that $L \overset{4\gamma/3}{\blacklozenge} A = (1/2)(A^{-1} + \gamma\mathrm{Id}_{\mathcal{H}})^{-1}(\mathrm{Id}_{\mathcal{H}}/2) = (1/2)\,{}^{\gamma}\!A(\mathrm{Id}_{\mathcal{H}}/2)$. Therefore, Proposition 3.3(vi)–(vii) implies that ${}^{\gamma}\!A = 2(L \overset{4\gamma/3}{\blacklozenge} A)(2\mathrm{Id}_{\mathcal{H}}) = L \overset{\gamma/3}{\blacklozenge} B$, as claimed. ☐

Example 3.9 Let V be a closed vector subspace of $\mathcal{H}$, $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$, $B \colon \mathcal{G} \to 2^{\mathcal{G}}$, and $\gamma \in \,]0, +\infty[$. Suppose that L is surjective and that $L^* \circ L = \mathrm{proj}_V$. Then $L \overset{\gamma}{\diamond} B = L^* \triangleright B$, $L \overset{\gamma}{\blacklozenge} B = L^* \circ B \circ L$, and ${}^{\gamma}(L^* \circ B \circ L) = L^* \circ ({}^{\gamma}B) \circ L$.

Proof. In this case, L is a coisometry. Therefore, Proposition 3.6(ii) implies that $L \overset{\gamma}{\diamond} B = L^* \triangleright B$ and $L \overset{\gamma}{\blacklozenge} B = L^* \circ B \circ L$. Further, we use Proposition 3.3(xii) to deduce that ${}^{\gamma}(L^* \circ B \circ L) = {}^{\gamma}(L \overset{\gamma}{\blacklozenge} B) = L^* \circ ({}^{\gamma}B) \circ L$, which completes the proof. ☐

Example 3.10 Let $S \in \mathcal{B}(\mathcal{H}, \mathcal{G})$, let $A \colon \mathcal{G} \to 2^{\mathcal{G}}$, let $\gamma \in \,]0, +\infty[$, and let $\mu \in \,]0, +\infty[$. Suppose that $S \circ S^* = \mu\mathrm{Id}_{\mathcal{G}}$. Then the following hold:

(i) Set $L = S/\sqrt{\mu}$ and $B = \sqrt{\mu}A(\sqrt{\mu}\mathrm{Id}_{\mathcal{G}})$. Then $L \overset{\gamma}{\blacklozenge} B = S^* \circ A \circ S$.

(ii) $J_{\gamma S^*\circ A\circ S} = \mathrm{Id}_{\mathcal{H}} - \mu^{-1}S^*\circ(\mathrm{Id}_{\mathcal{G}} - J_{\mu\gamma A})\circ S$.

Proof. (i): In this case, L is a coisometry and Proposition 3.6(ii) yields $L \overset{\gamma}{\diamond} B = L^*\circ B\circ L = S^*\circ A\circ S$.

(ii): By (i), $S^*\circ A\circ S = L \overset{\gamma}{\diamond} B$. Therefore, the result follows from Proposition 3.4(ii) and basic resolvent calculus. ☐

Example 3.11 Suppose that $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ satisfies $\|L\| < 1$, let $B\colon \mathcal{G} \to 2^{\mathcal{G}}$, and let $\gamma \in \left]0,+\infty\right[$. Let $\mathcal{X} = \mathcal{H}\oplus\mathcal{G}$, set $\Psi = \mathrm{Id}_{\mathcal{G}} - L\circ L^*$, and set $L_\Psi\colon \mathcal{X}\to\mathcal{G}\colon (x,y)\mapsto Lx + \Psi^{1/2}y$. Then

$$L_\Psi \overset{\gamma}{\diamond} B\colon \mathcal{X}\to 2^{\mathcal{X}}\colon (x,y)\mapsto \Big(L^*\big(B(Lx+\Psi^{1/2}y)\big)\Big)\times\Big(\Psi^{1/2}\big(B(Lx+\Psi^{1/2}y)\big)\Big). \tag{3.29}$$

Proof. Note that Ψ is self-adjoint and that

$$(\forall y\in\mathcal{G})\quad \langle \Psi y\mid y\rangle_{\mathcal{G}} = \|y\|_{\mathcal{G}}^2 - \|L^*y\|_{\mathcal{H}}^2 \geqslant (1-\|L\|^2)\|y\|_{\mathcal{G}}^2. \tag{3.30}$$

Thus, Ψ is strongly monotone and $\Psi^{1/2}$ is well defined. Further, since $L_\Psi^*\colon \mathcal{G}\to\mathcal{X}\colon y\mapsto (L^*y,\Psi^{1/2}y)$, we deduce that

$$(\forall y\in\mathcal{G})\quad L_\Psi(L_\Psi^*y) = L_\Psi(L^*y,\Psi^{1/2}y) = L(L^*y)+\Psi y = y. \tag{3.31}$$

Therefore, L_Ψ is a coisometry, and it follows from Proposition 3.6(ii) that $L_\Psi \overset{\gamma}{\diamond} B = L_\Psi^*\circ B\circ L_\Psi$, which establishes (3.29). ☐

Example 3.12 (resolvent mixture) Let $0\neq p\in\mathbb{N}$ and let $\gamma\in\left]0,+\infty\right[$. For every $k\in\{1,\ldots,p\}$, let $\mathcal{G}_k$ be a real Hilbert space, let $L_k\in\mathcal{B}(\mathcal{H},\mathcal{G}_k)$, let $B_k\colon\mathcal{G}_k\to 2^{\mathcal{G}_k}$, and let $\alpha_k\in\left]0,+\infty\right[$. Let $\mathcal{G}=\bigoplus_{k=1}^p\mathcal{G}_k$, and set

$$L\colon\mathcal{H}\to\mathcal{G}\colon x\mapsto\big(\sqrt{\alpha_1}L_1x,\ldots,\sqrt{\alpha_p}L_px\big) \tag{3.32}$$

and

$$B\colon\mathcal{G}\to 2^{\mathcal{G}}\colon (y_k)_{1\leqslant k\leqslant p}\mapsto\Big(\sqrt{\alpha_1}B_1\big(y_1/\sqrt{\alpha_1}\big)\Big)\times\cdots\times\Big(\sqrt{\alpha_p}B_p\big(y_p/\sqrt{\alpha_p}\big)\Big). \tag{3.33}$$

Then Proposition 3.4 yields

$$J_{\gamma(L\overset{\gamma}{\diamond}B)} = \sum_{k=1}^p \alpha_k L_k^*\circ J_{\gamma B_k}\circ L_k \tag{3.34}$$

and

$$J_{\gamma(L\overset{\gamma}{\diamond}B)} = \mathrm{Id}_{\mathcal{H}} - \sum_{k=1}^p \alpha_k L_k^*\circ(\mathrm{Id}_{\mathcal{G}_k} - J_{\gamma B_k})\circ L_k. \tag{3.35}$$

The operators

$$\overset{\diamond}{\mathsf{M}}_\gamma(L_k,B_k)_{1\leqslant k\leqslant p} = L \overset{\gamma}{\diamond} B = \Bigg(\sum_{k=1}^{p}\alpha_k L_k^* \circ \big(B_k+\gamma^{-1}\mathrm{Id}_{\mathcal{G}_k}\big)^{-1}\circ L_k\Bigg)^{-1} - \gamma^{-1}\mathrm{Id}_{\mathcal{H}} \tag{3.36}$$

and

$$\overset{\blacklozenge}{\mathsf{M}}_\gamma(L_k,B_k)_{1\leqslant k\leqslant p} = L \overset{\gamma}{\blacklozenge} B = \Bigg(\Bigg(\sum_{k=1}^{p}\alpha_k L_k^* \circ \big(B_k^{-1}+\gamma\mathrm{Id}_{\mathcal{G}_k}\big)^{-1}\circ L_k\Bigg)^{-1} - \gamma\mathrm{Id}_{\mathcal{H}}\Bigg)^{-1} \tag{3.37}$$

are called the *resolvent mixture* and *comixture*, respectively, and were initially introduced in [11, Example 3.4] (see also [8] for further developments).

Example 3.13 (resolvent average) In the context of Example 3.12, suppose that $\sum_{k=1}^{p}\alpha_k = 1$ and that, for every $k\in\{1,\ldots,p\}$, $\mathcal{G}_k=\mathcal{H}$ and $L_k=\mathrm{Id}_{\mathcal{H}}$. Since $L^*\colon\mathcal{G}\to\mathcal{H}\colon(y_k)_{1\leqslant k\leqslant p}\mapsto \sum_{k=1}^{p}\sqrt{\alpha_k}y_k$, the linear operator L is an isometry. Thus, Proposition 3.6(i) yields $L\overset{\gamma}{\diamond}B = L\overset{\gamma}{\blacklozenge}B$. This operator is called the *resolvent average* of $(B_k)_{1\leqslant k\leqslant p}$ and $(\alpha_k)_{1\leqslant k\leqslant p}$, denoted by $\mathsf{rav}_\gamma(B_k,\alpha_k)_{1\leqslant k\leqslant p}$, which has been studied in [6–8, 11, 23], namely,

$$L\overset{\gamma}{\diamond}B = \Bigg(\sum_{k=1}^{p}\alpha_k\big(B_k+\gamma^{-1}\mathrm{Id}_{\mathcal{H}}\big)^{-1}\Bigg)^{-1} - \gamma^{-1}\mathrm{Id}_{\mathcal{H}} = \mathsf{rav}_\gamma(B_k,\alpha_k)_{1\leqslant k\leqslant p}. \tag{3.38}$$

In addition, (3.34) yields $J_{\gamma\,\mathsf{rav}_\gamma(B_k,\alpha_k)_{1\leqslant k\leqslant p}} = \sum_{k=1}^{p}\alpha_k J_{\gamma B_k}$.

3.1.5 Monotonicity of resolvent compositions

We leverage the results of Section 3.1.4 to establish conditions that ensure the monotonicity of resolvent compositions.

Proposition 3.14 *Suppose that $0\neq L\in\mathcal{B}(\mathcal{H},\mathcal{G})$, let $B\colon\mathcal{G}\to 2^{\mathcal{G}}$, let $\gamma\in\left]0,+\infty\right[$, let $\alpha\in\left[-1/\gamma,+\infty\right[$, and set $\beta=(\alpha+\gamma^{-1})\|L\|^{-2}-\gamma^{-1}$. Suppose that $B-\alpha\mathrm{Id}_{\mathcal{G}}$ is monotone. Then $L\overset{\gamma}{\diamond}B-\beta\mathrm{Id}_{\mathcal{H}}$ is monotone.*

Proof. Set $\Psi=\mathrm{Id}_{\mathcal{G}}-L\circ L^*$ and $\mathcal{M}=(B-\alpha\mathrm{Id}_{\mathcal{G}})+(\alpha+\gamma^{-1})\|L\|^{-2}(\|L\|^2\mathrm{Id}_{\mathcal{G}}-L\circ L^*)$. It follows from Proposition 3.5(i) and Lemma 3.2(v) that

$$\begin{aligned}\big(L\overset{\gamma}{\diamond}B\big)-\beta\mathrm{Id}_{\mathcal{H}} &= L^*\rhd\big(B+\gamma^{-1}\Psi\big)-\beta\mathrm{Id}_{\mathcal{H}}\\ &= L^*\rhd\big(B+\gamma^{-1}\Psi+L\circ(-\beta\mathrm{Id}_{\mathcal{H}})\circ L^*\big)\\ &= L^*\rhd\big((B-\alpha\mathrm{Id}_{\mathcal{G}})+(\alpha+\gamma^{-1})\|L\|^{-2}(\|L\|^2\mathrm{Id}_{\mathcal{G}}-L\circ L^*)\big)\end{aligned}$$

$$= L^* \triangleright \mathcal{M}. \tag{3.39}$$

Since $\alpha + \gamma^{-1} \geqslant 0$ and the operators $B - \alpha \mathrm{Id}_{\mathcal{G}}$ and $\|L\|^2 \mathrm{Id}_{\mathcal{G}} - L \circ L^*$ are monotone, [7, Propositions 20.10] implies that $\mathcal{M}$ is monotone. Therefore, the assertion follows from (3.39) and [7, Proposition 25.41(ii)]. ☐

Corollary 3.15 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ satisfies $0 < \|L\| \leqslant 1$, let $B \colon \mathcal{G} \to 2^{\mathcal{G}}$, let $\gamma \in \left]0, +\infty\right[$, let $\alpha \in \left]0, +\infty\right[$, and set $\beta = (\alpha + \gamma^{-1})\|L\|^{-2} - \gamma^{-1}$. Suppose that B is α-strongly monotone. Then $L \overset{\gamma}{\diamond} B$ is β-strongly monotone.*

Proof. Since $\beta = \alpha\|L\|^{-2} + \gamma^{-1}(\|L\|^{-2} - 1) > 0$, the conclusion follows from Proposition 3.14. ☐

Corollary 3.16 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ satisfies $0 < \|L\| < 1$, let $B \colon \mathcal{G} \to 2^{\mathcal{G}}$ be monotone, let $\gamma \in \left]0, +\infty\right[$, and set $\beta = \gamma^{-1}(\|L\|^{-2} - 1)$. Then $L \overset{\gamma}{\diamond} B$ is β-strongly monotone.*

Proof. This follows from Proposition 3.14 when $\alpha = 0$. ☐

Proposition 3.17 *Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$, $B \colon \mathcal{G} \to 2^{\mathcal{G}}$, and $\gamma \in \left]0, +\infty\right[$. Suppose that $B + \gamma^{-1}(\mathrm{Id}_{\mathcal{G}} - L \circ L^*)$ is monotone. Then the following hold:*

(i) *$L \overset{\gamma}{\diamond} B$ is monotone.*

(ii) *Suppose that $\operatorname{ran} L \subset \operatorname{ran}(\mathrm{Id}_{\mathcal{G}} + \gamma B)$. Then $L \overset{\gamma}{\diamond} B$ is maximally monotone.*

Proof. Set $\Psi = \mathrm{Id}_{\mathcal{G}} - L \circ L^*$ and recall from Proposition 3.5(i) that $L \overset{\gamma}{\diamond} B = L^* \triangleright (B + \gamma^{-1}\Psi)$.

(i): Since $B + \gamma^{-1}\Psi$ is monotone, we deduce from [7, Proposition 25.41(ii)] that $L \overset{\gamma}{\diamond} B$ is monotone.

(ii): Since monotonicity is preserved under multiplication by positive scalars, $\gamma(L \overset{\gamma}{\diamond} B)$ is monotone by (i). Further, by assumption, $\operatorname{ran} L \subset \operatorname{ran}(\mathrm{Id}_{\mathcal{G}} + \gamma B) = \operatorname{dom}(\mathrm{Id}_{\mathcal{G}} + \gamma B)^{-1} = \operatorname{dom} J_{\gamma B}$. Therefore, $\operatorname{dom}(L^* \circ J_{\gamma B} \circ L) = \operatorname{dom}(J_{\gamma B} \circ L) = \mathcal{H}$, and it follows from Proposition 3.4(i) that $\operatorname{ran}(\mathrm{Id}_{\mathcal{H}} + \gamma(L \overset{\gamma}{\diamond} B)) = \operatorname{dom} J_{\gamma(L \overset{\gamma}{\diamond} B)} = \mathcal{H}$. Altogether, we deduce from [7, Theorem 21.1 (Minty)] that $\gamma(L \overset{\gamma}{\diamond} B)$ is maximally monotone. Hence, $L \overset{\gamma}{\diamond} B$ is maximally monotone. ☐

The following result recovers [11, Proposition 4.4(i)–(ii) and Theorem 4.5(i)–(ii)], which were proven when $\gamma = 1$ using distinct approaches.

Corollary 3.18 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ satisfies $\|L\| \leqslant 1$, let $B \colon \mathcal{G} \to 2^{\mathcal{G}}$ be monotone, and let $\gamma \in \left]0, +\infty\right[$. Then the following hold:*

(i) *$L \overset{\gamma}{\diamond} B$ and $L \overset{\gamma}{\blacklozenge} B$ are monotone.*

(ii) *Suppose that B is maximally monotone. Then $L \overset{\gamma}{\diamond} B$ and $L \overset{\gamma}{\blacklozenge} B$ are maximally monotone.*

Proof. Set $\Psi = \mathrm{Id}_{\mathcal{G}} - L \circ L^*$. Since $\|L\| \leqslant 1$, Ψ is monotone. Thus, [7, Proposition 20.10] implies that $B + \gamma^{-1}\Psi$ and B^{-1} are monotone.

(i): By Proposition 3.17(i), $L \overset{\gamma}{\diamond} B$ and $L \overset{1/\gamma}{\diamond} B^{-1}$ are monotone. Therefore, we combine Definition 3.1 and [7, Proposition 20.10] to deduce that $L \overset{\gamma}{\blacklozenge} B = (L \overset{1/\gamma}{\diamond} B^{-1})^{-1}$ is monotone.

(ii): By [7, Proposition 20.22], γB and B^{-1} are maximally monotone. Therefore, [7, Theorem 21.1] yields $\operatorname{ran}(\mathrm{Id}_{\mathcal{G}} + \gamma B) = \mathcal{G}$, and, by Proposition 3.17(ii), $L \overset{\gamma}{\diamond} B$ is maximally monotone. Similarly, $L \overset{1/\gamma}{\diamond} B^{-1}$ is maximally monotone, and it follows from [7, Proposition 20.22] that $L \overset{\gamma}{\blacklozenge} B = (L \overset{1/\gamma}{\diamond} B^{-1})^{-1}$ is maximally monotone as well. ☐

Next, we state several instantiations of resolvent compositions which are monotone.

Example 3.19 Let $0 \neq p \in \mathbb{N}$ and let $\gamma \in \left]0, +\infty\right[$. For every $k \in \{1, \ldots, p\}$, let $\mathcal{G}_k$ be a real Hilbert space, let $L_k \in \mathcal{B}(\mathcal{H}, \mathcal{G}_k)$, let $B_k \colon \mathcal{G}_k \to 2^{\mathcal{G}_k}$ be maximally monotone, and let $\alpha_k \in \left]0, +\infty\right[$. Suppose that $\sum_{k=1}^{p} \alpha_k \|L_k\|^2 \leqslant 1$. Then $\overset{\diamond}{\mathsf{M}}_\gamma(B_k, L_k)_{1 \leqslant k \leqslant p}$ and $\overset{\blacklozenge}{\mathsf{M}}_\gamma(B_k, L_k)_{1 \leqslant k \leqslant p}$ are maximally monotone.

Proof. Define L as in (3.32) and B as in (3.33). Thus, the assumption $\sum_{k=1}^{p} \alpha_k \|L_k\|^2 \leqslant 1$ implies that $\|L\| \leqslant 1$, and, by [7, Proposition 23.18], B is maximally monotone. Since $\overset{\diamond}{\mathsf{M}}_\gamma(B_k, L_k)_{1 \leqslant k \leqslant p} = L \overset{\gamma}{\diamond} B$ and $\overset{\blacklozenge}{\mathsf{M}}_\gamma(B_k, L_k)_{1 \leqslant k \leqslant p} = L \overset{\gamma}{\blacklozenge} B$, the conclusion follows from Corollary 3.18(ii). ☐

Example 3.20 Let $A_1 \colon \mathcal{H} \to 2^{\mathcal{H}}$ and $A_2 \colon \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone, let $\mathcal{G} = \mathcal{H} \oplus \mathcal{H}$, set $L \colon \mathcal{H} \to \mathcal{G} \colon x \mapsto (x, -x)$, and set

$$B \colon \mathcal{G} \to 2^{\mathcal{G}} \colon (x, y) \mapsto \big(A_1 x\big) \times \big(A_2^{-1} y - 2x\big). \tag{3.40}$$

Then $L \diamond B$ is maximally monotone and

$$J_{L \diamond B} = \frac{1}{2}\mathrm{Id}_{\mathcal{H}} + \frac{1}{2}\big(2J_{A_2} - \mathrm{Id}_{\mathcal{H}}\big) \circ \big(2J_{A_1} - \mathrm{Id}_{\mathcal{H}}\big). \tag{3.41}$$

In other words, the resolvent of $L \diamond B$ is the *Douglas–Rachford splitting operator* of A_2 and A_1 (see [7, 18]).

Proof. Set $\Psi = \mathrm{Id}_{\mathcal{G}} - L \circ L^*$, $\mathcal{M} \colon \mathcal{G} \to 2^{\mathcal{G}} \colon (x, y) \mapsto (A_1 x) \times (A_2^{-1} y)$, and $S \colon \mathcal{G} \to \mathcal{G} \colon (x, y) \mapsto (y, -x)$. Note that $\Psi \colon (x, y) \mapsto (y, x)$, that S is monotone, and that, since A_1 and A_2 are monotone, $\mathcal{M}$ is monotone as well. Thus, [7, Proposition 20.10] implies that $B + \Psi = \mathcal{M} + S$ is monotone. Further, it is straightforward verify that

$$(\forall (x, y) \in \mathcal{G}) \quad J_B(x, y) = \big(J_{A_1} x, J_{A_2^{-1}}(y + 2J_{A_1} x)\big). \tag{3.42}$$

Therefore, it follows from Proposition 3.17(ii) that $L \diamond B$ is maximally monotone. In addition, since $L^*\colon \mathcal{G} \to \mathcal{H}\colon (x,y) \mapsto x - y$, Proposition 3.4(i), (3.42), and (3.5) yield

$$
\begin{aligned}
(\forall x \in \mathcal{H}) \quad J_{L \diamond B} x &= L^*\big(J_B(x,-x)\big) \\
&= L^*\big(J_{A_1}x, J_{A_2^{-1}}(2J_{A_1}x - x)\big) \\
&= J_{A_1}x - J_{A_2^{-1}}(2J_{A_1}x - x) \\
&= J_{A_1}x - \big(2J_{A_1}x - x - J_{A_2}(2J_{A_1}x - x)\big) \\
&= \frac{1}{2}x + \frac{1}{2}\big(2J_{A_2} - \mathrm{Id}_{\mathcal{H}}\big)(2J_{A_1}x - x),
\end{aligned} \tag{3.43}
$$

which establishes (3.41). ☐

Remark 3.21 Consider the setting of Example 3.20. Then the operator B is not necessarily monotone (take $A_1 = 0$ and $A_2 = N_{\{0\}}$) and the norm of the linear operator is greater than 1 ($\|L\| = \sqrt{2}$). As a result, the resolvent composition $L \diamond B$ can be maximally monotone, even in cases when B is not monotone and $\|L\| > 1$.

The following example recovers the operator used in [20] for finding a zero in the sum of $p \geqslant 2$ maximally monotone operators. For the sake of simplicity, we represent operators using matrices.

Example 3.22 Let $\gamma \in \,]0,+\infty[$, let $\mathcal{K}$ be a real Hilbert space, let $p \in \mathbb{N} \smallsetminus \{0,1\}$, and let $(A_k)_{1 \leqslant k \leqslant p}$ be a family of maximally monotone operators in $\mathcal{K}$. Let $\mathcal{H} = \bigoplus_{k=1}^{p-1} \mathcal{K}$, let $\mathcal{G} = \bigoplus_{k=1}^{p} \mathcal{K}$, set

$$
L = \begin{bmatrix}
\mathrm{Id}_{\mathcal{K}} & & & \\
-\mathrm{Id}_{\mathcal{K}} & \mathrm{Id}_{\mathcal{K}} & & \\
& \ddots & \ddots & \\
& & -\mathrm{Id}_{\mathcal{K}} & \mathrm{Id}_{\mathcal{K}} \\
& & & -\mathrm{Id}_{\mathcal{K}}
\end{bmatrix} \in \mathcal{B}\,(\mathcal{H},\mathcal{G}), \tag{3.44}
$$

and set

$$
B = \gamma \begin{bmatrix}
A_1 & & & & \\
-\mathrm{Id}_{\mathcal{K}} & A_2 & & & \\
& \ddots & \ddots & & \\
& & -\mathrm{Id}_{\mathcal{K}} & A_{p-1} & \\
-\mathrm{Id}_{\mathcal{K}} & & & -\mathrm{Id}_{\mathcal{K}} & A_p
\end{bmatrix} : \mathcal{G} \to 2^{\mathcal{G}}. \tag{3.45}
$$

Then, for every $\boldsymbol{z} = (z_k)_{1 \leqslant k \leqslant p-1} \in \mathcal{H}$,

$$J_{\gamma\left((\gamma L) \overset{\gamma}{\blacklozenge} B^{-1}\right)} \boldsymbol{z} = \boldsymbol{z} + \gamma^2 \begin{bmatrix} x_2 - x_1 \\ x_3 - x_2 \\ \vdots \\ x_p - x_{p-1} \end{bmatrix},$$

$$\text{where} \quad \begin{cases} x_1 = J_{A_1} z_1, \\ (\forall k \in \{2, \ldots, p-1\}) \; x_k = J_{A_k}(z_k + x_{k-1} - z_{k-1}), \\ x_p = J_{A_p}(x_1 + x_{p-1} - z_{p-1}). \end{cases} \tag{3.46}$$

Proof. Let $\boldsymbol{z} = (z_k)_{1 \leqslant k \leqslant p-1} \in \mathcal{H}$. It is straightforward verify that $J_{\gamma^{-1}B}(L\boldsymbol{z}) = (x_k)_{1 \leqslant k \leqslant p} = \boldsymbol{x}$. On the other hand, recall from [7, Proposition 23.20] that $\mathrm{Id}_{\mathcal{G}} - J_{\gamma B^{-1}} = \gamma J_{\gamma^{-1}B}(\gamma^{-1}\mathrm{Id}_{\mathcal{G}})$. Further, note that

$$L^* \boldsymbol{x} = \begin{bmatrix} x_1 - x_2 \\ x_2 - x_3 \\ \vdots \\ x_{p-1} - x_p \end{bmatrix} \tag{3.47}$$

Altogether, we deduce from Proposition 3.4(ii) and (3.47) that

$$J_{\gamma\left((\gamma L) \overset{\gamma}{\blacklozenge} B^{-1}\right)} \boldsymbol{z} = \boldsymbol{z} - (\gamma L)^* \big(\gamma J_{\gamma^{-1}B}(L\boldsymbol{z})\big) = \boldsymbol{z} - \gamma^2 L^* \boldsymbol{x} = \boldsymbol{z} + \gamma^2 \begin{bmatrix} x_2 - x_1 \\ x_3 - x_2 \\ \vdots \\ x_p - x_{p-1}, \end{bmatrix}, \tag{3.48}$$

which establishes (3.46). ☐

Remark 3.23 As shown in [20, Lemma 3], the operator given in (3.46) is γ^2-averaged when $\gamma \in \left]0, 1\right[$.

Example 3.24 Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ be such that $\operatorname{ran} L$ is closed, let $B \colon \mathcal{G} \to 2^{\mathcal{G}}$ be maximally monotone, and let $\gamma \in \left]0, +\infty\right[$. Let $\mathcal{X}$ be the real Hilbert space obtained by endowing $\mathcal{H}$ with the scalar product

$$\langle \cdot \mid \cdot \rangle_{\mathcal{X}} \colon \mathcal{X} \times \mathcal{X} \to \mathbb{R} \colon (x, y) \mapsto \langle Lx \mid Ly \rangle_{\mathcal{G}} + \langle x \mid \mathrm{proj}_{\ker L}\, y \rangle_{\mathcal{H}}, \tag{3.49}$$

and set $L_{\mathcal{X}} \colon \mathcal{X} \to \mathcal{G} \colon x \mapsto Lx$. Then the following hold:

(i) $L_{\mathcal{X}} \overset{\gamma}{\diamond} B$ and $L_{\mathcal{X}} \overset{\gamma}{\blacklozenge} B$ are maximally monotone.

(ii) $J_{\gamma\left(L_{\mathcal{X}} \overset{\gamma}{\diamond} B\right)} = L^{\dagger} \circ J_{\gamma B} \circ L$.

(iii) $J_{\gamma(L_{\mathcal{X}} \overset{\gamma}{\blacklozenge} B)} = \mathrm{Id}_{\mathcal{X}} - L^{\dagger} \circ (\mathrm{Id}_{\mathcal{G}} - J_{\gamma B}) \circ L$.

(iv) Suppose that $\ker L = \{0\}$. Then $L_{\mathcal{X}} \overset{\gamma}{\diamond} B = L_{\mathcal{X}} \overset{\gamma}{\blacklozenge} B$.

(v) Suppose that $\operatorname{ran} L = \mathcal{G}$. Then $L_{\mathcal{X}} \overset{\gamma}{\diamond} B = L^{\dagger} \rhd B$ and $L_{\mathcal{X}} \overset{\gamma}{\blacklozenge} B = L^{\dagger} \circ B \circ L$.

Proof. Let $x \in \mathcal{H}$ and $y \in \mathcal{G}$. It follows from (3.49) that $\|L_{\mathcal{X}}\| \leqslant 1$ since

$$(\forall z \in \mathcal{H}) \quad \|L_{\mathcal{X}} z\|_{\mathcal{G}}^2 = \|Lz\|_{\mathcal{G}}^2 \leqslant \|Lz\|_{\mathcal{G}}^2 + \|\operatorname{proj}_{\ker L} z\|_{\mathcal{H}}^2 = \|z\|_{\mathcal{X}}^2. \tag{3.50}$$

Further, the identities $L^* y = L^*(L(L^{\dagger} y))$ and $L^{\dagger} y \in (\ker L)^{\perp}$ [7, Proposition 3.30(i)] imply that

$$\begin{aligned} \langle L_{\mathcal{X}} x \mid y \rangle_{\mathcal{G}} &= \langle x \mid L^* y \rangle_{\mathcal{H}} \\ &= \langle x \mid L^*\big(L(L^{\dagger} y)\big) \rangle_{\mathcal{H}} \\ &= \langle Lx \mid L(L^{\dagger} y) \rangle_{\mathcal{G}} + \langle x \mid \operatorname{proj}_{\ker L}(L^{\dagger} y) \rangle_{\mathcal{H}} \\ &= \langle x \mid L^{\dagger} y \rangle_{\mathcal{X}}. \end{aligned} \tag{3.51}$$

In turn, $L_{\mathcal{X}}^* \colon \mathcal{G} \to \mathcal{X} \colon y \mapsto L^{\dagger} y$.

(i): A consequence of Corollary 3.18(ii).

(ii)–(iii): A consequence of Proposition 3.4.

(iv): By [7, Proposition Corollary 3.32(iv)], $L^{\dagger} \circ L = \mathrm{Id}_{\mathcal{H}}$. Therefore, $L_{\mathcal{X}}$ is an isometry, and the assertion follows from Proposition 3.6(i).

(v): By [7, Proposition 3.30(ii)], $L \circ L^{\dagger} = \mathrm{Id}_{\mathcal{G}}$. Therefore, $L_{\mathcal{X}}$ is a coisometry, and the assertion follows from Proposition 3.6(ii). ☐

Remark 3.25 When $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ satisfies $L^{\dagger} \circ L = \mathrm{Id}_{\mathcal{H}}$, the operator $T = L^{\dagger} \circ J_{\gamma \partial \|\cdot\|_{\mathcal{H}}} \circ L$ has been used to enhance the performance of wavelet-domain denoising [9]. Consequently, Example 3.24(ii) shows that this method implicitly involves resolvent compositions.

Example 3.26 Let $U \in \mathcal{B}(\mathcal{H})$ be a self-adjoint and strongly monotone operator. In the context of Example 3.24, assume that $\mathcal{G} = \mathcal{H}$ and that $L = U^{-1/2}$. Then

$$L_{\mathcal{X}} \overset{\gamma}{\diamond} B = U^{1/2} \circ B \circ U^{-1/2}. \tag{3.52}$$

Proof. In this case, L is invertible and $L^{\dagger} = L^{-1} = U^{1/2}$. Therefore, Example 3.24(iv)–(v) implies that $L_{\mathcal{X}} \overset{\gamma}{\diamond} B = L_{\mathcal{X}} \overset{\gamma}{\blacklozenge} B = U^{1/2} \circ B \circ U^{-1/2}$, as claimed. ☐

Example 3.27 Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ satisfies $0 < \|L\| \leqslant 1$, let $g \colon \mathcal{G} \to \left]-\infty, +\infty\right]$ be a proper function that admits a continuous affine minorant, let $\gamma \in \left]0, +\infty\right[$, let

$$L \overset{\gamma}{\diamond} g = \Big(\tfrac{1}{\gamma}(g^*) \circ L\Big)^* - \frac{1}{2\gamma}\|\cdot\|_{\mathcal{H}}^2 \tag{3.53}$$

be the *proximal composition* of g and L, and let $L \overset{\gamma}{\blacklozenge} g = (L \overset{1/\gamma}{\diamond} g^*)^*$ be the *proximal cocomposition* of g and L (see [8, 11, 14]). Then the following hold:

(i) $L \overset{\gamma}{\diamond} \partial g^{**} = \partial(L \overset{\gamma}{\diamond} g)$.

(ii) $L \overset{\gamma}{\blacklozenge} \partial g^{**} = \partial(L \overset{\gamma}{\blacklozenge} g)$.

Proof. Set $\Psi = \mathrm{Id}_{\mathcal{G}} - L \circ L^*$ and recall from [14, Lemma 2.1(v)] that $g^* \in \Gamma_0(\mathcal{G})$.

(i): By [14, Proposition 3.11(i)] and Proposition 3.5(i), $\partial(L \overset{\gamma}{\diamond} g) = L^* \triangleright (\partial g^{**} + \gamma^{-1}\Psi) = L \overset{\gamma}{\diamond} \partial g^{**}$.

(ii): Note that (3.11) and the identity $g^{***} = g^*$ yield $(\partial g^{**})^{-1} = \partial g^{***} = \partial g^*$. Therefore, it follows from [14, Proposition 3.11(ii)] and Proposition 3.5(ii) that $\partial(L \overset{\gamma}{\blacklozenge} g) = L^* \circ (\partial g^* + \gamma\Psi)^{-1} \circ L = L \overset{\gamma}{\blacklozenge} \partial g^{**}$, which provides the desired identity. ☐

We conclude this section by examining the resolvent composition of uniformly monotone operators as well as the Fitzpatrick function of resolvent compositions.

Proposition 3.28 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ satisfies $0 < \|L\| \leqslant 1$, let $B\colon \mathcal{G} \to 2^{\mathcal{G}}$ be maximally monotone, and let $\gamma \in \left]0, +\infty\right[$. Suppose that B is uniformly monotone with modulus $\phi\colon \left[0, +\infty\right[\to [0, +\infty]$, i.e., ϕ is increasing, vanishes only at 0, and*

$$\big(\forall (y_1, y_1^*) \in \mathrm{gra}\, B\big)\big(\forall (y_2, y_2^*) \in \mathrm{gra}\, B\big) \quad \langle y_1 - y_2 \,|\, y_1^* - y_2^* \rangle_{\mathcal{G}} \geqslant \phi(\|y_1 - y_2\|_{\mathcal{G}}), \tag{3.54}$$

and set $\phi_L = L \overset{\gamma/2}{\diamond} (\phi \circ \|\cdot\|_{\mathcal{G}})$. Then

$$\big(\forall (x_1, x_1^*) \in \mathrm{gra}(L \overset{\gamma}{\diamond} B)\big)\big(\forall (x_2, x_2^*) \in \mathrm{gra}(L \overset{\gamma}{\diamond} B)\big) \quad \langle x_1 - x_2 \,|\, x_1^* - x_2^* \rangle_{\mathcal{G}} \geqslant \phi_L(x_1 - x_2). \tag{3.55}$$

Proof. Note that $\phi \circ \|\cdot\|_{\mathcal{G}} \geqslant 0$ and that $\phi(\|0\|_{\mathcal{G}}) = 0$. Thus, $\phi \circ \|\cdot\|_{\mathcal{G}}$ is a proper function minorized by the affine function 0. Further, by [7, Proposition 13.16], $\phi \circ \|\cdot\|_{\mathcal{G}} \geqslant (\phi \circ \|\cdot\|_{\mathcal{G}})^{**}$. On the other hand, recall from Corollary 3.18(ii) that $L \overset{\gamma}{\diamond} B$ is maximally monotone. Let $(x_1, x_1^*) \in \mathrm{gra}(L \overset{\gamma}{\diamond} B)$ and $(x_2, x_2^*) \in \mathrm{gra}(L \overset{\gamma}{\diamond} B)$. It follows from Proposition 3.4(i) that

$$\begin{aligned}
(\forall k \in \{1,2\}) \quad x_k^* \in \big(L \overset{\gamma}{\diamond} B\big)x_k &\Leftrightarrow J_{\gamma(L \overset{\gamma}{\diamond} B)}(x_k + \gamma x_k^*) = x_k \\
&\Leftrightarrow \begin{cases} (\exists\, p_k \in \mathcal{G})\ L^* p_k = x_k \\ J_{\gamma B}\big(L(x_k + \gamma x_k^*)\big) = p_k \end{cases} \\
&\Leftrightarrow \begin{cases} (\exists\, p_k \in \mathcal{G})\ L^* p_k = x_k \\ \big(p_k, L(\gamma^{-1} x_k + x_k^*) - \gamma^{-1} p_k\big) \in \mathrm{gra}\, B. \end{cases}
\end{aligned} \tag{3.56}$$

Since B is uniformly monotone with modulus ϕ, we deduce that

$$\gamma^{-1}\|x_1 - x_2\|_{\mathcal{H}}^2 + \langle x_1 - x_2 \,|\, x_1^* - x_2^* \rangle_{\mathcal{H}} - \gamma^{-1}\|p_1 - p_2\|_{\mathcal{G}}^2 = \langle p_1 - p_2 \,|\, \gamma^{-1} L(x_1 - x_2) + L(x_1^* - x_2^*) \rangle_{\mathcal{G}}$$

$$
\begin{aligned}
&- \gamma^{-1}\langle p_1 - p_2 \,|\, p_1 - p_2\rangle_{\mathcal{G}} \\
&\geqslant \phi(\|p_1 - p_2\|_{\mathcal{G}}).
\end{aligned} \tag{3.57}
$$

Therefore, since $L^*(p_1 - p_2) = x_1 - x_2$, we deduce from (3.57) and [14, Proposition 3.2(i)] that

$$
\begin{aligned}
\langle x_1 - x_2 \,|\, x_1^* - x_2^*\rangle_{\mathcal{H}} &\geqslant \phi(\|p_1 - p_2\|_{\mathcal{G}}) + \gamma^{-1}\big(\|p_1 - p_2\|_{\mathcal{G}}^2 - \|x_1 - x_2\|_{\mathcal{H}}^2\big) \\
&\geqslant \inf_{\substack{v \in \mathcal{G} \\ L^*v = x_1 - x_2}} \Big(\phi(\|v\|_{\mathcal{G}}) + \gamma^{-1}\big(\|v\|_{\mathcal{G}} - \|L^*v\|_{\mathcal{H}}^2\big)\Big) \\
&\geqslant \inf_{\substack{v \in \mathcal{G} \\ L^*v = x_1 - x_2}} \Big(\big(\phi \circ \|\cdot\|_{\mathcal{G}}\big)^{**}(v) + \gamma^{-1}\big(\|v\|_{\mathcal{G}} - \|L^*v\|_{\mathcal{H}}^2\big)\Big) \\
&= \big(L \overset{\gamma/2}{\diamond} (\phi \circ \|\cdot\|_{\mathcal{G}})\big)(x_1 - x_2),
\end{aligned} \tag{3.58}
$$

which completes the proof. ☐

Proposition 3.29 (Fitzpatrick function) *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ satisfies $\|L\| \leqslant 1$, let $B\colon \mathcal{G} \to 2^{\mathcal{G}}$ be maximally monotone, let*

$$
F_B\colon \mathcal{G} \times \mathcal{G} \to [-\infty, +\infty]\colon (x, x^*) \mapsto \sup_{(v, v^*) \in \operatorname{gra} B} \big(\langle v \,|\, x^*\rangle_{\mathcal{G}} + \langle x \,|\, v^*\rangle_{\mathcal{G}} - \langle v \,|\, v^*\rangle_{\mathcal{G}}\big) \tag{3.59}
$$

be its Fitzpatrick function, and let $\gamma \in \left]0, +\infty\right[$. Then the following hold:

(i) *Let $x \in \ker(\mathrm{Id}_{\mathcal{H}} - L^* \circ L)$ and $x^* \in \mathcal{H}$. Then $F_{\gamma(L \overset{\gamma}{\diamond} B)}(x, x^*) \leqslant F_{\gamma B}(Lx, Lx^*)$.*

(ii) *Let $x \in \mathcal{H}$ and $x^* \in \ker(\mathrm{Id}_{\mathcal{H}} - L^* \circ L)$. Then $F_{\gamma(L \overset{\gamma}{\blacklozenge} B)}(x, x^*) \leqslant F_{\gamma B}(Lx, Lx^*)$.*

Proof. We recall from Corollary 3.18(ii) that $L \overset{\gamma}{\diamond} B$ and $L \overset{\gamma}{\blacklozenge} B$ are maximally monotone.

(i): By Minty's parametrization [7, Remark 23.23(ii)],

$$
F_{\gamma(L \overset{\gamma}{\diamond} B)}(x, x^*) = \sup_{y \in \mathcal{H}} \bigg(\Big\langle J_{\gamma(L \overset{\gamma}{\diamond} B)} y \,\Big|\, x^* \Big\rangle_{\mathcal{H}} + \Big\langle x \,\Big|\, y - J_{\gamma(L \overset{\gamma}{\diamond} B)} y \Big\rangle_{\mathcal{H}} - \Big\langle J_{\gamma(L \overset{\gamma}{\diamond} B)} y \,\Big|\, y - J_{\gamma(L \overset{\gamma}{\diamond} B)} y \Big\rangle_{\mathcal{H}}\bigg). \tag{3.60}
$$

Thus, by virtue of Proposition 3.4(i) and $\|L\| \leqslant 1$, we deduce that, for every $y \in \mathcal{H}$,

$$
\begin{aligned}
&\Big\langle x \,\Big|\, y - J_{\gamma(L \overset{\gamma}{\diamond} B)} y \Big\rangle_{\mathcal{H}} + \Big\langle J_{\gamma(L \overset{\gamma}{\diamond} B)} y \,\Big|\, x^* \Big\rangle_{\mathcal{H}} - \Big\langle J_{\gamma(L \overset{\gamma}{\diamond} B)} y \,\Big|\, y - J_{\gamma(L \overset{\gamma}{\diamond} B)} y \Big\rangle_{\mathcal{H}} \\
&= \Big\langle x \,\Big|\, y - L^*\big(J_{\gamma B}(Ly)\big) \Big\rangle_{\mathcal{H}} + \Big\langle L^*\big(J_{\gamma B}(Ly)\big) \,\Big|\, x^* \Big\rangle_{\mathcal{H}} - \Big\langle L^*\big(J_{\gamma B}(Ly)\big) \,\Big|\, y - L^*\big(J_{\gamma B}(Ly)\big) \Big\rangle_{\mathcal{H}} \\
&\leqslant \langle x - L^*(Lx) \,|\, y\rangle_{\mathcal{H}} + \Big(\langle Lx \,|\, Ly - J_{\gamma B}(Ly)\rangle_{\mathcal{G}} + \langle J_{\gamma B}(Ly) \,|\, Lx^*\rangle_{\mathcal{G}} - \langle J_{\gamma B}(Ly) \,|\, Ly - J_{\gamma B}(Ly)\rangle_{\mathcal{G}}\Big) \\
&\leqslant \sup_{v \in \mathcal{G}} \Big(\langle Lx \,|\, v - J_{\gamma B} v\rangle_{\mathcal{G}} + \langle J_{\gamma B} v \,|\, Lx^*\rangle_{\mathcal{G}} - \langle J_{\gamma B} v \,|\, v - J_{\gamma B} v\rangle_{\mathcal{G}}\Big) \\
&= F_{\gamma B}(Lx, Lx^*).
\end{aligned} \tag{3.61}
$$

Therefore, taking the supremum over $y \in \mathcal{H}$ in (3.61), the conclusion follows from (3.60).

(ii): By Proposition 3.3(vi), Definition 3.1, (3.59), and (i),

$$F_{\gamma(L \overset{\gamma}{\blacklozenge} B)}(x, x^*) = F_{(L \diamond (\gamma B)^{-1})^{-1}}(x, x^*) = F_{L \diamond (\gamma B)^{-1}}(x^*, x)$$
$$\leqslant F_{(\gamma B)^{-1}}(Lx^*, Lx) = F_{\gamma B}(Lx, Lx^*), \tag{3.62}$$

which completes the proof. ☐

Corollary 3.30 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ is an isometry, let $B \colon \mathcal{G} \to 2^{\mathcal{G}}$ be maximally monotone, and let $\gamma \in \left]0, +\infty\right[$. Then*

$$(\forall x \in \mathcal{H})(\forall x^* \in \mathcal{H}) \quad F_{\gamma(L \overset{\gamma}{\diamond} B)}(x, x^*) \leqslant F_{\gamma B}(Lx, Lx^*) \tag{3.63}$$

Proof. Since L is an isometry, $\ker(\mathrm{Id}_{\mathcal{G}} - L^* \circ L) = \mathcal{H}$. Therefore, the conclusion is a consequence of Proposition 3.29(i). ☐

Corollary 3.31 ([6, Theorem 2.13]) *Let $0 \neq p \in \mathbb{N}$ and let $\gamma \in \left]0, +\infty\right[$. For every $k \in \{1, \ldots, p\}$, let $B_k \colon \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone and let $\alpha_k \in \left]0, +\infty\right[$. Suppose that $\sum_{k=1}^{p} \alpha_k = 1$. Then*

$$F_{\gamma \, \mathsf{rav}_{\gamma}(B_k, \alpha_k)_{1 \leqslant k \leqslant p}} \leqslant \sum_{k=1}^{p} \alpha_k F_{\gamma B_k}. \tag{3.64}$$

Proof. Define L and B as in Example 3.13 and recall that $L \overset{\gamma}{\diamond} B = \mathsf{rav}_{\gamma}(B_k, \alpha_k)_{1 \leqslant k \leqslant p}$. In this case, L is an isometry, and it follows from Corollary 3.30 that, for every $x \in \mathcal{H}$ and $x^* \in \mathcal{H}$,

$$F_{\gamma \, \mathsf{rav}_{\gamma}(B_k, \alpha_k)_{1 \leqslant k \leqslant p}}(x, x^*) \leqslant F_{\gamma B}(Lx, Lx^*) = \sum_{k=1}^{p} \alpha_k F_{\gamma B_k}(x, x^*), \tag{3.65}$$

as announced. ☐

3.1.6 Asymptotic behavior of resolvent compositions

We examine the convergence of the operators $L \overset{\gamma}{\diamond} B$ and $L \overset{\gamma}{\blacklozenge} B$ when γ varies, studying their corresponding graph. We begin by recalling some definitions related to set-convergence, which enable us to characterize the convergence of operators through their graphs.

3.1.6.1 Set-convergence

Definition 3.32 (Painlevé–Kuratowski) Let $(C_n)_{n \in \mathbb{N}}$ be a sequence of subsets of $\mathcal{H}$. The *lower limit* of the sequence $(C_n)_{n \in \mathbb{N}}$ is the closed subset of $\mathcal{H}$ defined by

$$\underline{\lim}\, C_n = \big\{ x \in \mathcal{H} \mid (\exists\, (x_n)_{n \in \mathbb{N}} \text{ in } \mathcal{H})(\forall n \in \mathbb{N})\ x_n \in C_n \text{ and } x_n \to x \big\}. \tag{3.66}$$

The *upper limit* of the sequence $(C_n)_{n\in\mathbb{N}}$ is the closed subset of $\mathcal{H}$ defined by

$$\overline{\lim}\, C_n = \big\{x \in \mathcal{H} \mid (\exists\,(x_n)_{n\in\mathbb{N}} \text{ in } \mathcal{H})(\exists\,(k_n)_{n\in\mathbb{N}} \text{ in } \mathbb{N})(\forall n \in \mathbb{N})\ x_n \in C_{k_n} \text{ and } x_n \to x\big\}. \tag{3.67}$$

The sequence $(C_n)_{n\in\mathbb{N}}$ is *Painlevé–Kuratowski* convergent if its upper limit coincides with its lower limit. The limit set in this case is given by

$$\lim_{n\to+\infty} C_n = \underline{\lim}\, C_n = \overline{\lim}\, C_n. \tag{3.68}$$

Let C and D be subsets of $\mathcal{H}$. The *excess function* of C on D is defined by

$$e(C, D) = \sup_{x\in C} d_D(x), \tag{3.69}$$

with the convention that $e(\varnothing, D) = 0$.

Definition 3.33 (ρ-Hausdorff distance [3, 5]) Let C and D be subsets of $\mathcal{H}$, let $\rho \in [0, +\infty[$, and set $C_\rho = C \cap B(0; \rho)$ and $D_\rho = D \cap B(0; \rho)$ The ρ-*Hausdorff distance* between C and D is

$$\text{haus}_\rho(C, D) = \max\{e(C_\rho, D), e(D_\rho, C)\}. \tag{3.70}$$

A sequence $(C_n)_{n\in\mathbb{N}}$ of subsets of $\mathcal{H}$ converges with respect to the ρ-Hausdorff distance to the subset C of $\mathcal{H}$ if

$$(\forall \rho \in\,]0, +\infty[) \qquad \lim_{n\to+\infty} \text{haus}_\rho(C_n, C) = 0. \tag{3.71}$$

3.1.6.2 Graph-convergence of operators

Definition 3.34 Let $(A_n)_{n\in\mathbb{N}}$ and A be set-valued operators from $\mathcal{H}$ to $2^{\mathcal{H}}$. Then $(A_n)_{n\in\mathbb{N}}$ *graph-converges* to A, denoted by $A_n \xrightarrow{g} A$, if $(\text{gra}\, A_n)_{n\in\mathbb{N}}$ converges to $\text{gra}\, A$ in the Painlevé–Kuratowski sense.

Definition 3.35 Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$, $B\colon \mathcal{H} \to 2^{\mathcal{H}}$, and $\rho \in [0, +\infty[$. The ρ-*Hausdorff distance* between A and B is

$$\text{haus}_\rho(A, B) = \text{haus}_\rho(\text{gra}\, A, \text{gra}\, B). \tag{3.72}$$

A sequence $(A_n)_{n\in\mathbb{N}}$ of operators from $\mathcal{H}$ to $2^{\mathcal{H}}$ converges with respect to the ρ-Hausdorff distance to the operator $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ if

$$(\forall \rho \in\,]0, +\infty[) \qquad \lim_{n\to+\infty} \text{haus}_\rho(A_n, A) = 0. \tag{3.73}$$

Some equivalences of graph-convergence for maximally monotone operators are summarized in the following result (see e.g. [2]).

Lemma 3.36 *Let $(A_n)_{n\in\mathbb{N}}$ and A be maximally monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$. Then the following are equivalent:*

(i) $A_n \xrightarrow{g} A$.

(ii) $(\forall\gamma \in\,]0,+\infty[)(\forall x\in\mathcal{H})\ J_{\gamma A_n}x \to J_{\gamma A}x$.

(iii) $(\exists\,\gamma \in\,]0,+\infty[)(\forall x\in\mathcal{H})\ J_{\gamma A_n}x \to J_{\gamma A}x$.

Lemma 3.37 *Let $(A_n)_{n\in\mathbb{N}}$ and A be maximally monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$, and let $(\gamma_n)_{n\in\mathbb{N}}$ and γ be in $]0,+\infty[$. Suppose that $A_n \xrightarrow{g} A$ and $\gamma_n\to\gamma$. Then the following hold:*

(i) $A_n^{-1} \xrightarrow{g} A^{-1}$.

(ii) $\gamma_n A_n \xrightarrow{g} \gamma A$

Proof. (i): This follows from (3.5) and Lemma 3.36.

(ii): Let $x\in\mathcal{H}$ and set $(\forall n\in\mathbb{N})\ \theta_n = 1-\gamma_n/\gamma$. By [7, Proposition 23.31(iii)],

$$\begin{aligned}\|J_{\gamma_n A_n}x - J_{\gamma A}x\|_{\mathcal{H}} &\leqslant \|J_{\gamma_n A_n}x - J_{\gamma A_n}x\|_{\mathcal{H}} + \|J_{\gamma A_n}x - J_{\gamma A}x\|_{\mathcal{H}}\\ &\leqslant |\theta_n|\|x - J_{\gamma A_n}x\|_{\mathcal{H}} + \|J_{\gamma A_n}x - J_{\gamma A}x\|_{\mathcal{H}}. \end{aligned}\tag{3.74}$$

Further, $A_n \xrightarrow{g} A$ and Lemma 3.36 yield $J_{\gamma A_n}x\to J_{\gamma A}x$. Altogether, since $\theta_n\to 0$, we deduce from (3.74) that $J_{\gamma_n A_n}x \to J_{\gamma A}x$. Finally, invoking Lemma 3.36 one more, we obtain the assertion. ☐

Lemma 3.38 ([4, Propositions 1.1 and 1.2]) *Let $A_1\colon\mathcal{H}\to 2^{\mathcal{H}}$ and $A_2\colon\mathcal{H}\to 2^{\mathcal{H}}$ be maximally monotone, and let $\gamma\in\,]0,+\infty[$. Consider*

$$(\forall\delta\in[0,+\infty[)\quad d_{\gamma,\delta}(A_1,A_2) = \sup_{x\in B(0;\delta)} \|J_{\gamma A_1}x - J_{\gamma A_2}x\|_{\mathcal{H}}. \tag{3.75}$$

Then the following hold:

(i) $(\forall\rho\in[0,+\infty[)\ \operatorname{haus}_\rho(A_1,A_2)\leqslant \max\{1,\gamma^{-1}\}d_{\gamma,(1+\gamma)\rho}(A_1,A_2)$.

(ii) *Set $\rho=\max\{\delta+\|J_{\gamma A_1}0\|_{\mathcal{H}},\gamma^{-1}(\delta+\|J_{\gamma A_1}0\|_{\mathcal{H}}\}$. Then*

$$d_{\gamma,\delta}(A_1,A_2)\leqslant (2+\gamma)\operatorname{haus}_\rho(A_1,A_2). \tag{3.76}$$

3.1.6.3 Convergence of resolvent compositions

We proceed to study the graph-convergence and convergence with respect to the ρ-Hausdorff distance of resolvent compositions.

Proposition 3.39 *Let $(L_n)_{n\in\mathbb{N}}$ and L be in $\mathcal{B}(\mathcal{H},\mathcal{G})$, let $(B_n)_{n\in\mathbb{N}}$ and B be maximally monotone operators from $\mathcal{G}$ to $2^{\mathcal{G}}$, and let $(\gamma_n)_{n\in\mathbb{N}}$ and γ be in $]0,+\infty[$. Suppose that $L_n\to L$, $B_n \xrightarrow{g} B$, $\gamma_n\to\gamma$, and $(\forall n\in\mathbb{N})\ \|L_n\|\leqslant 1$. Then the following hold:*

(i) $L_n \overset{\gamma_n}{\diamond} B_n \overset{g}{\to} L \overset{\gamma}{\diamond} B$.

(ii) $L_n \overset{\gamma_n}{\blacklozenge} B_n \overset{g}{\to} L \overset{\gamma}{\blacklozenge} B$.

Proof. We recall from Corollary 3.18(ii) that the operators $(L_n \overset{\gamma_n}{\diamond} B_n)_{n\in\mathbb{N}}$, $L \overset{\gamma}{\diamond} B$, $(L_n \overset{\gamma_n}{\blacklozenge} B_n)_{n\in\mathbb{N}}$, and $L \overset{\gamma}{\blacklozenge} B$ are maximally monotone. Therefore, by Lemma 3.36, it is enough to verify the convergence of the resolvent of these operators.

(i): Let $x \in \mathcal{H}$ and set $(\forall n \in \mathbb{N})\ \theta_n = 1 - \gamma/\gamma_n$. It follows from [7, Proposition 23.31(iii)], Proposition 3.4(i), and Lemma 3.37 that

$$
\begin{aligned}
&\|J_{\gamma(L_n \overset{\gamma_n}{\diamond} B_n)}x - J_{\gamma(L \overset{\gamma}{\diamond} B)}x\|_{\mathcal{H}} \\
&\leqslant \|J_{\gamma(L_n \overset{\gamma_n}{\diamond} B_n)}x - J_{\gamma_n(L_n \overset{\gamma_n}{\diamond} B_n)}x\|_{\mathcal{H}} + \|J_{\gamma_n(L_n \overset{\gamma_n}{\diamond} B_n)}x - J_{\gamma(L \overset{\gamma}{\diamond} B)}x\|_{\mathcal{H}} \\
&\leqslant |\theta_n|\, \|x - J_{\gamma_n(L_n \overset{\gamma_n}{\diamond} B_n)}x\|_{\mathcal{H}} + \|L_n^*\big(J_{\gamma B_n}(L_n x)\big) - L^*\big(J_{\gamma B}(Lx)\big)\|_{\mathcal{H}} \\
&= |\theta_n|\, \|x - L_n^*\big(J_{\gamma_n B_n}(L_n x)\big)\|_{\mathcal{H}} + \|L_n^*\big(J_{\gamma B_n}(L_n x)\big) - L^*\big(J_{\gamma B}(Lx)\big)\|_{\mathcal{H}}.
\end{aligned}
\tag{3.77}
$$

Further, nonexpansiveness of $J_{\gamma_n B_n}$ implies that

$$
\begin{aligned}
\|J_{\gamma_n B_n}(L_n x) - J_{\gamma B}(Lx)\|_{\mathcal{G}} &\leqslant \|J_{\gamma_n B_n}(L_n x) - J_{\gamma_n B_n}(Lx)\|_{\mathcal{G}} + \|J_{\gamma_n B_n}(Lx) - J_{\gamma B}(Lx)\|_{\mathcal{G}} \\
&\leqslant \|L_n x - Lx\|_{\mathcal{G}} + \|J_{\gamma_n B_n}(Lx) - J_{\gamma B}(Lx)\|_{\mathcal{G}}.
\end{aligned}
\tag{3.78}
$$

On the other hand, by Lemma 3.37(ii), $\gamma_n B \overset{g}{\to} \gamma B$. Since $\theta_n \to 0$ and $L_n \to L$, we combine (3.77) and (3.78) to obtain $J_{\gamma(L_n \overset{\gamma_n}{\diamond} B_n)}x \to J_{\gamma(L \overset{\gamma}{\diamond} B)}x$. Therefore, the conclusion follows from Lemma 3.36.

(ii): By Lemma 3.37(i), $B_n^{-1} \overset{g}{\to} B^{-1}$. Therefore, (i) yields $L_n \overset{1/\gamma_n}{\diamond} B_n^{-1} \overset{g}{\to} L \overset{1/\gamma}{\diamond} B^{-1}$. Altogether, by Definition 3.1 and Lemma 3.37(i) once more,

$$
L_n \overset{\gamma_n}{\blacklozenge} B_n = \big(L_n \overset{1/\gamma_n}{\diamond} B_n^{-1}\big)^{-1} \overset{g}{\to} \big(L \overset{1/\gamma}{\diamond} B^{-1}\big)^{-1} = L \overset{\gamma}{\blacklozenge} B,
\tag{3.79}
$$

as asserted. ☐

Proposition 3.40 *Let $(L_n)_{n\in\mathbb{N}}$ and L be in $\mathcal{B}(\mathcal{H},\mathcal{G})$, let $B\colon \mathcal{G} \to 2^{\mathcal{G}}$ be maximally monotone, and let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $]0,+\infty[$. Suppose that $L_n \to L$ and that $(\forall n \in \mathbb{N})\ \|L_n\| \leqslant 1$. Then the following hold:*

(i) *Suppose that $\gamma_n \downarrow 0$. Then the following are satisfied:*

  (a) $L_n \diamond (\gamma_n B) \overset{g}{\to} L \diamond N_{\overline{\operatorname{dom} B}}$.

  (b) $L_n \blacklozenge (\gamma_n B) \overset{g}{\to} L \blacklozenge N_{\overline{\operatorname{dom} B}}$.

(ii) *Suppose that $\gamma_n \uparrow +\infty$ and that* $\operatorname{zer} B \neq \varnothing$. *Then the following are satisfied:*

  (a) $L_n \diamond (\gamma_n B) \overset{g}{\to} L \diamond N_{\operatorname{zer} B}$.

(b) $L_n \blacklozenge (\gamma_n B) \overset{g}{\to} L \blacklozenge N_{\operatorname{zer} B}$.

Proof. (i): Let $y \in \mathcal{G}$. We recall from [7, Corollary 21.14] that $\overline{\operatorname{dom}}\, B$ is closed and convex. Further, by [7, Theorem 23.48], $J_{\gamma_n B} y \to \operatorname{proj}_{\overline{\operatorname{dom}}\, B}\, y = J_{N_{\overline{\operatorname{dom}}\, B}} y$. Therefore, it follows from Lemma 3.36 that $\gamma_n B \overset{g}{\to} N_{\overline{\operatorname{dom}}\, B}$, and the conclusion is a consequence of Proposition 3.39.

(ii): Let $y \in \mathcal{G}$. We recall from [7, Proposition 23.39] that $\operatorname{zer} B$ is closed and convex. Further, by [7, Theorem 23.48], $J_{\gamma_n B} y \to \operatorname{proj}_{\operatorname{zer} B}\, y = J_{N_{\operatorname{zer} B}} y$. Therefore, it follows from Lemma 3.36 that $\gamma_n B \overset{g}{\to} N_{\operatorname{zer} B}$, and the conclusion is a consequence of Proposition 3.39. ☐

The following proposition shows that, for a fixed $\gamma \in \left]0, +\infty\right[$, resolvent compositions are nonexpansive with respect to $d_{\gamma,\delta}$.

Proposition 3.41 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ satisfies $\|L\| \leqslant 1$, let $B_1 \colon \mathcal{G} \to 2^{\mathcal{G}}$ and $B_2 \colon \mathcal{G} \to 2^{\mathcal{G}}$ be maximally monotone, let $\gamma \in \left]0, +\infty\right[$, and let $\delta \in \left]0, +\infty\right[$. Then the following hold:*

(i) $d_{\gamma,\delta}(L \overset{\gamma}{\diamond} B_1, L \overset{\gamma}{\diamond} B_2) = d_{\gamma,\delta}(L \overset{\gamma}{\blacklozenge} B_1, L \overset{\gamma}{\blacklozenge} B_2)$.

(ii) $d_{\gamma,\delta}(L \overset{\gamma}{\diamond} B_1, L \overset{\gamma}{\diamond} B_2) \leqslant \|L\|\, d_{\gamma,\|L\|\delta}(B_1, B_2)$.

(iii) $d_{\gamma,\delta}(L \overset{\gamma}{\blacklozenge} B_1, L \overset{\gamma}{\blacklozenge} B_2) \leqslant \|L\|\, d_{\gamma,\|L\|\delta}(B_1, B_2)$.

Proof. We recall from Corollary 3.18(ii) that, for every $k \in \{1, 2\}$, $L \overset{\gamma}{\diamond} B_k$ and $L \overset{\gamma}{\blacklozenge} B_k$ are maximally monotone.

(i): By Proposition 3.4, $J_{\gamma(L \overset{\gamma}{\diamond} B_1)} - J_{\gamma(L \overset{\gamma}{\diamond} B_2)} = J_{\gamma(L \overset{\gamma}{\blacklozenge} B_1)} - J_{\gamma(L \overset{\gamma}{\blacklozenge} B_2)}$. Therefore, the conclusion follows from (3.75).

(ii): Let $x \in B(0; \delta)$. It follows from Proposition 3.4(i) that

$$
\begin{aligned}
\|J_{\gamma(L \overset{\gamma}{\diamond} B_1)} x - J_{\gamma(L \overset{\gamma}{\diamond} B_2)} x\|_{\mathcal{H}} &= \|L^*\big(J_{\gamma B_1}(Lx)\big) - L^*\big(J_{\gamma B_2}(Lx)\big)\|_{\mathcal{H}} \\
&\leqslant \|L\|\, \|J_{\gamma B_1}(Lx) - J_{\gamma B_2}(Lx)\|_{\mathcal{G}} \\
&\leqslant \|L\| \sup_{u \in B(0; \|L\|\delta)} \|J_{\gamma B_1} u - J_{\gamma B_2} u\|_{\mathcal{G}} \\
&= \|L\|\, d_{\gamma,\|L\|\delta}(B_1, B_2).
\end{aligned}
\tag{3.80}
$$

Therefore, by taking the supremum over all $x \in B(0; \delta)$, we obtain the assertion.

(iii): A consequence of (i) and (ii). ☐

Proposition 3.42 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ satisfies $\|L\| \leqslant 1$, let $B_1 \colon \mathcal{G} \to 2^{\mathcal{G}}$ and $B_2 \colon \mathcal{G} \to 2^{\mathcal{G}}$ be maximally monotone, let $\gamma \in \left]0, +\infty\right[$, and let $\rho \in \left]0, +\infty\right[$. Then*

$$
\operatorname{haus}_{\rho}(L \overset{\gamma}{\diamond} B_1, L \overset{\gamma}{\diamond} B_2) \leqslant \max\{1, \gamma^{-1}\} \|L\| d_{\gamma,\|L\|(1+\gamma)\rho}(B_1, B_2). \tag{3.81}
$$

Proof. Combine Corollary 3.18(ii), Lemma 3.38(i), and Proposition 3.41(ii). ☐

Proposition 3.43 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ satisfies $\|L\| \leqslant 1$ and let $B \colon \mathcal{G} \to 2^{\mathcal{G}}$ be maximally monotone. Assume that $L^* \circ B \circ L$ is maximally monotone. Then the following hold:*

(i) *$L \overset{\gamma}{\blacklozenge} B \overset{g}{\to} L^* \circ B \circ L$ as $0 < \gamma \to 0$.*

(ii) *Suppose that one of the following is satisfied:*

(a) *$\operatorname{ran}(B \circ L)$ is bounded.*

(b) *There exists $S \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that $S \circ L^*$ is invertible.*

Then

$$(\forall \rho \in \left]0, +\infty\right[) \quad \lim_{\gamma \to 0} \operatorname{haus}_\rho(L \overset{\gamma}{\blacklozenge} B, L^* \circ B \circ L) = 0. \tag{3.82}$$

Proof. Let $\gamma \in \left]0, 1\right[$ and recall from Corollary 3.18(ii) that $L \overset{\gamma}{\blacklozenge} B$ is maximally monotone. Let $x \in \mathcal{H}$, let $\rho \in \left]0, +\infty\right[$, and suppose that $x \in B(0; 2\rho)$. Set $\Psi = \mathrm{Id}_\mathcal{G} - L \circ L^*$, set $p = J_{L^* \circ B \circ L} x$, and set $p_\gamma = J_{L \overset{\gamma}{\blacklozenge} B} x$. We deduce from Proposition 3.5(ii) that

$$\begin{aligned} x - p_\gamma \in \big(L \overset{\gamma}{\blacklozenge} B\big) p_\gamma &\Leftrightarrow x - p_\gamma \in L^*\Big(\big(B^{-1} + \gamma\Psi\big)^{-1}(Lp_\gamma)\Big) \\ &\Leftrightarrow \begin{cases} (\exists\, y_\gamma \in \mathcal{G}) \quad x - p_\gamma = L^* y_\gamma \\ Lp_\gamma \in \big(B^{-1} + \gamma\Psi\big) y_\gamma \end{cases} \\ &\Leftrightarrow \begin{cases} (\exists\, y_\gamma \in \mathcal{G}) \quad x - p_\gamma = L^* y_\gamma \\ y_\gamma \in B(Lp_\gamma - \gamma\Psi y_\gamma). \end{cases} \end{aligned} \tag{3.83}$$

On the other hand,

$$x - p \in L^*\big(B(Lp)\big) \Leftrightarrow (\exists\, y \in \mathcal{G}) \quad x - p = L^* y \quad \text{and} \quad y \in B(Lp). \tag{3.84}$$

Altogether, monotonicity of B, (3.83), and (3.84) yield

$$\begin{aligned} \langle (Lp_\gamma - \gamma\Psi y_\gamma) - Lp \,|\, y_\gamma - y \rangle_\mathcal{G} \geqslant 0 &\Leftrightarrow \langle p_\gamma - p \,|\, L^*(y_\gamma - y) \rangle_\mathcal{H} - \gamma \langle \Psi y_\gamma \,|\, y_\gamma - y \rangle_\mathcal{G} \geqslant 0 \\ &\Leftrightarrow \langle p_\gamma - p \,|\, p - p_\gamma \rangle_\mathcal{H} - \gamma \langle \Psi y_\gamma \,|\, y_\gamma - y \rangle_\mathcal{G} \geqslant 0 \\ &\Leftrightarrow \|p_\gamma - p\|_\mathcal{H}^2 + \gamma \langle \Psi y_\gamma \,|\, y_\gamma - y \rangle_\mathcal{G} \leqslant 0. \end{aligned} \tag{3.85}$$

Further, since $L^* y_\gamma = x - p_\gamma$ and $L^* y = x - p$, by Cauchy–Schwarz inequality [7, Fact 2.11],

$$\begin{aligned} \langle \Psi y_\gamma \,|\, y_\gamma - y \rangle_\mathcal{G} &= \langle y_\gamma - L(x - p_\gamma) \,|\, y_\gamma - y \rangle_\mathcal{G} \\ &= \langle y_\gamma \,|\, y_\gamma - y \rangle_\mathcal{G} - \langle x - p_\gamma \,|\, L^*(y_\gamma - y) \rangle_\mathcal{H} \\ &= \|y_\gamma\|_\mathcal{G}^2 - \langle y_\gamma \,|\, y \rangle_\mathcal{G} - \langle x - p_\gamma \,|\, p - p_\gamma \rangle_\mathcal{H} \\ &\geqslant \|y_\gamma\|_\mathcal{G}^2 - \|y_\gamma\|_\mathcal{G} \, \|y\|_\mathcal{G} - \big(\langle p - p_\gamma \,|\, p - p_\gamma \rangle_\mathcal{H} + \langle x - p \,|\, p - p_\gamma \rangle_\mathcal{H}\big) \\ &\geqslant \|y_\gamma\|_\mathcal{G}^2 - \|y_\gamma\|_\mathcal{G} \, \|y\|_\mathcal{G} - \|p - p_\gamma\|_\mathcal{H}^2 - \|x - p\|_\mathcal{H} \, \|p - p_\gamma\|_\mathcal{H} \\ &\geqslant \min_{\alpha \in \mathbb{R}} \big(\alpha^2 - \alpha \|y\|_\mathcal{G}\big) - \|p - p_\gamma\|_\mathcal{H}^2 - \|x - p\|_\mathcal{H} \, \|p - p_\gamma\|_\mathcal{H} \end{aligned}$$

$$= -\frac{1}{4}\|y\|_{\mathcal{G}}^2 - \|p - p_\gamma\|_{\mathcal{H}}^2 - \|x - p\|_{\mathcal{H}} \|p - p_\gamma\|_{\mathcal{H}}. \tag{3.86}$$

Set $\delta = 2\rho + \|J_{L^* \circ B \circ L} 0\|_{\mathcal{H}}$. Since $L^* \circ B \circ L$ is maximally monotone, nonexpansiveness of $J_{L^* \circ B \circ L}$ yields

$$\begin{aligned}\|p\|_{\mathcal{H}} &\leqslant \|J_{L^* \circ B \circ L} x - J_{L^* \circ B \circ L} 0\|_{\mathcal{H}} + \|J_{L^* \circ B \circ L} 0\|_{\mathcal{H}} \leqslant \|x\|_{\mathcal{H}} + \|J_{L^* \circ B \circ L} 0\|_{\mathcal{H}} \\ &\leqslant 2\rho + \|J_{L^* \circ B \circ L} 0\|_{\mathcal{H}} = \delta. \end{aligned} \tag{3.87}$$

Thus, we combine (3.85), (3.86), and (3.87) to deduce that

$$\begin{aligned} & \|p - p_\gamma\|_{\mathcal{H}}^2 - \frac{\gamma}{4}\|y\|_{\mathcal{G}}^2 - \gamma\|p - p_\gamma\|_{\mathcal{H}}^2 - \gamma\|x - p\|_{\mathcal{H}} \|p - p_\gamma\|_{\mathcal{H}} \leqslant 0 \\ & \Leftrightarrow (1-\gamma)\|p - p_\gamma\|_{\mathcal{H}}^2 - \gamma\|x - p\|_{\mathcal{H}} \|p - p_\gamma\|_{\mathcal{H}} - \frac{\gamma}{4}\|y\|_{\mathcal{G}}^2 \leqslant 0 \\ & \Leftrightarrow \|p - p_\gamma\|_{\mathcal{H}}^2 - \frac{\gamma}{1-\gamma}\|x - p\|_{\mathcal{H}} \|p - p_\gamma\|_{\mathcal{H}} - \frac{\gamma}{4(1-\gamma)}\|y\|_{\mathcal{G}}^2 \leqslant 0 \\ & \Rightarrow \|p - p_\gamma\|_{\mathcal{H}}^2 - \frac{\gamma}{1-\gamma}(\|x\|_{\mathcal{H}} + \|p\|_{\mathcal{H}})\|p - p_\gamma\|_{\mathcal{H}} - \frac{\gamma}{4(1-\gamma)}\|y\|_{\mathcal{G}}^2 \leqslant 0 \\ & \Rightarrow \|p - p_\gamma\|_{\mathcal{H}}^2 - \frac{\gamma}{1-\gamma}(2\rho + \delta)\|p - p_\gamma\|_{\mathcal{H}} - \frac{\gamma}{4(1-\gamma)}\|y\|_{\mathcal{G}}^2 \leqslant 0 \\ & \Rightarrow \left(\|p - p_\gamma\|_{\mathcal{H}} - \frac{\gamma}{2(1-\gamma)}(2\rho + \delta)\right)^2 \leqslant \frac{\gamma^2}{4(1-\gamma)^2}(2\rho+\delta)^2 + \frac{\gamma}{4(1-\gamma)}\|y\|_{\mathcal{G}}^2 \\ & \Rightarrow \|p - p_\gamma\|_{\mathcal{H}} \leqslant \left(\frac{\gamma^2}{4(1-\gamma)^2}(2\rho+\delta)^2 + \frac{\gamma}{4(1-\gamma)}\|y\|_{\mathcal{G}}^2\right)^{1/2} + \frac{\gamma}{2(1-\gamma)}(2\rho+\delta). \end{aligned} \tag{3.88}$$

(i): By (3.88), $J_{L \overset{\gamma}{\blacklozenge} B} x \to J_{L^* \circ B \circ L} x$ as $0 < \gamma \to 0$, and the conclusion follows from Lemma 3.36.

(ii): Assumption (ii)(a) implies that there exits $\eta \in]0, +\infty[$ such that, for every $z \in \operatorname{ran}(B \circ L)$, $\|z\|_{\mathcal{G}} \leqslant \eta$. In particular, (3.84) yields $\|y\|_{\mathcal{G}} \leqslant \eta$. On the other hand, Assumption (ii)(b) and (3.84) imply that $y = (S \circ L^*)^{-1}(S(x - p))$. Thus, $\|y\|_{\mathcal{G}} \leqslant \|(S \circ L^*)^{-1}\| \, \|S\|(2\rho + \delta)$. Therefore, either Assumption (ii)(a) or Assumption (ii)(b) implies that there exists $\eta \in]0, +\infty[$ such that $\|y\|_{\mathcal{G}} \leqslant \eta$. Altogether, we deduce from Lemma 3.38(i) and (3.88) that

$$\begin{aligned} \operatorname{haus}_\rho(L \overset{\gamma}{\blacklozenge} B, L^* \circ B \circ L) &\leqslant d_{1,2\rho}(L \overset{\gamma}{\blacklozenge} B, L^* \circ B \circ L) \\ &= \sup_{u \in B(0;2\rho)} \|J_{L \overset{\gamma}{\blacklozenge} B} u - J_{L^* \circ B \circ L} u\|_{\mathcal{H}} \\ &\leqslant \left(\frac{\gamma^2}{4(1-\gamma)^2}(2\rho+\delta)^2 + \frac{\gamma}{4(1-\gamma)}\eta^2\right)^{1/2} + \frac{\gamma}{2(1-\gamma)}(2\rho+\delta) \\ &\to 0 \ \text{ as } \ 0 < \gamma \to 0, \end{aligned} \tag{3.89}$$

which completes the proof. ☐

Corollary 3.44 *Let $0 \neq p \in \mathbb{N}$ and, for every $k \in \{1, \ldots, p\}$, let $\mathcal{G}_k$ be a real Hilbert space, let*

$L_k \in \mathcal{B}(\mathcal{H}, \mathcal{G}_k)$, *let* $B_k \colon \mathcal{G}_k \to 2^{\mathcal{G}_k}$ *be maximally monotone, and let* $\alpha_k \in \left]0, +\infty\right[$. *Suppose that* $\sum_{k=1}^{p} \alpha_k \|L_k\|^2 \leqslant 1$ *and that* $\sum_{k=1}^{p} \alpha_k L_k^* \circ B_k \circ L_k$ *is maximally monotone. Then*

$$\overset{\blacklozenge}{\mathsf{M}}_\gamma(B_k, L_k)_{1 \leqslant k \leqslant p} \overset{g}{\to} \sum_{k=1}^{p} \alpha_k L_k^* \circ B_k \circ L_k \quad \textit{as} \quad 0 < \gamma \to 0. \tag{3.90}$$

Proof. Define L as in (3.32) and B as in (3.33), and recall that from Example 3.19 that, for every $\gamma \in \left]0, +\infty\right[$, $\overset{\blacklozenge}{\mathsf{M}}_\gamma(B_k, \alpha_k)_{1 \leqslant k \leqslant p} = L \overset{\gamma}{\blacklozenge} B$ is maximally monotone. Therefore, since $\sum_{k=1}^{p} \alpha_k L_k^* \circ B_k \circ L_k = L^* \circ B \circ L$, the conclusion follows from Proposition 3.43(i). ☐

Corollary 3.45 ([4, Proposition 1.4]) *Let* $A \colon \mathcal{H} \to 2^{\mathcal{H}}$ *be maximally monotone. Then*

$$(\forall \rho \in \left]0, +\infty\right[) \quad \lim_{\gamma \to 0} \operatorname{haus}_\rho({}^{\gamma}\!A, A) = 0. \tag{3.91}$$

Proof. Set $L = \mathrm{Id}_{\mathcal{H}}/2$ and $B = 2A(2\mathrm{Id}_{\mathcal{H}})$. Thus, $L^* \circ B \circ L = A$. Further, by Example 3.8, $(\forall \gamma \in \left]0, +\infty\right[)\ {}^{\gamma}\!A = L \overset{\gamma/3}{\blacklozenge} B$. Since $\mathrm{Id}_{\mathcal{H}} \circ L^*$ is invertible, we derive from Proposition 3.43(ii)(b) that

$$(\forall \rho \in \left]0, +\infty\right[) \quad \operatorname{haus}_\rho({}^{\gamma}\!A, A) = \operatorname{haus}_\rho(L \overset{\gamma/3}{\blacklozenge} B, L^* \circ B \circ L) \to 0 \ \text{ as } \ 0 < \gamma \to 0, \tag{3.92}$$

which establishes (3.91). ☐

Corollary 3.46 *Suppose that* $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ *satisfies* $0 < \|L\| \leqslant 1$ *and let* $g \in \Gamma_0(\mathcal{G})$. *Assume that* $0 \in \operatorname{sri}(\operatorname{dom} g - \operatorname{ran} L)$. *Then the following hold:*

(i) $\partial(L \overset{\gamma}{\blacklozenge} g) \overset{g}{\to} \partial(g \circ L)$ *as* $0 < \gamma \to 0$.

(ii) *Suppose that* $g \colon \mathcal{G} \to \mathbb{R}$ *is* β*-Lipschitzian for some* $\beta \in \left]0, +\infty\right[$. *Then*

$$(\forall \rho \in \left]0, +\infty\right[) \quad \lim_{\gamma \to 0} \operatorname{haus}_\rho\big(\partial(L \overset{\gamma}{\blacklozenge} g), \partial(g \circ L)\big) = 0. \tag{3.93}$$

Proof. Invoking [7, Corollaries 13.38 and 16.53(i)], $g^{**} = g$ and $\partial(g \circ L) = L^* \circ (\partial g) \circ L$.

(i): It follows from Example 3.27(ii) and Proposition 3.43(i) that

$$\partial(L \overset{\gamma}{\blacklozenge} g) = L \overset{\gamma}{\blacklozenge} \partial g \overset{g}{\to} L^* \circ (\partial g) \circ L = \partial(g \circ L) \ \text{ as } \ 0 < \gamma \to 0. \tag{3.94}$$

(ii): Appealing to [7, Corollary 17.19], $\operatorname{ran} \partial g \subset B(0; \beta)$. Thus, $\operatorname{ran} \partial g$ is bounded, and the conclusion follows from Proposition 3.43(ii)(a). ☐

Acknowledgement

This work forms part of the author's Ph.D. dissertation. The author gratefully acknowledges the guidance of his Ph.D. advisor P. L. Combettes, throughout this work. The author also thanks the anonymous referees for their comments and suggestions, which improved the quality of this manuscript.

Funding Information

This work was supported by the National Science Foundation under grant CCF-2211123.

References

[1] F. J. Aragón-Artacho, R. Campoy, and M. K. Tam, Strengthened splitting methods for computing resolvents, *Comput. Optim. Appl.*, vol. 80, pp. 549–585, 2021.

[2] H. Attouch, *Variational Convergence for Functions and Operators*. Pitman, Boston, MA, 1984.

[3] H. Attouch, R. Lucchetti, and R. Wets, The topology of the ρ-Hausdorff distance, *Ann. Mat. Pura Appl.*, vol. 160, pp. 303–320, 1991.

[4] H. Attouch, A. Moudafi, and H. Riahi, Quantitative stability analysis for maximal monotone operators and semi-groups of contractions, *Nonlinear Anal.*, vol. 21, pp. 697–723, 1993.

[5] D. Azé and J. P. Penot, Operations on convergent families of sets and functions, *Optimization*, vol. 21, pp. 521–534, 1990.

[6] S. Bartz, H. H. Bauschke, S. M. Moffat, and X. Wang, The resolvent average of monotone operators: Dominant and recessive properties, *SIAM J. Optim.*, vol. 26, pp. 602–634, 2016.

[7] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed. Springer, New York, 2017.

[8] M. N. Bùi and P. L. Combettes, Integral resolvent and proximal mixtures, *J. Optim. Theory Appl.*, vol. 203, pp. 2328–2353, 2024.

[9] R. R. Coifman and D. L. Donoho, Translation-invariant de-noising, *Lect. Notes Stat.*, vol. 103, pp 125–150, 1995.

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

[11] P. L. Combettes, Resolvent and proximal compositions, *Set-Valued Var. Anal.*, vol. 31, art. 22, 29 pp., 2023.

[12] P. L. Combettes, The geometry of monotone operator splitting methods, *Acta Numer.*, vol. 33, pp. 487–632, 2024.

[13] P. L. Combettes and D. J. Cornejo, Signal recovery with proximal comixtures, *Proc. Europ. Signal Process. Conf.*, pp. 2637–2641, Lyon, France, August 26–30, 2024.

[14] P. L. Combettes and D. J. Cornejo, Variational analysis of proximal compositions and integral proximal mixtures, *Evol. Equ. Control Theory*, vol. 17, pp. 106–139, 2026.

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

[16] P. L. Combettes and J.-C. Pesquet, Fixed point strategies in data science, *IEEE Trans. Signal Process.*, vol. 69, pp. 3878–3905, 2021.

[17] P. L. Combettes and Z. C. Woodstock, A variational inequality model for the construction of signals from inconsistent nonlinear equations, *SIAM J. Imaging Sci.*, vol. 15, pp. 84–109, 2022.

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

[19] R. Glowinski and P. L. Tallec (eds.), *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics.* SIAM, Philadelphia, PA, 1989.

[20] Y. Malitsky and M. K. Tam, Resolvent splitting for sums of monotone operators with minimal lifting, *Math. Program.*, vol. 201, pp. 231–262, 2023.

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

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

[23] X. Wang, Self-dual regularization of monotone operators via the resolvent average, *SIAM J. Optim.*, vol. 21, pp. 438–462, 2011.

Chapter 4

RESOLVENT COMPOSITIONS FOR POSITIVE LINEAR OPERATORS

4.1 Introduction and context

In this chapter, we focus on question **(Q3)** of Chapter 1 and present new results on resolvent compositions for positive linear operators. In particular, we show that the resolvent compositions are nonexpansive with respect to the Thompson metric and study nonlinear equations based on these operators.

This chapter presents the following journal article:

D. J. Cornejo, Resolvent compositions for positive linear operators, *Positivity*, to appear.

4.2 Article: Resolvent compositions for positive linear operators

Abstract. Resolvent compositions were recently introduced as monotonicity-preserving operations that combine a set-valued monotone operator and a bounded linear operator. They generalize in particular the notion of a resolvent average. We analyze the resolvent compositions when the monotone operator is a positive linear operator. We establish several new properties, including Löwner partial order relations, concavity, and asymptotic behavior. In addition, we show that the resolvent composition operations are nonexpansive with respect to the Thompson metric. We also introduce a new form of geometric interpolation and explore its connections to resolvent compositions. Finally, we study two nonlinear equations based on resolvent compositions.

4.2.1 Introduction

Throughout, $\mathcal{H}$ is a real Hilbert space with identity operator $\mathrm{Id}_{\mathcal{H}}$, scalar product $\langle \cdot \mid \cdot \rangle_{\mathcal{H}}$, and associated norm $\|\cdot\|_{\mathcal{H}}$. In addition, $\mathcal{G}$ is a real Hilbert space, the set of bounded linear operators from $\mathcal{H}$ to $\mathcal{G}$ is denoted by $\mathcal{B}(\mathcal{H},\mathcal{G})$, and $\mathcal{B}(\mathcal{H}) = \mathcal{B}(\mathcal{H},\mathcal{H})$. The adjoint of $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ is denoted by L^*. The set $\mathcal{P}(\mathcal{H})$ of positive operators on $\mathcal{H}$ is the collection of self-adjoint operators $A \in \mathcal{B}(\mathcal{H})$ such that $(\forall x \in \mathcal{H})\ \langle Ax \mid x\rangle_{\mathcal{H}} \geqslant 0$. The Löwner partial ordering between two self-adjoint operators A and B in $\mathcal{B}(\mathcal{H})$ is defined by $A \preccurlyeq B \Leftrightarrow B - A \in \mathcal{P}(\mathcal{H})$, and the set of self-adjoint strongly positive operators on $\mathcal{H}$ is

$$\mathcal{S}(\mathcal{H}) = \big\{A \in \mathcal{P}(\mathcal{H}) \mid (\exists\, \alpha \in \left]0,+\infty\right[)\ \ \alpha\mathrm{Id}_{\mathcal{H}} \preccurlyeq A\big\}. \tag{4.1}$$

A fundamental operator associated with a monotone operator $B\colon \mathcal{G} \to 2^{\mathcal{G}}$ is its resolvent

$$J_B = (\mathrm{Id}_{\mathcal{G}} + B)^{-1}, \tag{4.2}$$

which plays a central role in monotone operator theory and convex optimization, especially through its use in operator splitting algorithms [3, 7, 10]. In many applications, monotone operators arise in combination with linear operators, which motivates the study of operations that combine the monotone operator B and a linear operator $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ while preserving monotonicity. Recently, [6] introduced two monotonicity-preserving operations called the *resolvent composition* and the *resolvent cocomposition* of B and L, defined respectively by

$$L \overset{\gamma}{\diamond} B = L^* \triangleright \big(B + \gamma^{-1}\mathrm{Id}_{\mathcal{G}}\big) - \gamma^{-1}\mathrm{Id}_{\mathcal{H}} \tag{4.3}$$

and

$$L \overset{\gamma}{\blacklozenge} B = \Big(L \overset{1/\gamma}{\diamond} B^{-1}\Big)^{-1}, \tag{4.4}$$

where $L^*\triangleright B = (L^*\circ B^{-1}\circ L)^{-1}$ is the *parallel composition* of B by L^* [3], and $\gamma \in \left]0,+\infty\right[$. An attractive property of resolvent compositions is that their resolvent operators can be explicitly expressed through L and the resolvent of B [6, Propositions 1.2 and 4.1(v)], namely,

$$J_{\gamma(L\overset{\gamma}{\diamond}B)} = L^* \circ J_{\gamma B} \circ L \quad \text{and} \quad J_{\gamma(L\overset{\gamma}{\blacklozenge}B)} = \mathrm{Id}_{\mathcal{H}} - L^* \circ \big(\mathrm{Id}_{\mathcal{G}} - J_{\gamma B}\big) \circ L. \tag{4.5}$$

This feature, in turn, significantly facilitates the design and implementation of algorithms for monotone inclusion and convex optimization problems [5–8, 12]. Special cases can also be implicitly found in concrete applications such as image recovery [8], neural networks [14], inverse problems [15], and machine learning [29, 32]. For further motivation, let us consider the following examples.

Example 4.1 (resolvent mixtures) Let $0 \neq p \in \mathbb{N}$ and let $\gamma \in \left]0,+\infty\right[$. For every $k \in$

$\{1,\ldots,p\}$, let $\mathcal{G}_k$ be a real Hilbert space, let $L_k \in \mathcal{B}(\mathcal{H},\mathcal{G}_k)$ be such that $0 < \|L_k\| \leqslant 1$, let $B_k \in \mathcal{S}(\mathcal{G}_k)$, and let $\alpha_k \in \left]0,+\infty\right[$. Suppose that $\sum_{k=1}^p \alpha_k = 1$. Then the *resolvent mixture* and the *resolvent comixture* are defined, respectively, by

$$\overset{\diamond}{\mathsf{M}}_\gamma(L_k,B_k)_{1\leqslant k\leqslant p} = \left(\sum_{k=1}^p \alpha_k L_k^* \circ \left(B_k + \gamma^{-1}\mathrm{Id}_{\mathcal{G}_k}\right)^{-1} \circ L_k\right)^{-1} - \gamma^{-1}\mathrm{Id}_{\mathcal{H}} \tag{4.6}$$

and

$$\overset{\blacklozenge}{\mathsf{M}}_\gamma(L_k,B_k)_{1\leqslant k\leqslant p} = \left(\left(\sum_{k=1}^p \alpha_k L_k^* \circ \left(B_k^{-1} + \gamma^{-1}\mathrm{Id}_{\mathcal{G}_k}\right)^{-1} \circ L_k\right)^{-1} - \gamma^{-1}\mathrm{Id}_{\mathcal{H}}\right)^{-1} \tag{4.7}$$

Resolvent mixtures were introduced in [6, Example 3.4], and subsequently studied in [5, 12]. They are a particular case of resolvent compositions. Specifically, let $\mathcal{G} = \bigoplus_{k=1}^p \mathcal{G}_k$ and set

$$L\colon \mathcal{H} \to \mathcal{G}\colon x \mapsto \left(\sqrt{\alpha_k} L_k x\right)_{1\leqslant k\leqslant p} \quad \text{and} \quad B\colon \mathcal{G} \to \mathcal{G}\colon (y_k)_{1\leqslant k\leqslant p} \mapsto (B_k y_k)_{1\leqslant k\leqslant p}. \tag{4.8}$$

Then $L \overset{\gamma}{\diamond} B = \overset{\diamond}{\mathsf{M}}_\gamma(L_k,B_k)_{1\leqslant k\leqslant p}$ and $L \overset{\gamma}{\blacklozenge} B = \overset{\blacklozenge}{\mathsf{M}}_\gamma(L_k,B_k)_{1\leqslant k\leqslant p}$. Further, as shown in [12, Proposition 5.13(i)], $\overset{\blacklozenge}{\mathsf{M}}_\gamma(L_k,B_k)_{1\leqslant k\leqslant p}$ graph converges to $\sum_{k=1}^p \alpha_k L_k^* \circ B_k \circ L_k$ as $\gamma \to 0$.

Example 4.2 (arithmetic, harmonic, and resolvent average) In the context of Example 4.1, suppose that, for every $k \in \{1,\ldots,p\}$, $\mathcal{G}_k = \mathcal{H}$ and $L_k = \mathrm{Id}_{\mathcal{H}}$. Then the *arithmetic average* and the *harmonic average* are given, respectively, by

$$L^* \circ B \circ L = \sum_{k=1}^p \alpha_k B_k \quad \text{and} \quad L^* \triangleright B = \left(\sum_{k=1}^p \alpha_k B_k^{-1}\right)^{-1}. \tag{4.9}$$

An alternative averaging operation is the *resolvent average*, introduced in [4] and further studied in [2, 6, 31], given by

$$\mathsf{rav}_\gamma(B_k)_{1\leqslant k\leqslant p} = \left(\sum_{k=1}^p \alpha_k \left(B_k + \gamma^{-1}\mathrm{Id}_{\mathcal{H}}\right)^{-1}\right)^{-1} - \gamma^{-1}\mathrm{Id}_{\mathcal{H}}, \quad \text{where} \quad \gamma \in \left]0,+\infty\right[. \tag{4.10}$$

The resolvent average is as a special case of the resolvent mixtures [6, Example 1.3], to wit,

$$\overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathrm{Id}_{\mathcal{H}},B_k)_{1\leqslant k\leqslant p} = \overset{\diamond}{\mathsf{M}}_\gamma(\mathrm{Id}_{\mathcal{H}},B_k)_{1\leqslant k\leqslant p} = \mathsf{rav}_\gamma(B_k)_{1\leqslant k\leqslant p}. \tag{4.11}$$

It was established in [19, 22] that, when $\mathcal{S}(\mathcal{H})$ is endowed with the *Thompson metric* $d_T^{\mathcal{H}}$ (see (4.49)), the averages (4.9) and (4.10) are nonexpansive. This property is important as it ensures stability of the averaging processes and is particularly useful in the study of nonlinear equations [19]. Further, in the finite-dimensional setting, [4, Corollary 4.6] shows that the resolvent average is concave, and [4, Theorem 4.2] establishes, by means of a pointwise convergence proof, that the resolvent average interpolates between the arithmetic average ($0 < \gamma \to 0$) and the harmonic average ($\gamma \to +\infty$).

Example 4.3 (weighted $\mathscr{A}\#\mathscr{H}$–means) In the finite-dimensional setting of Example 4.2, a family $\big(\mathcal{L}_\gamma(B_k)_{1\leqslant k\leqslant p}\big)_{\gamma\in\mathbb{R}}$ of means interpolating between the arithmetic average and the harmonic average was introduced in [17] (see also [16]). These means, referred to as *weighted $\mathscr{A}\#\mathscr{H}$–means*, are closely related to resolvent averages [17, Proposition 3.5] through the ordering

$$\mathsf{rav}_\gamma(B_k)_{1\leqslant k\leqslant p} \preccurlyeq \mathcal{L}_{1/\gamma}(B_k)_{1\leqslant k\leqslant p} \preccurlyeq \sum_{k=1}^{p} \alpha_k B_k, \tag{4.12}$$

and themselves interpolate between the arithmetic average ($\gamma \to +\infty$) and the harmonic average ($\gamma \to -\infty$) [17, Proposition 3.4].

The aim of this paper is to investigate the operations (4.3) and (4.4) when $B \in \mathcal{S}(\mathcal{G})$. We establish several new properties, including Löwner partial order relations, concavity, nonexpansiveness, and asymptotic behavior. This specific setting leads to new results that, in particular, generalize the corresponding asymptotic properties in [12], as well as those of the proximal average established in [4, 17, 19].

The remainder of the paper is organized as follows. In Section 4.2.2, we provide our notation and necessary mathematical background. In Section 4.2.3, we present several new properties of $(L \overset{\gamma}{\blacklozenge} B)_{\gamma\in]0,+\infty[}$ and $(L \overset{\gamma}{\lozenge} B)_{\gamma\in]0,+\infty[}$. In particular, these operations are concave and

- $L \overset{\gamma}{\blacklozenge} B \preccurlyeq L^* \circ B \circ L$ and $L \overset{\gamma}{\blacklozenge} B \to L^* \circ B \circ L$ as $0 < \gamma \to 0$,
- $L^* \triangleright B \preccurlyeq L \overset{\gamma}{\lozenge} B$ and $L \overset{\gamma}{\lozenge} B \to L^* \triangleright B$ as $\gamma \to +\infty$.

In Section 4.2.4, we show that the resolvent compositions are nonexpansive with respect to the Thompson metric, in the sense that, for every $A \in \mathcal{S}(\mathcal{G})$ and $B \in \mathcal{S}(\mathcal{G})$,

$$d_T^{\mathcal{H}}\big(L \overset{\gamma}{\blacklozenge} A, L \overset{\gamma}{\blacklozenge} B\big) \leqslant d_T^{\mathcal{G}}(A,B) \quad \text{and} \quad d_T^{\mathcal{H}}\big(L \overset{\gamma}{\lozenge} A, L \overset{\gamma}{\lozenge} B\big) \leqslant d_T^{\mathcal{G}}(A,B). \tag{4.13}$$

Finally, in Section 4.2.5, we introduce the geometric interpolation $\mathcal{L}_\gamma(L,B)$ (see (4.64)) between $L^* \triangleright B$ and $L^* \circ B \circ L$ when L is an isometry, which generalizes the weighted $\mathscr{A}\#\mathscr{H}$–means. We establish the partial order relations

$$L^* \triangleright B \preccurlyeq \mathcal{L}_{-\gamma}(L,B) \preccurlyeq L \overset{\gamma}{\blacklozenge} B \preccurlyeq \mathcal{L}_{1/\gamma}(L,B) \preccurlyeq L^* \circ B \circ L, \tag{4.14}$$

and conclude by studying two nonlinear equations involving resolvent compositions.

4.2.2 Notation and background

The space $\mathcal{B}(\mathcal{H},\mathcal{G})$ is endowed with the topology induced by the operator norm

$$\big(\forall L\in\mathcal{B}(\mathcal{H},\mathcal{G})\big)\quad \|L\|=\sup_{\substack{x\in\mathcal{H}\\ \|x\|_{\mathcal{H}}\leqslant 1}}\|Lx\|_{\mathcal{G}}. \tag{4.15}$$

Let $L\in\mathcal{B}(\mathcal{H},\mathcal{G})$. Then L is an isometry if $L^*\circ L=\mathrm{Id}_{\mathcal{H}}$. Further, L is bounded below if there exists $\alpha\in\,]0,+\infty[$ such that $(\forall x\in\mathcal{H})\;\alpha\|x\|_{\mathcal{H}}\leqslant\|Lx\|_{\mathcal{G}}$. Equivalently, by [3, Fact 2.26], L is bounded below if and only if L is injective with closed range. In particular, when $\mathcal{H}$ and $\mathcal{G}$ are finite-dimensional, L is bounded below if and only if L is injective.

The quadratic kernel of $A\in\mathcal{P}(\mathcal{H})$ is $\mathscr{Q}_A\colon\mathcal{H}\to\mathbb{R}\colon x\mapsto(1/2)\langle x\mid Ax\rangle_{\mathcal{H}}$. The Legendre conjugate of $f\colon\mathcal{H}\to[-\infty,+\infty]$ is the function

$$f^*\colon\mathcal{H}\to[-\infty,+\infty]\colon x^*\mapsto\sup_{x\in\mathcal{H}}\big(\langle x\mid x^*\rangle_{\mathcal{H}}-f(x)\big), \tag{4.16}$$

and the Moreau envelope of $f\colon\mathcal{H}\to[-\infty,+\infty]$ of parameter $\gamma\in\,]0,+\infty[$ is

$$^{\gamma}f\colon\mathcal{H}\to[-\infty,+\infty]\colon x\mapsto\inf_{z\in\mathcal{H}}\Big(f(z)+\frac{1}{2\gamma}\|x-z\|_{\mathcal{H}}^2\Big). \tag{4.17}$$

The set of proper lower semicontinuous convex functions from $\mathcal{H}$ to $]-\infty,+\infty]$ is denoted by $\Gamma_0(\mathcal{H})$. Let $L\in\mathcal{B}(\mathcal{H},\mathcal{G})$ and $h\colon\mathcal{G}\to[-\infty,+\infty]$. The infimal postcomposition of h by L^* is

$$L^*\triangleright h\colon\mathcal{H}\to[-\infty,+\infty]\colon x\mapsto\inf_{\substack{y\in\mathcal{G}\\ L^*y=x}}h(y), \tag{4.18}$$

the proximal composition of h and L with parameter $\gamma\in\,]0,+\infty[$ (see [6, 9]) is

$$L\overset{\gamma}{\diamond}h=\Big({}^{\frac{1}{\gamma}}(h^*)\circ L\Big)^*-\frac{1}{2\gamma}\|\cdot\|_{\mathcal{H}}^2, \tag{4.19}$$

and the proximal cocomposition of h and L with parameter $\gamma\in\,]0,+\infty[$ is

$$L\overset{\gamma}{\blacklozenge}h=\big(L\overset{1/\gamma}{\diamond}h^*\big)^*. \tag{4.20}$$

The following facts will be used subsequently.

Lemma 4.4 *The following properties are satisfied:*

(i) *Let $A\in\mathcal{S}(\mathcal{G})$. Then $\mathscr{Q}_A^*=\mathscr{Q}_{A^{-1}}$.*

(ii) *Let $A\in\mathcal{S}(\mathcal{G})$ and $B\in\mathcal{S}(\mathcal{G})$. Then $A\preccurlyeq B\Leftrightarrow B^{-1}\preccurlyeq A^{-1}$.*

(iii) *Let $L\in\mathcal{B}(\mathcal{H},\mathcal{G})$, $A\in\mathcal{P}(\mathcal{G})$, and $B\in\mathcal{P}(\mathcal{G})$. Then $A\preccurlyeq B\Rightarrow L^*\circ A\circ L\preccurlyeq L^*\circ B\circ L$.*

(iv) *Let $A\in\mathcal{P}(\mathcal{G})$ and $B\in\mathcal{P}(\mathcal{G})$. Then $A\preccurlyeq B\Rightarrow\|A\|\leqslant\|B\|$.*

(v) *Let* $(A_n)_{n\in\mathbb{N}}$*,* $(B_n)_{n\in\mathbb{N}}$*,* A*, and* B *be self-adjoint operators in* $\mathcal{B}(\mathcal{G})$ *such that* $A_n \to A$*,* $B_n \to B$*, and* $(\forall n \in \mathbb{N})$ $A_n \preccurlyeq B_n$*. Then* $A \preccurlyeq B$*.*

Proof. (i)–(ii): See the proof of [3, Example 13.18(i)].

(iii): Let $x \in \mathcal{H}$. Since $A \preccurlyeq B$, $\langle x \,|\, L^*(A(Lx))\rangle_{\mathcal{H}} = \langle Lx \,|\, A(Lx)\rangle_{\mathcal{G}} \leqslant \langle Lx \,|\, B(Lx)\rangle_{\mathcal{G}} = \langle x \,|\, L^*(B(Lx))\rangle_{\mathcal{H}}$.

(iv): Since A and B are self-adjoint and $0 \preccurlyeq A \preccurlyeq B$, we deduce from [3, Fact 2.25(iii)] that

$$\|A\| = \sup_{\substack{x\in\mathcal{G}\\ \|x\|_{\mathcal{G}}\leqslant 1}} |\langle Ax \,|\, x\rangle_{\mathcal{G}}| = \sup_{\substack{x\in\mathcal{G}\\ \|x\|_{\mathcal{G}}\leqslant 1}} \langle Ax \,|\, x\rangle_{\mathcal{G}} \leqslant \sup_{\substack{x\in\mathcal{G}\\ \|x\|_{\mathcal{G}}\leqslant 1}} \langle Bx \,|\, x\rangle_{\mathcal{G}} = \sup_{\substack{x\in\mathcal{G}\\ \|x\|_{\mathcal{G}}\leqslant 1}} |\langle Bx \,|\, x\rangle_{\mathcal{G}}| = \|B\|. \tag{4.21}$$

(v): Since $A_n \to A$ and $B_n \to B$, convergence is in particular pointwise. Thus, for every $x \in \mathcal{G}$, $0 \leqslant \langle x \,|\, (B_n - A_n)x\rangle_{\mathcal{G}} \to \langle x \,|\, (B - A)x\rangle_{\mathcal{G}}$. Hence, $0 \preccurlyeq B - A$ or, equivalently, $A \preccurlyeq B$. ☐

Lemma 4.5 *Suppose that* $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ *satisfies* $0 < \|L\| \leqslant 1$*, let* $g \in \Gamma_0(\mathcal{G})$*, and let* $\gamma \in \,]0,+\infty[$*. Then the following hold:*

(i) $L \overset{\gamma}{\diamond} g = (L \overset{1/\gamma}{\blacklozenge} g^*)^*$.

(ii) $L \overset{\gamma}{\blacklozenge} g \leqslant \min\{ L \overset{\gamma}{\diamond} g \,,\, g \circ L \}$.

(iii) *Set* $\Phi = (1/2)\|\cdot\|^2_{\mathcal{G}} - (1/2)\|\cdot\|^2_{\mathcal{H}} \circ L^*$*. Then* $L \overset{\gamma}{\blacklozenge} g = (g^* + \gamma\Phi)^* \circ L$*.*

(iv) *Set* $\Phi = (1/2)\|\cdot\|^2_{\mathcal{G}} - (1/2)\|\cdot\|^2_{\mathcal{H}} \circ L^*$*. Then* $L \overset{\gamma}{\diamond} g = L^* \rhd (g + \Phi/\gamma)$*.*

Proof. Recall that $g = g^{**}$ [3, Corollary 13.38].

(i): [9, Proposition 3.7(iii)].

(ii): [9, Proposition 3.20(ii)–(iii)].

(iii)–(iv): [9, Proposition 3.2(i)–(ii)]. ☐

Lemma 4.6 *Let* $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ *and* $B \in \mathcal{P}(\mathcal{G})$*. Then the following hold:*

(i) $L^* \circ B \circ L \in \mathcal{P}(\mathcal{H})$.

(ii) $\mathcal{Q}_B \circ L = \mathcal{Q}_{L^*\circ B\circ L}$.

(iii) *Suppose that* $B \in \mathcal{S}(\mathcal{G})$ *and that* L *is bounded below. Then* $L^* \circ B \circ L \in \mathcal{S}(\mathcal{H})$ *and* $L^* \rhd B \in \mathcal{S}(\mathcal{H})$*.*

Proof. (i): Take $A = 0$ in Lemma 4.4(iii).

(ii): For every $x \in \mathcal{H}$, $\mathcal{Q}_B(Lx) = (1/2)\langle Lx \,|\, B(Lx)\rangle_{\mathcal{G}} = (1/2)\langle x \,|\, L^*(B(Lx))\rangle_{\mathcal{H}} = \mathcal{Q}_{L^*\circ B\circ L}(x)$.

(iii): Since $B \in \mathcal{S}(\mathcal{G})$, there exists $\alpha \in \,]0,+\infty[$ such that $\alpha\mathrm{Id}_{\mathcal{G}} \preccurlyeq B$. On the other hand, since L is bounded below, there exists $\beta \in \,]0,+\infty[$ such that $\beta^2\mathrm{Id}_{\mathcal{H}} \preccurlyeq L^* \circ L$. Therefore, Lemma 4.4(iii) yields

$$(\alpha\beta^2)\mathrm{Id}_{\mathcal{H}} \preccurlyeq \alpha(L^* \circ L) = L^* \circ (\alpha \mathrm{Id}_{\mathcal{G}}) \circ L \preccurlyeq L^* \circ B \circ L, \tag{4.22}$$

i.e., $L^* \circ B \circ L \in \mathcal{S}(\mathcal{H})$. Similarly, $L^* \circ B^{-1} \circ L \in \mathcal{S}(\mathcal{H})$, which implies that $L^* \triangleright B = (L^* \circ B^{-1} \circ L)^{-1} \in \mathcal{S}(\mathcal{H})$. ☐

Lemma 4.7 ([12, Proposition 3.3(ii)]) *Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$, let $B\colon \mathcal{G} \to 2^{\mathcal{G}}$, let $\gamma \in \left]0, +\infty\right[$, and set $\Psi = \mathrm{Id}_{\mathcal{G}} - L \circ L^*$. Then $L \overset{\gamma}{\blacklozenge} B = L^* \circ (B^{-1} + \gamma\Psi)^{-1} \circ L$.*

Lemma 4.8 ([12, Proposition 3.4(i)]) *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ is an isometry, let $B\colon \mathcal{G} \to 2^{\mathcal{G}}$, and let $\gamma \in \left]0, +\infty\right[$. Then $L \overset{\gamma}{\diamond} B = L \overset{\gamma}{\blacklozenge} B$.*

4.2.3 Resolvent compositions

In this section, we study the resolvent cocomposition operators when $B \in \mathcal{S}(\mathcal{G})$. We strengthen several results obtained in [12], as well as those established specifically for the resolvent average in [4]. The results obtained include comparisons among the different composite operations, as well as an analysis of the asymptotic behavior of $(L \overset{\gamma}{\blacklozenge} B)_{\gamma \in]0,+\infty[}$ and $(L \overset{\gamma}{\diamond} B)_{\gamma \in]0,+\infty[}$, as the parameter γ varies.

Proposition 4.9 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ satisfies $0 < \|L\| \leqslant 1$, let $B \in \mathcal{S}(\mathcal{G})$, and let $\gamma \in \left]0, +\infty\right[$. Then the following hold:*

(i) $L \overset{\gamma}{\blacklozenge} B \in \mathcal{P}(\mathcal{H})$.

(ii) $L \overset{\gamma}{\blacklozenge} \mathscr{Q}_B = \mathscr{Q}_{L \overset{\gamma}{\blacklozenge} B}$.

(iii) *Let $\lambda \in \left]0, 1\right[$. Then $T_\gamma\colon \mathcal{S}(\mathcal{G}) \to \mathcal{P}(\mathcal{H})\colon A \mapsto L \overset{\gamma}{\blacklozenge} A$ is concave in the sense that*

$$\big(\forall A \in \mathcal{S}(\mathcal{G})\big) \quad \lambda\big(L \overset{\gamma}{\blacklozenge} A\big) + (1-\lambda)\big(L \overset{\gamma}{\blacklozenge} B\big) \preccurlyeq L \overset{\gamma}{\blacklozenge} \big(\lambda A + (1-\lambda)B\big). \tag{4.23}$$

(iv) *Suppose that L is bounded below. Then the following are satisfied:*

  (a) $L \overset{\gamma}{\blacklozenge} B \in \mathcal{S}(\mathcal{H})$ *and* $L \overset{\gamma}{\diamond} B \in \mathcal{S}(\mathcal{H})$.

  (b) $L \overset{\gamma}{\diamond} \mathscr{Q}_B = \mathscr{Q}_{L \overset{\gamma}{\diamond} B}$.

  (c) *Let $\lambda \in \left]0, 1\right[$. Then $R_\gamma\colon \mathcal{S}(\mathcal{G}) \to \mathcal{S}(\mathcal{H})\colon A \mapsto L \overset{\gamma}{\diamond} A$ is concave in the sense that*

$$\big(\forall A \in \mathcal{S}(\mathcal{G})\big) \quad \lambda\big(L \overset{\gamma}{\diamond} A\big) + (1-\lambda)\big(L \overset{\gamma}{\diamond} B\big) \preccurlyeq L \overset{\gamma}{\diamond} \big(\lambda A + (1-\lambda)B\big). \tag{4.24}$$

Proof. Set $\Psi = \mathrm{Id}_{\mathcal{G}} - L \circ L^*$. Since $\|L\| \leqslant 1$, $\Psi \in \mathcal{P}(\mathcal{G})$, which yields $B^{-1} + \gamma\Psi \in \mathcal{S}(\mathcal{G})$. On the other hand, recall from Lemma 4.7 that

$$L \overset{\gamma}{\blacklozenge} B = L^* \circ \big(B^{-1} + \gamma\Psi\big)^{-1} \circ L. \tag{4.25}$$

(i): This follows from (4.25) and Lemma 4.6(i).

(ii): Set $\Phi = (1/2)\|\cdot\|_{\mathcal{G}}^2 - (1/2)\|\cdot\|_{\mathcal{H}}^2 \circ L^*$ and note that $\Phi = \mathscr{Q}_\Psi$. It follows from Lemma 4.5(iii), Lemma 4.4(i), Lemma 4.6(ii), and (4.25) that

$$\begin{aligned} L \overset{\gamma}{\blacklozenge} \mathscr{Q}_B &= \big(\mathscr{Q}_B^* + \gamma\Phi\big)^* \circ L \\ &= \big(\mathscr{Q}_{B^{-1}} + \gamma\mathscr{Q}_\Psi\big)^* \circ L \\ &= \mathscr{Q}_{B^{-1}+\gamma\Psi}^* \circ L \\ &= \mathscr{Q}_{L^*\circ\left(B^{-1}+\gamma\Psi\right)^{-1}\circ L} \\ &= \mathscr{Q}_{L\overset{\gamma}{\blacklozenge}B}. \end{aligned} \tag{4.26}$$

(iii): By (i), T_γ is well defined. Further, for every $A \in \mathcal{S}(\mathcal{G})$,

$$\begin{aligned} &\lambda\big(L \overset{\gamma}{\blacklozenge} A\big) + (1-\lambda)\big(L \overset{\gamma}{\blacklozenge} B\big) \preccurlyeq L \overset{\gamma}{\blacklozenge} \big(\lambda A + (1-\lambda)B\big) \\ &\Leftrightarrow (\forall x \in \mathcal{H})\ \lambda\Big\langle \big(L \overset{\gamma}{\blacklozenge} A\big)x \,\Big|\, x\Big\rangle_{\mathcal{H}} + (1-\lambda)\Big\langle \big(L \overset{\gamma}{\blacklozenge} B\big)x \,\Big|\, x\Big\rangle_{\mathcal{H}} \leqslant \Big\langle \Big(L \overset{\gamma}{\blacklozenge} \big(\lambda A + (1-\lambda)B\big)\Big)x \,\Big|\, x\Big\rangle_{\mathcal{H}} \\ &\Leftrightarrow (\forall x \in \mathcal{H})\ \lambda\mathscr{Q}_{L\overset{\gamma}{\blacklozenge}A}(x) + (1-\lambda)\mathscr{Q}_{L\overset{\gamma}{\blacklozenge}B}(x) \leqslant \mathscr{Q}_{\lambda(L\overset{\gamma}{\blacklozenge}A)+(1-\lambda)(L\overset{\gamma}{\blacklozenge}B)}(x). \end{aligned} \tag{4.27}$$

Therefore, it is enough to prove that, for every $x \in \mathcal{H}$, the function $\mathcal{S}(\mathcal{G}) \to \mathbb{R}\colon A \mapsto \mathscr{Q}_{L\overset{\gamma}{\blacklozenge}A}(x)$ is concave. Set $\Phi = (1/2)\|\cdot\|_{\mathcal{G}}^2 - (1/2)\|\cdot\|_{\mathcal{H}}^2 \circ L^*$. Because $\operatorname{dom}\Phi = \mathcal{G}$, the identity $(\gamma\Phi)^* = \Phi^*/\gamma$ and [3, Proposition 15.2] imply that

$$(\forall A \in \mathcal{S}(\mathcal{G}))\quad \big(\mathscr{Q}_A^* + \gamma\Phi\big)^* = \mathscr{Q}_A \,\square\, \big(\Phi^*/\gamma\big)\colon \mathcal{G} \to \left]-\infty,+\infty\right]\colon z \mapsto \inf_{y\in\mathcal{G}}\Big(\mathscr{Q}_A(y) + \frac{1}{\gamma}\Phi^*(z-y)\Big). \tag{4.28}$$

Thus, by virtue of (ii), Lemma 4.5(iii), and (4.28),

$$\begin{aligned} \big(\forall A \in \mathcal{S}(\mathcal{G})\big)(\forall x \in \mathcal{H})\quad \mathscr{Q}_{L\overset{\gamma}{\blacklozenge}A}(x) &= \big(L \overset{\gamma}{\blacklozenge} \mathscr{Q}_A\big)(x) \\ &= \big(\mathscr{Q}_A^* + \gamma\Phi\big)^*(Lx) \\ &= \inf_{y\in\mathcal{G}} \underbrace{\Big(\mathscr{Q}_A(y) + \frac{1}{\gamma}\Phi^*(Lx-y)\Big)}_{\text{affine in } A}. \end{aligned} \tag{4.29}$$

Hence, for every $x \in \mathcal{H}$, the function $\mathcal{S}(\mathcal{G}) \to \mathbb{R}\colon A \mapsto \mathscr{Q}_{L\overset{\gamma}{\blacklozenge}A}(x)$ is concave, as it can be expressed as the infimum of affine functions.

(iv)(a): It follows from (4.25) and Lemma 4.6(iii) that $L \overset{\gamma}{\blacklozenge} B \in \mathcal{S}(\mathcal{H})$. On the other hand, by (4.4) and applying the previous reasoning to B^{-1}, we obtain $L \overset{\gamma}{\lozenge} B = (L \overset{1/\gamma}{\blacklozenge} B^{-1})^{-1} \in \mathcal{S}(\mathcal{H})$.

(iv)(b): By Lemma 4.5(i), Lemma 4.4(i), (ii), and (4.4),

$$L \overset{\gamma}{\lozenge} \mathscr{Q}_B = \big(L \overset{1/\gamma}{\blacklozenge} \mathscr{Q}_B^*\big)^* = \big(L \overset{1/\gamma}{\blacklozenge} \mathscr{Q}_{B^{-1}}\big)^* = \mathscr{Q}^*_{L\overset{1/\gamma}{\blacklozenge}B^{-1}} = \mathscr{Q}_{\left(L\overset{1/\gamma}{\blacklozenge}B^{-1}\right)^{-1}} = \mathscr{Q}_{L\overset{\gamma}{\lozenge}B}. \tag{4.30}$$

(iv)(c): It follows from Lemma 4.5(iv) and (iv)(b) that

$$\big(\forall A\in\mathcal{S}(\mathcal{G})\big)\big(\forall x\in\mathcal{G}\big)\quad \mathscr{Q}_{L\overset{\gamma}{\diamond}A}(x)=\big(L\overset{\gamma}{\diamond}\mathscr{Q}_A\big)(x)=\inf_{\substack{y\in\mathcal{G}\\ L^*y=x}}\underbrace{\Big(\mathscr{Q}_A(y)+\frac{1}{\gamma}\Phi(y)\Big)}_{\text{affine in }A} \tag{4.31}$$

Thus, for every $x\in\mathcal{H}$, the function $\mathcal{S}(\mathcal{G})\to\mathbb{R}\colon A\mapsto\mathscr{Q}_{L\overset{\gamma}{\diamond}A}(x)$ is concave. As a consequence, as in the proof of (iii), R_γ is concave. ☐

The following example shows that, in the finite-dimensional setting, the resolvent composition admits a variational characterization. In particular, this holds for the resolvent average, as established in [4, Proposition 2.8].

Example 4.10 (variational characterization) Suppose that $\mathcal{H}$ and $\mathcal{G}$ are finite-dimensional and that $L\in\mathcal{B}(\mathcal{H},\mathcal{G})$ is injective and satisfies $\|L\|\leqslant 1$, let $B\in\mathcal{S}(\mathcal{G})$, and let $\gamma\in\,]0,+\infty[$. Define

$$f\colon\mathcal{S}(\mathcal{H})\to\mathbb{R}\colon X\mapsto -\ln\det\big(X+\gamma^{-1}\mathrm{Id}_{\mathcal{H}}\big) \tag{4.32}$$

and

$$F\colon\mathcal{S}(\mathcal{H})\to\mathbb{R}\colon X\mapsto f(X)+\big\langle L^*\circ(B+\gamma^{-1}\mathrm{Id}_{\mathcal{G}})^{-1}\circ L\,\big|\,X\big\rangle\,, \tag{4.33}$$

where $\det(X)$ denotes the determinant of X and $\langle X\mid B\rangle$ denotes the trace of $X\circ B$. Then $L\overset{\gamma}{\diamond}B$ is the unique minimizer of F. *Proof.* Let $\mathbb{S}(\mathcal{H})$ denote the set of self-adjoint operators on $\mathcal{H}$. Since $\mathbb{S}(\mathcal{H})\to\mathbb{S}(\mathcal{H})\colon X\mapsto X+\gamma^{-1}\mathrm{Id}_{\mathcal{H}}$ is affine, [3, Example 24.66 and Proposition 8.20] show that f is convex, differentiable, and that $(\forall X\in\mathcal{S}(\mathcal{H}))\;\nabla f(X)=-(X+\gamma^{-1}\mathrm{Id}_{\mathcal{H}})^{-1}$. Thus, F is also convex and differentiable, being the sum of f and an affine function. Therefore, by virtue of [3, Theorem 16.3 and Proposition 17.31(i)], it suffices to find the critical points of F, that is, to solve $\nabla F(X)=0$. Altogether, Proposition 4.9(iv)(a) ensures that $L\overset{\gamma}{\diamond}B\in\mathcal{S}(\mathcal{H})$, and

$$\begin{aligned}\nabla F(X)=0 &\Leftrightarrow -\big(X+\gamma^{-1}\mathrm{Id}_{\mathcal{H}}\big)^{-1}+L^*\circ\big(B+\gamma^{-1}\mathrm{Id}_{\mathcal{G}}\big)^{-1}\circ L=0\\ &\Leftrightarrow X+\gamma^{-1}\mathrm{Id}_{\mathcal{H}}=L^*\triangleright(B+\gamma^{-1}\mathrm{Id}_{\mathcal{G}})\\ &\Leftrightarrow X=L\overset{\gamma}{\diamond}B,\end{aligned} \tag{4.34}$$

which completes the proof. ☐

We now focus on Löwner partial ordering relations for resolvent compositions. These ordering relations will assist us in studying the convergence properties of resolvent compositions $L\overset{\gamma}{\diamond}B$ and $L\overset{\gamma}{\blacklozenge}B$, as well as of the new interpolation $\mathcal{L}_\gamma(L,B)$ introduced in Section 4.2.5, as γ varies.

Proposition 4.11 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ satisfies $0 < \|L\| \leqslant 1$, let $B \in \mathcal{S}(\mathcal{G})$, and let $\gamma \in \left]0, +\infty\right[$. Then the following hold:*

(i) *Set $\theta = 1/(1 + \gamma\|B\|)$. Then $\theta(L^* \circ B \circ L) \preccurlyeq L \overset{\gamma}{\blacklozenge} B \preccurlyeq L^* \circ B \circ L$.*

(ii) *Suppose that $A \in \mathcal{S}(\mathcal{G})$ satisfies $A \preccurlyeq B$. Then $L \overset{\gamma}{\blacklozenge} A \preccurlyeq L \overset{\gamma}{\blacklozenge} B$.*

(iii) *Let $\rho \in \left]0, +\infty\right[$ be such that $\rho \leqslant \gamma$. Then $L \overset{\gamma}{\blacklozenge} B \preccurlyeq L \overset{\rho}{\blacklozenge} B$.*

Proof. Set $\Psi = \mathrm{Id}_{\mathcal{G}} - L \circ L^*$ and recall that $L \overset{\gamma}{\blacklozenge} B = L^* \circ (B^{-1} + \gamma\Psi)^{-1} \circ L$ by Lemma 4.7.

(i): Note that $B \preccurlyeq \|B\|\ \mathrm{Id}_{\mathcal{G}}$ and that Lemma 4.4(ii) implies that $\mathrm{Id}_{\mathcal{G}} \preccurlyeq \|B\|\ B^{-1}$. Since $0 \preccurlyeq \Psi \preccurlyeq \mathrm{Id}_{\mathcal{G}}$,

$$B^{-1} \preccurlyeq B^{-1} + \gamma\Psi \preccurlyeq B^{-1} + \gamma\mathrm{Id}_{\mathcal{G}} \preccurlyeq (1 + \gamma\|B\|)\ B^{-1}, \tag{4.35}$$

and, by virtue of Lemma 4.4(ii),

$$\theta B \preccurlyeq \left(B^{-1} + \gamma\Psi\right)^{-1} \preccurlyeq B. \tag{4.36}$$

Hence, we deduce from (4.36) and Lemma 4.4(iii) that

$$\theta(L^* \circ B \circ L) \preccurlyeq L \overset{\gamma}{\blacklozenge} B \preccurlyeq L^* \circ B \circ L. \tag{4.37}$$

(ii): Since $\Psi \in \mathcal{P}(\mathcal{G})$, $A^{-1} + \gamma\Psi$ and $B^{-1} + \gamma\Psi$ are in $\mathcal{S}(\mathcal{G})$. Further, by Lemma 4.4(ii) and the fact that $A \preccurlyeq B$, $B^{-1} + \gamma\Psi \preccurlyeq A^{-1} + \gamma\Psi$. Thus, $(A^{-1} + \gamma\Psi)^{-1} \preccurlyeq (B^{-1} + \gamma\Psi)^{-1}$. Altogether, we deduce from Lemma 4.4(iii) that

$$L \overset{\gamma}{\blacklozenge} A = L^* \circ \left(A^{-1} + \gamma\Psi\right)^{-1} \circ L \preccurlyeq L^* \circ \left(B^{-1} + \gamma\Psi\right)^{-1} \circ L = L \overset{\gamma}{\blacklozenge} B. \tag{4.38}$$

(iii): Note that $B^{-1} + \gamma\Psi$ and $B^{-1} + \rho\Psi$ are in $\mathcal{S}(\mathcal{G})$ and that $B^{-1} + \rho\Psi \preccurlyeq B^{-1} + \gamma\Psi$. Therefore, Lemma 4.4(ii)-(iii) yields

$$L \overset{\gamma}{\blacklozenge} B = L^* \circ \left(B^{-1} + \gamma\Psi\right)^{-1} \circ L \preccurlyeq L^* \circ \left(B^{-1} + \rho\Psi\right)^{-1} \circ L = L \overset{\rho}{\blacklozenge} B, \tag{4.39}$$

as claimed. ☐

Corollary 4.12 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ is bounded below and satisfies $\|L\| \leqslant 1$, let $B \in \mathcal{S}(\mathcal{G})$, and let $\gamma \in \left]0, +\infty\right[$. Then the following hold:*

(i) *Set $\omega = 1 + \|B^{-1}\|/\gamma$. Then $L^* \triangleright B \preccurlyeq L \overset{\gamma}{\lozenge} B \preccurlyeq \omega(L^* \triangleright B)$.*

(ii) $L \overset{\gamma}{\blacklozenge} B \preccurlyeq L \overset{\gamma}{\lozenge} B$.

(iii) *Suppose that $A \in \mathcal{S}(\mathcal{G})$ satisfies $A \preccurlyeq B$. Then $L \overset{\gamma}{\lozenge} A \preccurlyeq L \overset{\gamma}{\lozenge} B$.*

(iv) *Let $\rho \in \left]0, +\infty\right[$ be such that $\rho \leqslant \gamma$. Then $L \overset{\gamma}{\lozenge} B \preccurlyeq L \overset{\rho}{\lozenge} B$.*

Proof. By Proposition 4.9(iv)(a), $L \overset{\gamma}{\lozenge} B \in \mathcal{S}(\mathcal{H})$. Further, recall that (4.4) yields $L \overset{\gamma}{\lozenge} B = (L \overset{1/\gamma}{\blacklozenge} B^{-1})^{-1}$.

[(i)](#): This follows from Lemma 4.4(ii) and Proposition 4.11(i) applied to B^{-1} and $1/\gamma$.

(ii): By Proposition 4.9(ii), Lemma 4.5(ii), and Proposition 4.9(iv)(b),

$$\mathscr{Q}_{L \overset{\gamma}{\blacklozenge} B} = L \overset{\gamma}{\blacklozenge} \mathscr{Q}_B \leqslant L \overset{\gamma}{\lozenge} \mathscr{Q}_B = \mathscr{Q}_{L \overset{\gamma}{\lozenge} B}. \tag{4.40}$$

Therefore, $L \overset{\gamma}{\blacklozenge} B \preccurlyeq L \overset{\gamma}{\lozenge} B$.

(iii): This follows from Lemma 4.4(ii) and Proposition 4.11(ii) applied to B^{-1} and $1/\gamma$.

(iv): This follows from Lemma 4.4(ii) and Proposition 4.11(iii) applied to B^{-1} and $1/\gamma$. $\square$

Corollary 4.13 *Suppose that $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ is bounded below and satisfies $\|L\| \leqslant 1$, let $B \in \mathcal{S}(\mathcal{G})$, and set $\kappa = \|B\|\,\|B^{-1}\|$ and $\rho = (1+\sqrt{\kappa})^2$. Then $L^* \circ B \circ L \preccurlyeq \rho(L^* \rhd B)$.*

Proof. Set $f\colon\,]0,+\infty[\, \to\,]0,+\infty[\colon \gamma \to (1+\gamma\|B\|)(1+\|B^{-1}\|/\gamma)$. By Proposition 4.11(i), Corollary 4.12(ii), and Corollary 4.12(i),

$$(\forall \gamma \in\,]0,+\infty[) \quad L^* \circ B \circ L \preccurlyeq f(\gamma)\big(L^* \rhd B\big). \tag{4.41}$$

Since $\rho = \min_{\gamma \in]0,+\infty[} f(\gamma)$, the assertion follows from (4.41). $\square$

We now present the main result of this section. In contrast to [12, Propositions 5.8 and 5.12(i)], which establish graph convergence of resolvent compositions, the following theorem provides asymptotic behavior of resolvent compositions in operator norm, which is stronger than graph convergence and therefore offers additional stability properties.

Theorem 4.14 *Suppose that $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ satisfies $0 < \|L\| \leqslant 1$, and let $B \in \mathcal{S}(\mathcal{G})$. Then the following hold:*

(i) *$L \overset{\gamma}{\blacklozenge} B \to L^* \circ B \circ L$ as $0 < \gamma \to 0$.*

(ii) *Suppose that L is bounded below. Then $L \overset{\gamma}{\lozenge} B \to L^* \rhd B$ as $\gamma \to +\infty$.*

Proof. (i): Set $(\forall \gamma \in\,]0,+\infty[)\ \theta_\gamma = 1/(1+\gamma\|B\|)$ and $D_\gamma = (L^* \circ B \circ L) - (L \overset{\gamma}{\blacklozenge} B)$. By Proposition 4.11(i),

$$0 \preccurlyeq D_\gamma \preccurlyeq \left(\frac{1-\theta_\gamma}{\theta_\gamma}\right)(L^* \circ B \circ L). \tag{4.42}$$

In addition, note that $\theta_\gamma \to 1$ as $0 < \gamma \to 0$. Therefore, it follows from (4.42) and Lemma 4.4(iv) that

$$\|D_\gamma\| \leqslant \left(\frac{1-\theta_\gamma}{\theta_\gamma}\right)\|L^* \circ B \circ L\| \to 0 \ \text{ as } \ 0 < \gamma \to 0. \tag{4.43}$$

(ii): Set $(\forall \gamma \in\,]0,+\infty[)\ \omega_\gamma = 1+\|B^{-1}\|/\gamma$ and $D_\gamma = (L \overset{\gamma}{\lozenge} B) - (L^* \rhd B)$. By Corollary 4.12(i),

$$0 \preccurlyeq D_\gamma \preccurlyeq (\omega_\gamma - 1)\,(L^* \rhd B). \tag{4.44}$$

Also, note that $\omega_\gamma \to 1$ as $\gamma \to +\infty$. Therefore, we combine (4.44) and Lemma 4.4(iv) to obtain

$$\|D_\gamma\| \leqslant (\omega_\gamma - 1)\,\|L^* \rhd B\| \to 0 \ \text{ as } \ 0 < \gamma \to +\infty, \tag{4.45}$$

which completes the proof. ☐

Corollary 4.15 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ is bounded below and satisfies $\|L\| \leqslant 1$. Then the operator $R\colon \mathcal{S}(\mathcal{G}) \to \mathcal{S}(\mathcal{H})\colon A \mapsto L^* \rhd A$ is concave in the sense that*

$$\big(\forall \lambda \in \,]0,1[\big)\big(\forall A \in \mathcal{S}(\mathcal{G})\big)\big(\forall B \in \mathcal{S}(\mathcal{G})\big) \quad \lambda(L^* \rhd A) + (1-\lambda)(L^* \rhd B) \preccurlyeq L^* \rhd \big(\lambda A + (1-\lambda)B\big). \tag{4.46}$$

Proof. By Proposition 4.9(iv)(c), $R_\gamma\colon \mathcal{S}(\mathcal{G}) \to \mathcal{S}(\mathcal{H})\colon A \mapsto L \overset{\gamma}{\diamond} A$ is concave, i.e.,

$$\big(\forall \lambda \in \,]0,1[\big)\big(\forall A \in \mathcal{S}(\mathcal{G})\big)\big(\forall B \in \mathcal{S}(\mathcal{G})\big) \quad \lambda(L \overset{\gamma}{\diamond} A) + (1-\lambda)(L \overset{\gamma}{\diamond} B) \preccurlyeq L^* \overset{\gamma}{\diamond} \big(\lambda A + (1-\lambda)B\big). \tag{4.47}$$

Hence, letting $\gamma \to +\infty$ in (4.47) and invoking Theorem 4.14(ii) together with Lemma 4.4(v), we obtain (4.46). ☐

Remark 4.16 In the context of the resolvent averages of Example 4.2, Theorem 4.14 and Corollary 4.15 generalize [4, Theorem 4.2 and Corollary 4.6], which were established in the finite-dimensional context using different techniques.

Corollary 4.17 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ is an isometry, and let $B \in \mathcal{S}(\mathcal{G})$. Then the following hold:*

(i) $(\forall \gamma \in \,]0,+\infty[)\ L^* \rhd B \preccurlyeq L \overset{\gamma}{\blacklozenge} B \preccurlyeq L^* \circ B \circ L$.

(ii) $L \overset{\gamma}{\blacklozenge} B \to L^* \circ B \circ L$ *as* $0 < \gamma \to 0$.

(iii) $L \overset{\gamma}{\blacklozenge} B \to L^* \rhd B$ *as* $\gamma \to +\infty$.

Proof. Since L is an isometry, Lemma 4.8 yields $L \overset{\gamma}{\diamond} B = L \overset{\gamma}{\blacklozenge} B$.

(i): This follows from Proposition 4.11(i) and Corollary 4.12(i).

(ii): This follows from Theorem 4.14(i).

(iii): This follows from Theorem 4.14(ii). ☐

Corollary 4.18 (resolvent mixtures) *Consider the setting of Example 4.1. Then the following hold:*

(i) $\overset{\blacklozenge}{\mathsf{M}}_\gamma(L_k, B_k)_{1 \leqslant k \leqslant p} \preccurlyeq \sum_{k=1}^p \alpha_k L_k^* \circ B_k \circ L_k$.

(ii) $\overset{\blacklozenge}{\mathsf{M}}_\gamma(L_k, B_k)_{1 \leqslant k \leqslant p} \to \sum_{k=1}^p \alpha_k L_k^* \circ B_k \circ L_k$ *as* $0 < \gamma \to 0$.

(iii) *Suppose that L_j is bounded below for some $j \in \{1, \ldots, p\}$. Then the following are satisfied:*

(a) $\overset{\diamond}{\mathsf{M}}_\gamma(L_k, B_k)_{1 \leqslant k \leqslant p} \in \mathcal{S}(\mathcal{H})$ *and* $\overset{\blacklozenge}{\mathsf{M}}_\gamma(L_k, B_k)_{1 \leqslant k \leqslant p} \in \mathcal{S}(\mathcal{H})$.

(b) $\left(\sum_{k=1}^{p} \alpha_k L_k^* \circ B_k^{-1} \circ L_k\right)^{-1} \preccurlyeq \overset{\diamond}{\mathsf{M}}_\gamma(L_k, B_k)_{1\leqslant k\leqslant p}$.

(c) $\overset{\diamond}{\mathsf{M}}_\gamma(L_k, B_k)_{1\leqslant k\leqslant p} \to \left(\sum_{k=1}^{p} \alpha_k L_k^* \circ B_k^{-1} \circ L_k\right)^{-1}$ *as* $\gamma \to +\infty$.

Proof. Note that $L^* \circ B \circ L = \sum_{k=1}^{p} \alpha_l L_k^* \circ B_k \circ L_k$ and $L^* \triangleright B = \left(\sum_{k=1}^{p} \alpha_k L_k^* \circ B_k^{-1} \circ L_k\right)^{-1}$. Further, if L_j is bounded below for some $j \in \{1, \ldots, p\}$, then L is also bounded below. Indeed, there exists $\alpha \in \,]0, +\infty[$ such that $(\forall x \in \mathcal{H})\ \alpha\|x\|_{\mathcal{H}} \leqslant \|L_j x\|_{\mathcal{G}_j}$. Thus, L is bounded below since

$$(\forall x \in \mathcal{H}) \quad \|Lx\|_{\mathcal{G}} = \left(\sum_{k=1}^{p} \alpha_k \|L_k x\|_{\mathcal{G}_k}^2\right)^{1/2} \geqslant \left(\alpha_j \alpha^2 \|x\|_{\mathcal{H}}^2\right)^{1/2} = \left(\alpha_j^{1/2}\alpha\right)\|x\|_{\mathcal{H}}. \tag{4.48}$$

(i): This follows from Proposition 4.11(i).

(ii): This follows from Theorem 4.14(i).

(iii)(a): This follows from Proposition 4.9(iv)(a).

(iii)(b): This follows from Corollary 4.12(i).

(iii)(c): This follows from Theorem 4.14(ii). ☐

We conclude this section by deriving a result that recovers and extends [4, Theorem 4.2], which was proved in the finite-dimensional setting.

Corollary 4.19 *Consider the setting of Example 4.2. Then the following hold:*

(i) $\left(\sum_{k=1}^{p} \alpha_k B_k^{-1}\right)^{-1} \preccurlyeq \mathsf{rav}_\gamma(B_k)_{1\leqslant k\leqslant p} \preccurlyeq \sum_{k=1}^{p} \alpha_k B_k$.

(ii) $\mathsf{rav}_\gamma(B_k)_{1\leqslant k\leqslant p} \to \sum_{k=1}^{p} \alpha_k B_k$ *as* $0 < \gamma \to 0$.

(iii) $\mathsf{rav}_\gamma(B_k)_{1\leqslant k\leqslant p} \to \left(\sum_{k=1}^{p} \alpha_k B_k^{-1}\right)^{-1}$ *as* $\gamma \to +\infty$.

Proof. Recall that $\mathsf{rav}_\gamma(B_k)_{1\leqslant k\leqslant p} = \overset{\diamond}{\mathsf{M}}_\gamma(\mathrm{Id}_{\mathcal{H}}, B_k)_{1\leqslant k\leqslant p} = \overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathrm{Id}_{\mathcal{H}}, B_k)_{1\leqslant k\leqslant p}$.

(i): This follows from items (i) and (iii)(b) in Corollary 4.18.

(ii): This follows from Corollary 4.18(ii).

(iii): This follows from Corollary 4.18(iii)(c). ☐

4.2.4 Nonexpansiveness of resolvent compositions

In this section, we build on the results of Section 4.2.3 to prove that the resolvent composition operations are nonexpansive with respect to the Thompson metric [30] on $\mathcal{S}(\mathcal{H})$, defined by

$$\big(\forall A \in \mathcal{S}(\mathcal{H})\big)\big(\forall B \in \mathcal{S}(\mathcal{H})\big) \quad d_T^{\mathcal{H}}(A, B) = \ln\big(\max\{g(A, B), g(B, A)\}\big), \tag{4.49}$$

where $g(A, B) = \inf\big\{\lambda \in \,]0, +\infty[\ \big|\ A \preccurlyeq \lambda B\big\}$.

The Thompson metric was originally defined on cones in Banach spaces [30]. Since $\mathcal{S}(\mathcal{H})$ is contained in the cone of monotone self-adjoint operators on $\mathcal{B}(\mathcal{H})$, which is closed and hence complete in the operator norm topology, and since every $A \in \mathcal{S}(\mathcal{H})$ satisfies $\alpha\mathrm{Id}_{\mathcal{H}} \preccurlyeq A \preccurlyeq \|A\|\mathrm{Id}_{\mathcal{H}}$ for some $\alpha \in \,]0, +\infty[$, it follows from [30, Lemma 3] that $(\mathcal{S}(\mathcal{H}), d_T^{\mathcal{H}})$ is a complete metric space. The metric $d_T^{\mathcal{H}}$ provides a geometric structure on $\mathcal{S}(\mathcal{H})$ that plays a central role in the study of nonlinear matrix equations, especially for establishing existence and uniqueness results via Banach contraction mappings [23–26], and in various applications to nonlinear optimization [13, 21, 27]. In this context, the nonexpansiveness of resolvent compositions is crucial, as it ensures that the resulting operations preserve both the metric structure and the stability necessary for analysis. For instance, in Section 4.2.5, we present two nonlinear equations based on resolvent compositions that admit unique solutions.

Theorem 4.20 *Suppose that* $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ *is bounded below and satisfies* $\|L\| \leqslant 1$*, and let* $\gamma \in \,]0, +\infty[$*. Then the following hold:*

(i) $T_\gamma \colon \big(\mathcal{S}(\mathcal{G}), d_T^{\mathcal{G}}\big) \to \big(\mathcal{S}(\mathcal{H}), d_T^{\mathcal{H}}\big) \colon B \mapsto L \overset{\gamma}{\blacklozenge} B$ *is nonexpansive, i.e.,*

$$\big(\forall A \in \mathcal{S}(\mathcal{G})\big)\big(\forall B \in \mathcal{S}(\mathcal{G})\big) \quad d_T^{\mathcal{H}}\big(L \overset{\gamma}{\blacklozenge} A, L \overset{\gamma}{\blacklozenge} B\big) \leqslant d_T^{\mathcal{G}}(A, B). \tag{4.50}$$

(ii) $R_\gamma \colon \big(\mathcal{S}(\mathcal{G}), d_T^{\mathcal{G}}\big) \to \big(\mathcal{S}(\mathcal{H}), d_T^{\mathcal{H}}\big) \colon B \mapsto L \overset{\gamma}{\lozenge} B$ *is nonexpansive, i.e.,*

$$\big(\forall A \in \mathcal{S}(\mathcal{G})\big)\big(\forall B \in \mathcal{S}(\mathcal{G})\big) \quad d_T^{\mathcal{H}}\big(L \overset{\gamma}{\lozenge} A, L \overset{\gamma}{\lozenge} B\big) \leqslant d_T^{\mathcal{G}}(A, B). \tag{4.51}$$

Proof. Let A and B be in $\mathcal{S}(\mathcal{G})$, and set $g(A, B) = \inf\big\{\lambda \in \,]0, +\infty[\;\big|\; A \preccurlyeq \lambda B\big\}$.

(i): Note that the operator T_γ is well defined by Proposition 4.9(iv)(a). By virtue of (4.49),

$$A \preccurlyeq e^{d_T^{\mathcal{G}}(A,B)} B. \tag{4.52}$$

On the other hand, it follows from [12, Proposition 3.1(vi)] and Proposition 4.11(iii) that

$$(\forall \rho \in [1, +\infty[) \quad L \overset{\gamma}{\blacklozenge} (\rho B) = \rho\big(L \overset{\gamma\rho}{\blacklozenge} B\big) \preccurlyeq \rho\big(L \overset{\gamma}{\blacklozenge} B\big). \tag{4.53}$$

Since $e^{d_T^{\mathcal{G}}(A,B)} \geqslant 1$, we combine Proposition 4.11(ii), (4.52), and (4.53) to obtain

$$L \overset{\gamma}{\blacklozenge} A \preccurlyeq L \overset{\gamma}{\blacklozenge} \big(e^{d_T^{\mathcal{G}}(A,B)} B\big) \preccurlyeq e^{d_T^{\mathcal{G}}(A,B)} \big(L \overset{\gamma}{\blacklozenge} B\big). \tag{4.54}$$

In turn,

$$g\big(L \overset{\gamma}{\blacklozenge} A, L \overset{\gamma}{\blacklozenge} B\big) = \inf\big\{\lambda \in \,]0, +\infty[\;\big|\; L \overset{\gamma}{\blacklozenge} A \preccurlyeq \lambda(L \overset{\gamma}{\blacklozenge} B)\big\} \leqslant e^{d_T^{\mathcal{G}}(A,B)}. \tag{4.55}$$

By the same argument,

$$g\big(L \overset{\gamma}{\blacklozenge} B, L \overset{\gamma}{\blacklozenge} A\big) \leqslant e^{d_T^{\mathcal{G}}(A,B)}. \tag{4.56}$$

Altogether, it follows from (4.49), (4.55), and (4.56) that

$$d_T^{\mathcal{H}}\big(L \overset{\gamma}{\blacklozenge} A, L \overset{\gamma}{\blacklozenge} B\big) = \max\big\{\ln g\big(L \overset{\gamma}{\blacklozenge} A, L \overset{\gamma}{\blacklozenge} B\big), \ln g\big(L \overset{\gamma}{\blacklozenge} B, L \overset{\gamma}{\blacklozenge} A\big)\big\} \leqslant d_T^{\mathcal{G}}(A, B). \tag{4.57}$$

(ii): Note that R_γ is well defined by Proposition 4.9(iv)(a). Since $d_T^{\mathcal{G}}(A, B) =$

$d_T^{\mathcal{G}}(A^{-1}, B^{-1})$, we deduce from (i) and (4.4) that

$$d_T^{\mathcal{H}}\big(L \overset{\gamma}{\diamond} A, L \overset{\gamma}{\diamond} B\big) = d_T^{\mathcal{H}}\big(L \overset{1/\gamma}{\blacklozenge} A^{-1}, L \overset{1/\gamma}{\blacklozenge} B^{-1}\big) \leqslant d_T^{\mathcal{G}}(A^{-1}, B^{-1}) = d_T^{\mathcal{G}}(A, B), \tag{4.58}$$

as announced. ☐

Corollary 4.21 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ is bounded below and satisfies $\|L\| \leqslant 1$, and let $\gamma \in \,]0, +\infty[$. Then the following hold:*

(i) $T_0 \colon \big(\mathcal{S}(\mathcal{G}), d_T^{\mathcal{G}}\big) \to \big(\mathcal{S}(\mathcal{H}), d_T^{\mathcal{H}}\big) \colon B \mapsto L^* \circ B \circ L$ *is nonexpansive.*

(ii) $R_{+\infty} \colon \big(\mathcal{S}(\mathcal{G}), d_T^{\mathcal{G}}\big) \to \big(\mathcal{S}(\mathcal{H}), d_T^{\mathcal{H}}\big) \colon B \mapsto L^* \triangleright B$ *is nonexpansive.*

Proof. By Lemma 4.6(iii), the operators $\mathcal{S}(\mathcal{G}) \to \mathcal{S}(\mathcal{H}) \colon B \mapsto L^* \circ B \circ L$ and $\mathcal{S}(\mathcal{G}) \to \mathcal{S}(\mathcal{H}) \colon B \mapsto L^* \triangleright B$ are well defined.

(i): This follows from Theorem 4.20(i) and Theorem 4.14(i).

(ii): This follows from Theorem 4.20(ii) and Theorem 4.14(ii). ☐

Corollary 4.22 *Consider the setting of Example 4.1. Suppose that L_j is bounded below for some $j \in \{1, \ldots, p\}$ and that, for every $k \in \{1, \ldots, p\}$, $A_k \in \mathcal{S}(\mathcal{G}_k)$, and set $A \colon \mathcal{G} \to \mathcal{G} \colon (y_k)_{1 \leqslant k \leqslant p} \mapsto (A_k y_k)_{1 \leqslant k \leqslant p}$. Then*

$$d_T^{\mathcal{H}}\Big(\overset{\diamond}{\mathsf{M}}_\gamma(L_k, A_k)_{1 \leqslant k \leqslant p}, \overset{\diamond}{\mathsf{M}}_\gamma(L_k, B_k)_{1 \leqslant k \leqslant p}\Big) \leqslant d_T^{\mathcal{G}}(A, B) = \max_{1 \leqslant k \leqslant p} d_{\mathcal{G}_k}(A_k, B_k) \tag{4.59}$$

and

$$d_T^{\mathcal{H}}\Big(\overset{\blacklozenge}{\mathsf{M}}_\gamma(L_k, A_k)_{1 \leqslant k \leqslant p}, \overset{\blacklozenge}{\mathsf{M}}_\gamma(L_k, B_k)_{1 \leqslant k \leqslant p}\Big) \leqslant d_T^{\mathcal{G}}(A, B) = \max_{1 \leqslant k \leqslant p} d_{\mathcal{G}_k}(A_k, B_k). \tag{4.60}$$

In other words, the resolvent mixtures are nonexpansive for the Thompson metric.

Proof. It is straightforward to verify that $d_T^{\mathcal{G}}(A, B) = \max_{1 \leqslant k \leqslant p} d_{\mathcal{G}_k}(A_k, B_k)$. On the other hand, $L \overset{\gamma}{\diamond} A = \overset{\diamond}{\mathsf{M}}_\gamma(L_k, A_k)_{1 \leqslant k \leqslant p}$ and $L \overset{\gamma}{\blacklozenge} A = \overset{\blacklozenge}{\mathsf{M}}_\gamma(L_k, A_k)_{1 \leqslant k \leqslant p}$. Hence, the assertion follows from Theorem 4.20. ☐

Corollary 4.23 ([19, Theorem 3.5]) *Consider the setting of Example 4.2. Suppose that, for every $k \in \{1, \ldots, p\}$, $A_k \in \mathcal{S}(\mathcal{H})$, and set $A \colon \mathcal{G} \to \mathcal{G} \colon (y_k)_{1 \leqslant k \leqslant p} \mapsto (A_k y_k)_{1 \leqslant k \leqslant p}$. Then*

$$d_T^{\mathcal{H}}\big(\mathsf{rav}_\gamma(A_k)_{1 \leqslant k \leqslant p}, \mathsf{rav}_\gamma(B_k)_{1 \leqslant k \leqslant p}\big) \leqslant d_T^{\mathcal{G}}(A, B). \tag{4.61}$$

In other words, the resolvent average is nonexpansive for the Thompson metric.

Proof. Since $\mathsf{rav}_\gamma(A_k)_{1 \leqslant k \leqslant p} = \overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathrm{Id}_{\mathcal{H}}, A_k)_{1 \leqslant k \leqslant p}$, the conclusion follows from Corollary 4.22. ☐

4.2.5 Geometric means and nonlinear equations

Let $A \in \mathcal{S}(\mathcal{G})$. Since A is strongly positive, there exists $\alpha \in \left]0,+\infty\right[$ such that $\alpha\mathrm{Id}_{\mathcal{G}} \preccurlyeq A$. Consequently, the spectrum $\sigma(A)$ is contained in the compact interval $[\alpha, \|A\|]$ (see [28, Theorem VI.6 and Problem VII.12]). Hence, for every $t \in \mathbb{R}$, the function $f_t\colon \sigma(A) \to \mathbb{R}\colon \lambda \to \lambda^t$ is well defined and continuous, and the operator $f_t(A) = A^t$ is therefore defined according to the continuous functional calculus (see [28, Section VII]).

Given $A \in \mathcal{S}(\mathcal{G})$ and $B \in \mathcal{S}(\mathcal{G})$, an important instance of Kubo-Ando's operator means [18] is the *t-weighted geometric mean* [1, 19, 22, 26] of A and B, defined by

$$(\forall t \in [0,1]) \quad A\#_t B = A^{1/2} \circ \left(A^{-1/2} \circ B \circ A^{-1/2}\right)^t \circ A^{1/2}. \tag{4.62}$$

From a geometric viewpoint, the curve $t \mapsto A\#_t B$ describes a minimal geodesic between A and B with respect to the Thompson metric (see, e.g., [22, Lemma 2.2(iv)]), in the sense that

$$(\forall t \in [0,1])(\forall s \in [0,1]) \quad d_T^{\mathcal{G}}(A\#_t B, A\#_s B) = |t-s|\, d_T^{\mathcal{G}}(A,B). \tag{4.63}$$

In particular, the geometric mean $A\#B = A\#_{1/2}B$ is the metric midpoint of the arithmetic mean $(A+B)/2$ and the harmonic mean $2(A^{-1}+B^{-1})^{-1}$ for the Thompson metric (see [11, 20]).

The following result introduces a new interpolation between $L^* \triangleright B$ and $L^* \circ B \circ L$, which generalizes the weighted $\mathscr{A}\#\mathscr{H}$–mean discussed in Example 4.3.

Proposition 4.24 *Suppose that $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ is an isometry, let $B \in \mathcal{S}(\mathcal{G})$, and let $\gamma \in \left]0,+\infty\right[$. Define*

$$\mathcal{L}_\gamma(L,B) = \big(L^* \circ (B+\gamma\mathrm{Id}_{\mathcal{G}}) \circ L\big)\#\big(L^* \triangleright (B+\gamma\mathrm{Id}_{\mathcal{G}})\big) - \gamma\mathrm{Id}_{\mathcal{H}} \tag{4.64}$$

and

$$\mathcal{L}_{-\gamma}(L,B) = \Big(\mathcal{L}_\gamma\big(L,B^{-1}\big)\Big)^{-1}. \tag{4.65}$$

Then the following hold:

(i) $L^* \triangleright B \preccurlyeq \mathcal{L}_{-\gamma}(L,B) \preccurlyeq L \overset{\gamma}{\blacklozenge} B \preccurlyeq \mathcal{L}_{1/\gamma}(L,B) \preccurlyeq L^* \circ B \circ L$.

(ii) $\mathcal{L}_\gamma(L,B) \to L^* \circ B \circ L$ *as* $\gamma \to +\infty$.

(iii) $\mathcal{L}_\gamma(L,B) \to L^* \triangleright B$ *as* $\gamma \to -\infty$.

Proof. (i): Since L is an isometry, $L^* \circ L = \mathrm{Id}_{\mathcal{H}}$ and Lemma 4.8 yields $L \overset{\gamma}{\lozenge} B = L \overset{\gamma}{\blacklozenge} B$. By Corollary 4.17(i), (4.64), and the fact that $B\#B = B$,

$$\begin{aligned}\mathcal{L}_{1/\gamma}(L,B) &\preccurlyeq \big(L^* \circ (B+\gamma^{-1}\mathrm{Id}_{\mathcal{G}}) \circ L\big)\#\big(L^* \circ (B+\gamma^{-1}\mathrm{Id}_{\mathcal{G}}) \circ L\big) - \gamma^{-1}\mathrm{Id}_{\mathcal{H}}\\ &= \big(L^* \circ (B+\gamma^{-1}\mathrm{Id}_{\mathcal{G}}) \circ L\big) - \gamma^{-1}\mathrm{Id}_{\mathcal{H}}\\ &= L^* \circ B \circ L + \gamma^{-1}(L^* \circ L - \mathrm{Id}_{\mathcal{H}})\end{aligned}$$

$$= L^* \circ B \circ L. \tag{4.66}$$

Similarly, (4.3), Corollary 4.17(i), and (4.64), imply that

$$\begin{aligned}
L \overset{\gamma}{\diamond} B &= L^* \triangleright (B + \gamma^{-1}\mathrm{Id}_{\mathcal{G}}) - \gamma^{-1}\mathrm{Id}_{\mathcal{H}} \\
&= \big(L^* \triangleright (B + \gamma^{-1}\mathrm{Id}_{\mathcal{G}})\big) \# \big(L^* \triangleright (B + \gamma^{-1}\mathrm{Id}_{\mathcal{G}})\big) - \gamma^{-1}\mathrm{Id}_{\mathcal{H}} \\
&\preccurlyeq \big(L^* \circ (B + \gamma^{-1}\mathrm{Id}_{\mathcal{G}}) \circ L\big) \# \big(L^* \triangleright (B + \gamma^{-1}\mathrm{Id}_{\mathcal{G}})\big) - \gamma^{-1}\mathrm{Id}_{\mathcal{H}} \\
&= \mathcal{L}_{1/\gamma}(L, B).
\end{aligned} \tag{4.67}$$

Thus, (4.66) and (4.67) yield

$$L \overset{\gamma}{\blacklozenge} B \preccurlyeq \mathcal{L}_{1/\gamma}(L, B) \preccurlyeq L^* \circ B \circ L. \tag{4.68}$$

On the other hand, by virtue of Lemma 4.4(ii), (4.68) applied to B^{-1} and $1/\gamma$, (4.64), and (4.4),

$$L^* \triangleright B = (L^* \circ B^{-1} \circ L)^{-1} \preccurlyeq \mathcal{L}_{\gamma}(L, B^{-1})^{-1} = \mathcal{L}_{-\gamma}(L, B) \preccurlyeq \big(L \overset{1/\gamma}{\blacklozenge} B^{-1}\big)^{-1} = L \overset{\gamma}{\diamond} B = L \overset{\gamma}{\blacklozenge} B. \tag{4.69}$$

Hence, the result follows from (4.68) and (4.69).

(ii): This follows from (i) and Corollary 4.17(ii).

(iii): This follows from (i) and Corollary 4.17(iii). ☐

Remark 4.25 Note that the operator $\mathcal{L}_{\gamma}(L, B)$ is a type of weighted geometric mean that interpolates between the parallel composition $L^* \triangleright B$ ($\gamma \to -\infty$) and $L^* \circ B \circ L$ ($\gamma \to +\infty$). In the particular case where L and B are defined as in Example 4.2, $L^* \circ B \circ L = \sum_{k=1}^{p} \alpha_k B_k$ is the arithmetic average, $L^* \triangleright B = \big(\sum_{k=1}^{p} \alpha_k B_k^{-1}\big)^{-1}$ is the harmonic average, and $\mathcal{L}_{\gamma}(L, B)$ reduces to the *weighted* $\mathscr{A}\#\mathscr{H}$*-mean* with parameter γ of Example 4.3, with Proposition 4.24(ii)–(iii) recovering [17, Proposition 3.4].

We now focus on nonlinear equations that are based on resolvent compositions. The nonexpansive nature of these operations, as shown in Section 4.2.4, will play a key role in our subsequent analysis.

Proposition 4.26 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ is bounded below and satisfies $\|L\| \leqslant 1$, let $B \in \mathcal{S}(\mathcal{G})$, let $\gamma \in \,]0, +\infty[$, and let $t \in \,]0, 1[$. Set*

$$\varphi \colon \big(\mathcal{S}(\mathcal{G}), d_T^{\mathcal{G}}\big) \to \big(\mathcal{S}(\mathcal{H}), d_T^{\mathcal{H}}\big) \colon X \mapsto L \overset{\gamma}{\blacklozenge} (X \#_t B). \tag{4.70}$$

Then the following hold:

(i) *φ is $(1-t)$-Lipschitzian.*

(ii) *Suppose that $\mathcal{H} = \mathcal{G}$. Then the problem*

$$\text{find} \quad X \in \mathcal{S}(\mathcal{H}) \quad \text{such that} \quad X = L \overset{\gamma}{\blacklozenge} (X \#_t B) \tag{4.71}$$

admits a unique solution.

Proof. (i): It follows from Theorem 4.20(i) and [22, Lemma 2.2(iii)] that

$$\begin{aligned}\big(\forall X \in \mathcal{S}(\mathcal{G})\big)\big(\forall Y \in \mathcal{S}(\mathcal{G})\big) \quad d_T^{\mathcal{H}}\big(\varphi(X), \varphi(Y)\big) &= d_T^{\mathcal{H}}\big(L \overset{\gamma}{\blacklozenge} (X \#_t B), L \overset{\gamma}{\blacklozenge} (Y \#_t B)\big) \\ &\leqslant d_T^{\mathcal{G}}(X \#_t B, Y \#_t B) \\ &\leqslant (1-t) d_T^{\mathcal{G}}(X, Y) + t d_T^{\mathcal{G}}(B, B) \\ &= (1-t) d_T^{\mathcal{G}}(X, Y). \end{aligned} \tag{4.72}$$

(ii): Since $d_T^{\mathcal{H}}$ is a complete metric on $\mathcal{S}(\mathcal{H})$, (i) and the Banach–Picard theorem [3, Theorem 1.50] ensure that φ admits a unique fixed point, i.e., (4.71) admits a unique solution. ◻

Remark 4.27 Let $X \in \mathcal{S}(\mathcal{H})$ be the unique solution to (4.71). Since $(X \#_t B)^{-1} = X^{-1} \#_t B^{-1}$ and $L \overset{\gamma}{\blacklozenge} B = (L \overset{1/\gamma}{\diamond} B^{-1})^{-1}$, we note that X^{-1} is the unique solution to the problem

$$\text{find} \quad Y \in \mathcal{S}(\mathcal{H}) \quad \text{such that} \quad Y = L \overset{1/\gamma}{\diamond} (Y \#_t B^{-1}). \tag{4.73}$$

Proposition 4.28 *Suppose that $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ is bounded below and satisfies $\|L\| \leqslant 1$, let $B \in \mathcal{S}(\mathcal{G})$, let $\gamma \in \left]0, +\infty\right[$, let $t \in \left]-1, 1\right[$, and set*

$$\varphi \colon \big(\mathcal{S}(\mathcal{G}), d_T^{\mathcal{G}}\big) \to \big(\mathcal{S}(\mathcal{H}), d_T^{\mathcal{H}}\big) \colon X \mapsto L \overset{\gamma}{\blacklozenge} (B^* \circ X^t \circ B). \tag{4.74}$$

Then the following hold:

(i) *φ is $|t|$-Lipschitzian.*

(ii) *Suppose that $\mathcal{H} = \mathcal{G}$. Then the problem*

$$\text{find} \quad X \in \mathcal{S}(\mathcal{H}) \quad \text{such that} \quad X = L \overset{\gamma}{\blacklozenge} (B^* \circ X^t \circ B) \tag{4.75}$$

admits a unique solution.

Proof. Since $B \in \mathcal{S}(\mathcal{G})$, for every $X \in \mathcal{S}(\mathcal{G})$, $B^* \circ X^t \circ B \in \mathcal{S}(\mathcal{G})$. Hence, φ is well defined by virtue of Proposition 4.9(iv)(a).

(i): Let $X \in \mathcal{S}(\mathcal{G})$ and $Y \in \mathcal{S}(\mathcal{G})$. By [22, Lemma 2.2(i)],

$$d_T^{\mathcal{G}}(B^* \circ X^t \circ B, B^* \circ Y^t \circ B) = d_T^{\mathcal{G}}(X^t, Y^t) = d_T^{\mathcal{G}}(X^{|t|}, Y^{|t|}). \tag{4.76}$$

Thus, combining Theorem 4.20(i), (4.76), and [22, Lemma 2.2(iii)],

$$\begin{aligned} d_T^{\mathcal{H}}\big(L \overset{\gamma}{\blacklozenge} (B^* \circ X^t \circ B), L \overset{\gamma}{\blacklozenge} (B^* \circ Y^t \circ B)\big) &\leqslant d_T^{\mathcal{G}}(B^* \circ X^t \circ B, B^* \circ Y^t \circ B) \\ &= d_T^{\mathcal{G}}(X^{|t|}, Y^{|t|}) \\ &= d_T^{\mathcal{G}}(\mathrm{Id}_{\mathcal{G}} \#_{|t|} X, \mathrm{Id}_{\mathcal{G}} \#_{|t|} Y) \\ &\leqslant |t| d_T^{\mathcal{G}}(X, Y). \end{aligned} \tag{4.77}$$

(ii): This follows from (i) and the Banach–Picard theorem. ☐

Corollary 4.29 *Consider the setting of Example 4.1. Suppose that L_j is bounded below for some $j \in \{1, \ldots, p\}$ and that, for every $k \in \{1, \ldots, p\}$, $\mathcal{G}_k = \mathcal{H}$, and let $s \in \left]0,1\right[$ and $t \in \left]-1,1\right[$. Then the problems*

$$\textit{find} \quad X \in \mathcal{S}(\mathcal{H}) \quad \textit{such that} \quad X = \overset{\blacklozenge}{\mathsf{M}}_\gamma\big(L_k, X \#_s B_k\big)_{1 \leqslant k \leqslant p} \tag{4.78}$$

and

$$\textit{find} \quad X \in \mathcal{S}(\mathcal{H}) \quad \textit{such that} \quad X = \overset{\blacklozenge}{\mathsf{M}}_\gamma\big(L_k, B_k^* \circ X^t \circ B_k\big)_{1 \leqslant k \leqslant p} \tag{4.79}$$

admit unique solutions.

Proof. Set $R\colon \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{G})\colon X \mapsto \boldsymbol{\mathcal{X}}$, where $\boldsymbol{\mathcal{X}}\colon \mathcal{G} \to \mathcal{G}\colon (y_k) \mapsto (X y_k)_{1 \leqslant k \leqslant p}$, and set

$$\varphi_1\colon \big(\mathcal{S}(\mathcal{G}), d_T^{\mathcal{G}}\big) \to \big(\mathcal{S}(\mathcal{H}), d_T^{\mathcal{H}}\big)\colon X \mapsto L \overset{\gamma}{\blacklozenge} (X \#_s B) \tag{4.80}$$

and

$$\varphi_2\colon \big(\mathcal{S}(\mathcal{G}), d_T^{\mathcal{G}}\big) \to \big(\mathcal{S}(\mathcal{H}), d_T^{\mathcal{H}}\big)\colon X \mapsto L \overset{\gamma}{\blacklozenge} (B^* \circ X^t \circ B). \tag{4.81}$$

Note that $(\forall \lambda \in \left]0,+\infty\right[)\ X \preccurlyeq \lambda Y \Rightarrow \boldsymbol{\mathcal{X}} \preccurlyeq \lambda \boldsymbol{\mathcal{Y}}$. Thus,

$$g(\boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{Y}}) = \inf\big\{\lambda \in \left]0,+\infty\right[\bigm| \boldsymbol{\mathcal{X}} \preccurlyeq \lambda \boldsymbol{\mathcal{Y}}\big\} \leqslant \inf\big\{\lambda \in \left]0,+\infty\right[\bigm| X \preccurlyeq \lambda Y\big\} = g(X, Y), \tag{4.82}$$

and it follows from (4.49) that

$$d_T^{\mathcal{G}}\big(R(X), R(Y)\big) = d_T^{\mathcal{G}}(\boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{Y}}) \leqslant d_T^{\mathcal{H}}(X, Y). \tag{4.83}$$

Now, given that R is nonexpansive, Propositions 4.26(i) implies that $\varphi_1 \circ R$ is $(1-s)$-Lipschitzian, whereas Proposition 4.28(i) implies that $\varphi_2 \circ R$ is $|t|$-Lipschitzian. Further, since $\boldsymbol{\mathcal{X}} \#_s B\colon \mathcal{G} \to \mathcal{G}\colon (y_k)_{1 \leqslant k \leqslant p} \mapsto \big((X \#_s B_k) y_k\big)_{1 \leqslant k \leqslant p}$ and $B^* \circ X^t \circ B\colon \mathcal{G} \to \mathcal{G}\colon (y_k)_{1 \leqslant k \leqslant p} \mapsto$

$\big((B_k^* \circ X^t \circ B_k)y_k\big)_{1 \leqslant k \leqslant p}$, we deduce that

$$\varphi_1 \circ T \colon \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H}) \colon X \mapsto L \overset{\gamma}{\blacklozenge} (\boldsymbol{X} \#_s B) = \overset{\blacklozenge}{\mathsf{M}}_\gamma(L_k, X \#_s B_k)_{1 \leqslant k \leqslant p} \tag{4.84}$$

and

$$\varphi_2 \circ T \colon \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H}) \colon X \mapsto L \overset{\gamma}{\blacklozenge} (B^* \circ X^t \circ B) = \overset{\blacklozenge}{\mathsf{M}}_\gamma(L_k, B_k^* \circ X^t \circ B_k)_{1 \leqslant k \leqslant p}. \tag{4.85}$$

Altogether, it follows from the Banach–Picard theorem that $\varphi_1 \circ T$ and $\varphi_2 \circ T$ admit unique fixed points, i.e., the problems (4.78) and (4.79) admit unique solutions. ☐

Corollary 4.30 *Consider the setting of Corollary 4.29. Then the problems*

$$\textit{find} \quad X \in \mathcal{S}(\mathcal{H}) \quad \textit{such that} \quad X = \sum_{k=1}^{p} \alpha_k L_k^* \circ (X \#_s B_k) \circ L_k \tag{4.86}$$

and

$$\textit{find} \quad X \in \mathcal{S}(\mathcal{H}) \quad \textit{such that} \quad X = \sum_{k=1}^{p} \alpha_k L_k^* \circ (B_k^* \circ X^t \circ B_k) \circ L_k \tag{4.87}$$

admit unique solutions.

Proof. As shown in the proof of Corollary 4.29, the operators

$$\mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H}) \colon X \mapsto \overset{\blacklozenge}{\mathsf{M}}_\gamma(L_k, X \#_s B_k)_{1 \leqslant k \leqslant p} \tag{4.88}$$

and

$$\mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H}) \colon X \mapsto \overset{\blacklozenge}{\mathsf{M}}_\gamma(L_k, B_k^* \circ X^t \circ B_k)_{1 \leqslant k \leqslant p} \tag{4.89}$$

are $(1-s)$-Lipschitzian and $|t|$-Lipchitzian, respectively. Therefore, by virtue of Lemma 4.6(iii) and Corollary 4.18(ii), letting $\gamma \to 0$, we deduce that operators

$$\mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H}) \colon X \mapsto \sum_{k=1}^{p} \alpha_k L_k^* \circ (X \#_s B_k) \circ L_k \tag{4.90}$$

and

$$\mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H}) \colon X \mapsto \sum_{k=1}^{p} \alpha_k L_k^* \circ (B_k^* \circ X^t \circ B_k) \circ L_k \tag{4.91}$$

are well defined and are $(1-s)$-Lipschitzian and $|t|$-Lipchitzian, respectively. Finally, the conclusion follows from the Banach–Picard theorem. ☐

Corollary 4.31 ([19, Theorem 4.2]) *Consider the setting of Example 4.2, and let $s \in \left]0,1\right[$ and*

$t \in \,]-1,1[$. *Then the problems*

$$\text{find} \quad X \in \mathcal{S}(\mathcal{H}) \quad \text{such that} \quad X = \mathsf{rav}_\gamma(X \#_s B_k)_{1 \leqslant k \leqslant p} \tag{4.92}$$

and

$$\text{find} \quad X \in \mathcal{S}(\mathcal{H}) \quad \text{such that} \quad X = \mathsf{rav}_\gamma(B_k^* \circ X^t \circ B_k)_{1 \leqslant k \leqslant p} \tag{4.93}$$

admit unique solutions.

Proof. A direct consequence of Corollary 4.29. ☐

Corollary 4.32 ([22, Theorem 3.1]) *Consider the setting of Example 4.2 and let* $s \in \,]0,1[$. *Then the problem*

$$\text{find} \quad X \in \mathcal{S}(\mathcal{H}) \quad \text{such that} \quad X = \sum_{k=1}^{p} \alpha_k (X \#_s B_k) \tag{4.94}$$

admits a unique solution.

Proof. A direct consequence of Corollary 4.30. ☐

Acknowledgments

This work is part of the author's Ph.D. dissertation. The author acknowledges the guidance of his Ph.D. advisor P. L. Combettes, throughout this work. The author also thanks the anonymous referee for valuable suggestions that improved the quality of the paper.

Funding Information

This work was supported by the National Science Foundation under grant DMS-2513409.

References

[1] T. Ando, On some operator inequalities, *Math. Ann.*, vol. 279, pp. 157–159, 1987.

[2] S. Bartz, H. H. Bauschke, S. M. Moffat, and X. Wang, The resolvent average of monotone operators: Dominant and recessive properties, *SIAM J. Optim.*, vol. 26, pp. 602–634, 2016.

[3] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces,* 2nd ed. Springer, New York, 2017.

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

[5] M. N. Bùi and P. L. Combettes, Integral resolvent and proximal mixtures, *J. Optim. Theory Appl.*, vol. 203, pp. 2328–2353, 2024.

[6] P. L. Combettes, Resolvent and proximal compositions, *Set-Valued Var. Anal.*, vol. 31, art. 22, 29 pp., 2023.

[7] P. L. Combettes, The geometry of monotone operator splitting methods, *Acta Numer.*, vol. 33, pp. 487–632, 2024.

[8] P. L. Combettes and D. J. Cornejo, Signal recovery with proximal comixtures, *Proc. Europ. Signal Process. Conf.*, pp. 2637–2641, Lyon, France, August 26–30, 2024.

[9] P. L. Combettes and D. J. Cornejo, Variational analysis of proximal compositions and integral proximal mixtures. *Evol. Equ. Control Theory*, vol. 17, pp. 106–139, 2026.

[10] P. L. Combettes and J.-C. Pesquet, Fixed point strategies in data science, *IEEE Trans. Signal Process.*, vol. 69, pp. 3878–3905, 2021.

[11] G. Corach, H. Porta, and L. Recht, Convexity of the geodesic distance on spaces of positive operators, *Illinois J. Math.*, vol. 38, pp. 87–94, 1994.

[12] D. J. Cornejo, Parametrized families of resolvent compositions, *Set-Valued Var. Anal.*, vol. 33, art. 6, 24 pp., 2025.

[13] S. Gaubert and Z. Qu, The contraction rate in Thompson's part metric of order-preserving flows on a cone – Application to generalized Riccati equations, *J. Differential Equations*, vol. 256, pp. 2902–2948, 2014.

[14] M. Hasannasab, J. Hertrich, S. Neumayer, G. Plonka, S. Setzer, and G. Steidl, Parseval proximal neural networks, *J. Fourier Anal. Appl.*, vol. 26, art. 59, 31 pp., 2020.

[15] U. S. Kamilov, A parallel proximal algorithm for anisotropic total variation minimization, *IEEE Trans. Image Process.*, vol. 26, pp. 539–548, 2017.

[16] S. Kim, Mixture and interpolation of the parametrized ordered means, *J. Inequal. Appl.*, vol. 2022, art. 119, 16 pp., 2022.

[17] S. Kim, J. Lawson, and Y. Lim, The matrix geometric mean of parametrized, weighted arithmetic and harmonic means, *Linear Algebra Appl.*, vol. 435, pp. 2114–2131, 2011.

[18] F. Kubo and T. Ando, Means of positive linear operators, *Math. Ann.*, vol. 246, pp. 205–224, 1980.

[19] S. Kum and Y. Lim, Nonexpansiveness of the resolvent average, *J. Math. Anal. Appl.*, vol. 432, pp. 918–927, 2015.

[20] J. D. Lawson and Y. Lim, The geometric mean, matrices, metrics, and more, *Amer. Math. Monthly*, vol. 108, pp. 797–812, 2001.

[21] J. Lawson and Y. Lim, A Birkhoff contraction formula with applications to Riccati equations, *SIAM J. Control Optim.*, vol. 46, pp. 930–951, 2007.

[22] J. Lawson and Y. Lim, Karcher means and Karcher equations of positive definite operators, *Trans. Amer. Math. Soc. Ser. B*, vol. 1, pp. 1–22, 2014.

[23] H. Lee and Y. Lim, Invariant metrics, contractions and nonlinear matrix equations, *Nonlinearity*, vol. 21, pp. 857–878, 2008.

[24] B. Lemmens and M. Roelands, Unique geodesics for Thompson's metric, *Ann. Inst. Fourier (Grenoble)*, vol. 65, pp. 315–348, 2015.

[25] Y. Lim, Solving the nonlinear matrix equation $X = Q + \sum_{i=1}^{m} M_i X^{\delta_i} M_i^*$ via a contraction principle, *Linear Algebra Appl.*, vol. 430, pp. 1380–1383, 2009.

[26] Y. Lim and M. Pálfia, Matrix power means and the Karcher mean, *J. Funct. Anal.*, vol. 262, pp. 1498–1514, 2012.

[27] L. Montrucchio, Thompson metric, contraction property and differentiability of policy functions, *J. Econ. Behav. Organ.*, vol. 33, pp. 449–466, 1998.

[28] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. I: Functional Analysis*. Academic Press, New York, 1980.

[29] L. Shen, W. Liu, J. Huang, Y.-G. Jiang, and S. Ma, Adaptive proximal average approximation for composite convex minimization, *Proc. AAAI Conf. Artif. Intell.*, vol. 31, pp. 2513–2519, 2017.

[30] A. C. Thompson, On certain contraction mappings in a partially ordered vector space, *Proc. Amer. Math. Soc.*, vol. 14, pp. 438–443, 1963.

[31] X. Wang, Self-dual regularization of monotone operators via the resolvent average, *SIAM J. Optim.*, vol. 21, pp. 438–462, 2011.

[32] Y.-L. Yu, Better approximation and faster algorithm using the proximal average, *Proc. Conf. Adv. Neural Inform. Process. Syst.*, pp. 458–466, 2013.

Chapter 5

PROXIMAL COMIXTURE MINIMIZATION MODELS FOR IMAGE RECOVERY AND DATA ANALYSIS

5.1 Introduction and context

This chapter addresses question **(Q4)** of Chapter 1. We propose an alternative minimization model based on proximal comixtures.

This chapter presents the following journal article:

P. L. Combettes and D. J. Cornejo, Proximal comixture minimization models for image recovery and data analysis, submitted.

5.2 Article: Proximal comixture minimization models for image recovery and data analysis

Abstract. In minimization models for image recovery and data analysis problems, loss functions and linear operators are typically aggregated as an average of composite terms. Each term in the aggregate models a desired property of the ideal solution arising from the *a priori* knowledge and the observed data. We propose an alternative minimization model based on proximal comixtures, an operation which combines functions and linear operators in such a way that the proximity operator of the resulting function is computable explicitly in terms of the individual proximity and linear operators. The mathematical properties of this operation are analyzed and comparisons between proximal comixtures and standard composite averages are made. Numerical illustrations of the benefits of minimization models based on proximal

comixtures are provided in the context of image recovery and machine learning applications.

5.2.1 Introduction

The objective of image recovery and, more broadly, of various tasks in the areas of inverse problems and data analysis is to determine the value of a mathematical object (an image, a signal, a set of parameters, a distribution, a covariance matrix, a spectrum, etc.) conveying information of interest using experimental measurements and some *a priori* knowledge (data formation model, properties of the target solution, etc.). Over the past 60 years, convex optimization has established itself as one of the most reliable and efficient framework to formulate, analyze, and solve such problems [1, 11, 19, 21, 22, 24, 26, 30, 37, 43]. One typically aims at minimizing an aggregation of convex loss functions that model individually the desired properties of the ideal solution arising from the *a priori* knowledge and the observed data. The assumptions underlying the minimization models to be discussed are the following (see Section 5.2.2 for notation).

Assumption 5.1 *$\mathcal{H}$ is a real Hilbert space with scalar product $\langle\cdot\,|\,\cdot\rangle_{\mathcal{H}}$, associated norm $\|\cdot\|_{\mathcal{H}}$, and quadratic kernel $\mathscr{Q}_{\mathcal{H}}=\|\cdot\|^2_{\mathcal{H}}/2$, $f\in\Gamma_0(\mathcal{H})$, and $h\colon\mathcal{H}\to\mathbb{R}$ is a convex and differentiable function with a β^{-1}-Lipschitzian gradient for some $\beta\in\,]0,+\infty[$. Further, $0<p\in\mathbb{N}$ and, for every $k\in\{1,\ldots,p\}$, $\mathcal{G}_k$ is a real Hilbert space, $g_k\in\Gamma_0(\mathcal{G}_k)$, and $0\neq L_k\colon\mathcal{H}\to\mathcal{G}_k$ is a bounded linear operator. Finally, $(\alpha_k)_{1\leqslant k\leqslant p}$ are weights in $]0,+\infty[$ which satisfy $\sum_{k=1}^p\alpha_k\|L_k\|^2\leqslant 1$.*

The most prevalent minimization framework in data analysis consists in minimizing an objective function which aggregates the loss functions $(g_k)_{1\leqslant k\leqslant p}$ and the linear operators $(L_k)_{1\leqslant k\leqslant p}$ by means of a standard composite averaging operation as follows.

Problem 5.2 Under Assumption 5.1, the task is to

$$\underset{x\in\mathcal{H}}{\text{minimize}}\;\; f(x)+\big(\text{cav}(g_k,L_k)_{1\leqslant k\leqslant p}\big)(x)+h(x), \tag{5.1}$$

where

$$\text{cav}(g_k,L_k)_{1\leqslant k\leqslant p}=\sum_{k=1}^p\alpha_k g_k\circ L_k \tag{5.2}$$

is the *standard composite average* of $(g_k)_{1\leqslant k\leqslant p}$ and $(L_k)_{1\leqslant k\leqslant p}$.

An alternative way to aggregate the functions $(g_k)_{1\leqslant k\leqslant p}$ and the linear operators $(L_k)_{1\leqslant k\leqslant p}$ is via the proximal comixture operation. This operation is derived from the proximal mixture operation recently introduced in [15] by duality, and it is further studied in [8, 17]. The main objective of the present paper is to propose the use of this aggregation process as an alternative to the standard composite average of Problem 5.2. This brings us to the following minimization model, which involves the conjugation operation of (5.12).

Problem 5.3 Let $\gamma \in \left]0, +\infty\right[$. Under Assumption 5.1, the task is to

$$\underset{x \in \mathcal{H}}{\text{minimize}} \;\; f(x) + \big(\text{pcm}_\gamma(g_k, L_k)_{1 \leqslant k \leqslant p}\big)(x) + h(x), \tag{5.3}$$

where

$$\text{pcm}_\gamma(g_k, L_k)_{1 \leqslant k \leqslant p} = \left(\left(\sum_{k=1}^{p} \alpha_k \Big(g_k^* + \gamma \mathscr{Q}_{\mathcal{G}_k}\Big)^* \circ L_k\right)^* - \gamma \mathscr{Q}_{\mathcal{H}}\right)^* \tag{5.4}$$

is the *proximal comixture* of $(g_k)_{1 \leqslant k \leqslant p}$ and $(L_k)_{1 \leqslant k \leqslant p}$ with parameter γ.

Let us make some observations about these two formulations to motivate the latter.

- Both Problems 5.2 and 5.3 can be viewed as relaxations of the ideal feasibility problem

$$\text{find} \;\; x \in \mathcal{H} \quad \text{such that} \quad \begin{cases} x \in \text{Argmin}\, f \\ (\forall k \in \{1, \ldots, p\})\; L_k x \in \text{Argmin}\, g_k \\ x \in \text{Argmin}\, h, \end{cases} \tag{5.5}$$

  which aims at imposing all the desired properties exactly and hence at minimizing all the loss functions simultaneously. However, the set

$$Z = \text{Argmin}\, f \cap \left(\bigcap_{k=1}^{p} L_k^{-1}(\text{Argmin}\, g_k)\right) \cap \text{Argmin}\, h \tag{5.6}$$

  of solutions to (5.5) is typically empty and, as will be seen in Remark 5.11(ii), both Problems 5.2 and 5.3 are relaxations of (5.5) in the sense that, if $Z \neq \varnothing$, then Z is the set of solutions of these two problems.

- The aggregation in Problem 5.2 may not be robust to perturbations. For instance, let us consider the special case when $f = 0$, $h = 0$, and, for every $k \in \{1, \ldots, p\}$, $\mathcal{G}_k = \mathcal{H}$, $L_k = \text{Id}_{\mathcal{H}}$, and g_k is the indicator function of a nonempty closed convex set $C_k \subset \mathcal{H}$. This reduces (5.1) to

$$\text{find} \;\; x \in \bigcap_{k=1}^{p} C_k, \tag{5.7}$$

  which happens to coincide with (5.5) in this case. If the sets $(C_k)_{1 \leqslant k \leqslant p}$ are not specified exactly, no solution may exist [10, 14]. However, it will follow from Remark 5.11(i) that the solutions to Problem 5.3 minimize the average squared distance to the sets. Such solutions are classical surrogates in inconsistent feasibility problems [14], which go back to Legendre's least-squares method for systems of linear equations [32]. They exist under mild conditions, such as the boundedness of one of the sets.

- Problem 5.2 involves a simple aggregation process by averaging. However, the nonsmooth function (5.2) has no explicit proximity operator, and solving (5.1) therefore calls for sophisticated proximal splitting methods to decompose the p functions $(g_k)_{1\leqslant k\leqslant p}$ and the p linear operators $(L_k)_{1\leqslant k\leqslant p}$ individually [16]. Overall, solving Problem 5.2 requires the splitting of $2p+2$ terms, which can be expected to lead to algorithms that are slower and necessitate more memory storage than those that would split less terms. By contrast, the aggregated function (5.4) in Problem 5.3 is less intuitive but, as will be seen in Proposition 5.10(ii), its proximity operator can be computed explicitly in terms of those of the functions $(g_k)_{1\leqslant k\leqslant p}$ and of the linear operators $(L_k)_{1\leqslant k\leqslant p}$ as

$$\operatorname{prox}_{\gamma \operatorname{pcm}_\gamma (g_k, L_k)_{1\leqslant k\leqslant p}} = \operatorname{Id}_{\mathcal{H}} - \sum_{k=1}^{p} \alpha_k L_k^* \circ \left(\operatorname{Id}_{\mathcal{G}_k} - \operatorname{prox}_{\gamma g_k}\right) \circ L_k. \tag{5.8}$$

  In turn, solving Problem 5.3 requires the splitting of only 3 terms.

- In the special case when $\sum_{k=1}^{p} \alpha_k = 1$ and, for every $k \in \{1, \ldots, p\}$, $\mathcal{G}_k = \mathcal{H}$ and $L_k = \operatorname{Id}_{\mathcal{H}}$, the proximal comixture of (5.4) reduces to the proximal average

$$\operatorname{pav}_\gamma (g_k)_{1\leqslant k\leqslant p} = \left(\left(\sum_{k=1}^{p} \alpha_k \left(g_k^* + \gamma \mathscr{Q}_{\mathcal{H}} \right)^* \right)^* - \gamma \mathscr{Q}_{\mathcal{H}} \right)^*, \tag{5.9}$$

  and (5.8) gives

$$\operatorname{prox}_{\gamma \operatorname{pav}_\gamma (g_k)_{1\leqslant k\leqslant p}} = \sum_{k=1}^{p} \alpha_k \operatorname{prox}_{\gamma g_k}. \tag{5.10}$$

  This aggregation process has been implicitly introduced in [35], extensively studied in [6], and used in data analysis problems in [2, 13, 29, 34, 40, 45], where its benefits over the standard average

$$\operatorname{ave}(g_k)_{1\leqslant k\leqslant p} = \sum_{k=1}^{p} \alpha_k g_k, \tag{5.11}$$

  are discussed.

- Formally, evaluating (5.4) at $\gamma = 0$ gives back (5.3). This connection will be made mathematically precise in Theorem 5.13, where we study the asymptotic behavior as $\gamma \downarrow 0$.

Notation and background are provided in Section 5.2.2. In Section 5.2.3, we study the mathematical properties of proximal comixtures and investigate connections between Problems 5.2 and 5.3. Section 5.2.4 is devoted to algorithms for solving Problems 5.2 and 5.3. Finally, Section 5.2.5 provides numerical illustrations of proximal comixture models in concrete signal restoration, image reconstruction, and linear regression problems.

5.2.2 Notation and background

Throughout, $\mathcal{H}$ is a real Hilbert space with power set $2^{\mathcal{H}}$, identity operator $\mathrm{Id}_{\mathcal{H}}$, scalar product $\langle\cdot\mid\cdot\rangle_{\mathcal{H}}$, associated norm $\|\cdot\|_{\mathcal{H}}$, and quadratic kernel $\mathscr{Q}_{\mathcal{H}} = \|\cdot\|_{\mathcal{H}}^2/2$. For background on nonlinear analysis in Hilbert spaces, see [5].

Let $\varphi\colon \mathcal{H} \to [-\infty,+\infty]$. The conjugate of φ is

$$\varphi^*\colon \mathcal{H} \to [-\infty,+\infty]\colon x^* \mapsto \sup_{x\in\mathcal{H}}\big(\langle x \mid x^*\rangle_{\mathcal{H}} - \varphi(x)\big). \tag{5.12}$$

If $\inf_{y\in\mathcal{H}}\varphi(y) > -\infty$ and $\varepsilon \in [0,+\infty[$, the set of ε-minimizers of φ is

$$\varepsilon\text{–Argmin}\,\varphi = \big\{x\in\mathcal{H} \mid \varphi(x) \leqslant \inf_{y\in\mathcal{H}}\varphi(y) + \varepsilon\big\}. \tag{5.13}$$

In particular, the set of minimizers of φ is $\mathrm{Argmin}\,\varphi = 0\text{–Argmin}\,\varphi$. If $\mathrm{Argmin}\,\varphi$ is a singleton, its unique element is denoted by $\mathrm{argmin}_{x\in\mathcal{H}}\varphi(x)$. Moreover, φ is proper if $-\infty \notin \varphi(\mathcal{H})$ and $\mathrm{dom}\,\varphi = \big\{x\in\mathcal{H} \mid \varphi(x) < +\infty\big\} \neq \varnothing$. If φ is proper, its subdifferential is

$$\partial\varphi\colon \mathcal{H} \to 2^{\mathcal{H}}\colon x \mapsto \big\{x^*\in\mathcal{H} \mid (\forall y\in\mathcal{H})\ \langle y-x \mid x^*\rangle_{\mathcal{H}} + \varphi(x) \leqslant \varphi(y)\big\}. \tag{5.14}$$

Let $\gamma \in\,]0,+\infty[$. Then the *(lower) Moreau envelope* of φ with parameter γ is

$$\mathrm{lenv}_\gamma\,\varphi\colon \mathcal{H} \to [-\infty,+\infty]\colon x \mapsto \inf_{y\in\mathcal{H}}\Big(\varphi(y) + \frac{1}{\gamma}\mathscr{Q}_{\mathcal{H}}(x-y)\Big), \tag{5.15}$$

and the *upper Moreau envelope* of φ with parameter γ is

$$\mathrm{uenv}_\gamma\,\varphi\colon \mathcal{H} \to [-\infty,+\infty]\colon x \mapsto \sup_{y\in\mathcal{H}}\Big(\varphi(y) - \frac{1}{\gamma}\mathscr{Q}_{\mathcal{H}}(x-y)\Big). \tag{5.16}$$

Given $\rho\in\mathbb{R}$, $\Gamma_\rho(\mathcal{H})$ denotes the class of proper lower semicontinuous functions $\varphi\colon \mathcal{H} \to\,]-\infty,+\infty]$ such that $\varphi + \rho\mathscr{Q}_{\mathcal{H}}$ is convex. The proximity operator of $\varphi\in\Gamma_0(\mathcal{H})$ is

$$\mathrm{prox}_\varphi\colon \mathcal{H}\to\mathcal{H}\colon x \mapsto \mathrm{argmin}_{y\in\mathcal{H}}\big(\varphi(y) + \mathscr{Q}_{\mathcal{H}}(x-y)\big) \tag{5.17}$$

and it is characterized by

$$(\forall x\in\mathcal{H})(\forall p\in\mathcal{H})\quad p = \mathrm{prox}_\varphi\, x \quad\Leftrightarrow\quad x - p \in \partial\varphi(p). \tag{5.18}$$

An operator $T\colon \mathcal{H}\to\mathcal{H}$ is γ-cocoercive if

$$(\forall x\in\mathcal{H})(\forall y\in\mathcal{H})\quad \langle x-y \mid Tx - Ty\rangle_{\mathcal{H}} \geqslant \gamma\|Tx-Ty\|_{\mathcal{H}}^2. \tag{5.19}$$

Let $C\subset\mathcal{H}$. Then the indicator function of C is denoted by ι_C and the distance function to C

is denoted by d_C. If C is nonempty, closed, and convex, its projection operator is denoted by proj_C. The closed ball with center $x \in \mathcal{H}$ and radius $\rho \in \left]0,+\infty\right[$ is denoted by $B(x;\rho)$.

The following results will be useful.

Lemma 5.4 *Let $\varphi \in \Gamma_0(\mathcal{H})$ and let $\gamma \in \left]0,+\infty\right[$. Then the following hold:*

(i) $\operatorname{lenv}_\gamma \varphi = \big(\varphi^* + \gamma \mathscr{Q}_\mathcal{H}\big)^*$.

(ii) $\operatorname{uenv}_\gamma \varphi = \big(\varphi^* - \gamma \mathscr{Q}_\mathcal{H}\big)^*$.

Proof. Recall that, since $\varphi \in \Gamma_0(\mathcal{H})$, $\varphi^{**} = \varphi$ [5, Corollary 13.38].

(i): This follows from [5, Proposition 14.1].

(ii): Apply [27, Theorem 2.2] with $g = \varphi^*$ and $h = \gamma \mathscr{Q}_\mathcal{H}$. ☐

Lemma 5.5 ([7, Lemma 3]) *Let $\varphi\colon \mathcal{H} \to \left]-\infty,+\infty\right]$ be proper and let $\gamma \in \left]0,+\infty\right[$. Then the following are equivalent:*

(i) $\operatorname{uenv}_\gamma (\operatorname{lenv}_\gamma \varphi) = \varphi$.

(ii) $\varphi \in \Gamma_{1/\gamma}(\mathcal{H})$.

Lemma 5.6 *Let $\varphi\colon \mathcal{H} \to \mathbb{R}$ be continuous and convex, and let $\gamma \in \left]0,+\infty\right[$. Then the following are equivalent:*

(i) $\operatorname{lenv}_\gamma (\operatorname{uenv}_\gamma \varphi) = \varphi$.

(ii) $-\varphi \in \Gamma_{1/\gamma}(\mathcal{H})$.

(iii) $(\operatorname{uenv}_\gamma \varphi)^* = \varphi^* - \gamma \mathscr{Q}_\mathcal{H}$.

(iv) $\operatorname{uenv}_\gamma \varphi \in \Gamma_0(\mathcal{H})$, *$\varphi$ is Fréchet differentiable, and* $\operatorname{prox}_{\gamma(\operatorname{uenv}_\gamma \varphi)} = \operatorname{Id}_\mathcal{H} - \gamma \nabla \varphi$.

(v) *φ is Fréchet differentiable and $\nabla\varphi$ is γ-cocoercive.*

(vi) *φ is Fréchet differentiable and $\nabla\varphi$ is γ^{-1}-Lipschitzian.*

Proof. It follows from (5.15) and (5.16) that $\operatorname{lenv}_\gamma \varphi = -\operatorname{uenv}_\gamma(-\varphi)$. Therefore,

$$\operatorname{lenv}_\gamma (\operatorname{uenv}_\gamma \varphi) = -\operatorname{uenv}_\gamma \big(-\operatorname{uenv}_\gamma \varphi\big) = -\operatorname{uenv}_\gamma \big(\operatorname{lenv}_\gamma (-\varphi)\big). \tag{5.20}$$

(i)⇒(ii): By (5.20), $\operatorname{uenv}_\gamma (\operatorname{lenv}_\gamma(-\varphi)) = -\varphi$. Thus, Lemma 5.5 yields $-\varphi \in \Gamma_{1/\gamma}(\mathcal{H})$.

(ii)⇒(iii): By [5, Proposition 14.2], $\varphi^* \in \Gamma_{-\gamma}(\mathcal{H})$. Thus, the result follows from Lemma 5.4(ii) and [5, Corollary 13.38].

(iii)⇒(iv): Since $\varphi \in \Gamma_0(\mathcal{H})$, [5, Corollary 13.38] asserts that $\varphi^* \in \Gamma_0(\mathcal{H})$, which implies that $\operatorname{uenv}_\gamma \varphi$ is proper. In turn, it results from [5, Proposition 13.13] and Lemma 5.4(ii) that $\operatorname{uenv}_\gamma \varphi \in \Gamma_0(\mathcal{H})$. Thus, we derive from Lemma 5.4(i) and [5, Corollary 13.38] that

$$\operatorname{lenv}_\gamma (\operatorname{uenv}_\gamma \varphi) = \Big((\operatorname{uenv}_\gamma \varphi)^* + \gamma \mathscr{Q}_\mathcal{H}\Big)^* = \Big(\big(\varphi^* - \gamma \mathscr{Q}_\mathcal{H}\big) + \gamma \mathscr{Q}_\mathcal{H}\Big)^* = \varphi^{**} = \varphi. \tag{5.21}$$

Hence, [5, Proposition 12.30] guarantees that $\varphi = \operatorname{lenv}_\gamma\big(\operatorname{uenv}_\gamma \varphi\big)$ is Fréchet differentiable and that $\nabla\varphi = \gamma^{-1}(\mathrm{Id}_{\mathcal{H}} - \operatorname{prox}_{\gamma(\operatorname{uenv}_\gamma \varphi)})$.

(iv)⇒(v): This follows from [5, Proposition 12.28].

(v)⇒(vi): This follows from (5.19) and the Cauchy–Schwarz inequality.

(vi)⇒(i): It follows from the equivalence (i)⇔(vi) in [5, Theorem 18.15] that $-\varphi \in \Gamma_{1/\gamma}(\mathcal{H})$. We therefore deduce from Lemma 5.5 that $\operatorname{uenv}_\gamma(\operatorname{lenv}_\gamma(-\varphi)) = -\varphi$. We conclude via (5.20). ☐

Remark 5.7 The functions $\operatorname{lenv}_\gamma \varphi$ and $\operatorname{uenv}_\gamma \varphi$ are, respectively, the infimal and the supremal convolutions of φ and $\mathscr{Q}_{\mathcal{H}}/\gamma$. In [28, Definition 2.4], $\operatorname{uenv}_\gamma \varphi$ is called the deconvolution of φ by $\mathscr{Q}_{\mathcal{H}}/\gamma$. Furthermore, when $\gamma < \rho$, the functions $\operatorname{lenv}_\gamma(\operatorname{uenv}_\rho \varphi)$ and $\operatorname{uenv}_\gamma(\operatorname{lenv}_\rho \varphi)$ are known as the *Lasry-Lions regularizations* of φ, which were introduced in [31] and further studied in [4, 7, 38, 41], whereas $\operatorname{lenv}_\gamma(\operatorname{uenv}_\gamma \varphi)$ is called *proximal hull* of φ in [38, Example 1.44].

5.2.3 Properties of proximal comixtures

Throughout this section, Assumption 5.1 is in force. Recall from (5.2) that the associated standard composite average is

$$\operatorname{cav}(g_k, L_k)_{1\leqslant k\leqslant p} = \sum_{k=1}^{p} \alpha_k g_k \circ L_k, \tag{5.22}$$

and from (5.4) that the associated proximal comixture with parameter $\gamma \in \left]0, +\infty\right[$ is

$$\operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p} = \left(\left(\sum_{k=1}^{p} \alpha_k \Big(g_k^* + \gamma\mathscr{Q}_{\mathcal{G}_k}\Big)^* \circ L_k\right)^* - \gamma\mathscr{Q}_{\mathcal{H}}\right)^*. \tag{5.23}$$

We first provide reformulations of the proximal comixture of (5.23) in terms of the Moreau envelopes of (5.15) and (5.16) (see Fig. 5.1), and then in terms of the proximal average of (5.9) and of the following proximal cocomposition operation. These connections will not only shed new light on proximal comixtures but also play a role in forthcoming proofs.

Definition 5.8 ([15]) Let $\mathcal{G}$ be a real Hilbert space, let $g \in \Gamma_0(\mathcal{G})$, and let $L\colon \mathcal{H} \to \mathcal{G}$ be a bounded linear operator. The *proximal cocomposition* of g and L with parameter $\gamma \in \left]0, +\infty\right[$ is

$$L \overset{\gamma}{\blacklozenge} g = \left(\Big((g^* + \gamma\mathscr{Q}_{\mathcal{G}})^* \circ L\Big)^* - \gamma\mathscr{Q}_{\mathcal{H}}\right)^*. \tag{5.24}$$

Proposition 5.9 *Let $\gamma \in \left]0, +\infty\right[$. Then the following hold:*

(i) $\operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p} = \operatorname{uenv}_\gamma \sum_{k=1}^{p} \alpha_k(\operatorname{lenv}_\gamma g_k) \circ L_k$.

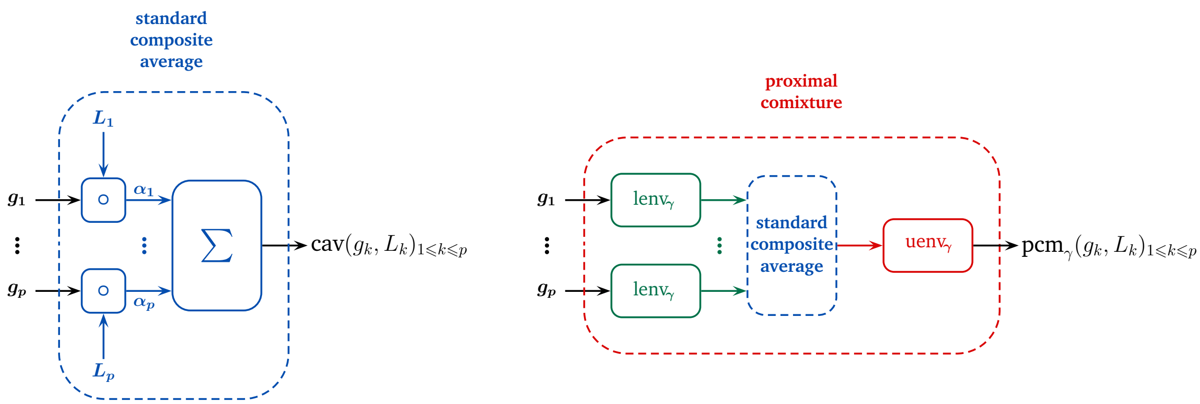


Figure 5.1 (left): Standard composite average (5.22). (right): Proximal comixture (5.23) in terms of the Moreau envelopes of (5.15) and (5.16) using Proposition 5.9(i).

(ii) *Suppose that* $\sum_{k=1}^p \alpha_k = 1$ *and* $\max_{1\leqslant k\leqslant p}\|L_k\| \leqslant 1$. *Then* $\mathrm{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p} = \mathrm{pav}_\gamma(L_k \overset{\gamma}{\blacklozenge} g_k)_{1\leqslant k\leqslant p}$.

(iii) *Let* $\mathcal{G}$ *be the standard product vector space* $\times_{k=1}^p \mathcal{G}_k$*, with generic element* $\boldsymbol{y} = (y_k)_{1\leqslant k\leqslant p}$*, and equipped with the scalar product*

$$\langle \cdot \mid \cdot \rangle_{\mathcal{G}} \colon (\boldsymbol{y}, \boldsymbol{y}') \mapsto \sum_{k=1}^p \alpha_k \langle y_k \mid y_k' \rangle_{\mathcal{G}_k}, \tag{5.25}$$

and set

$$L\colon \mathcal{H} \to \mathcal{G} \colon x \mapsto (L_k x)_{1\leqslant k\leqslant p} \quad \textit{and} \quad g\colon \mathcal{G} \to \left]-\infty, +\infty\right] \colon \boldsymbol{y} \mapsto \sum_{k=1}^p \alpha_k g_k(y_k). \tag{5.26}$$

Then $\mathrm{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p} = L \overset{\gamma}{\blacklozenge} g$.

Proof. Set $\varphi = \sum_{k=1}^p \alpha_k (\mathrm{lenv}_\gamma\, g_k) \circ L_k$. By virtue of [5, Proposition 12.30],

$$\nabla \varphi = \frac{1}{\gamma} \sum_{k=1}^p \alpha_k L_k^* \circ (\mathrm{Id}_{\mathcal{G}_k} - \mathrm{prox}_{\gamma g_k}) \circ L_k. \tag{5.27}$$

However, $\sum_{k=1}^p \alpha_k \|L_k\|^2 \leqslant 1$ by Assumption 5.1 and [5, Proposition 12.28] asserts that the operators $(\mathrm{Id}_{\mathcal{G}_k} - \mathrm{prox}_{\gamma g_k})_{1\leqslant k\leqslant p}$ are nonexpansive. Therefore, (5.27) implies that $\nabla\varphi$ is γ^{-1}-Lipschitzian.

(i): We derive from (5.23) and Lemma 5.4(i)–(ii) that

$$\mathrm{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p} = \left(\varphi^* - \gamma \mathscr{Q}_{\mathcal{H}}\right)^* = \mathrm{uenv}_\gamma\, \varphi. \tag{5.28}$$

(ii): It follows from [17, Proposition 3.13(ii)] that $(\forall k \in \{1,\ldots,p\})$ $\operatorname{lenv}_\gamma(L_k \overset{\gamma}{\blacklozenge} g_k) = (\operatorname{lenv}_\gamma g_k)\circ L_k$. Therefore, (i), Lemma 5.4(i)–(ii), and (5.9) yield

$$\begin{aligned}\operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p} &= \operatorname{uenv}_\gamma \varphi \\ &= \operatorname{uenv}_\gamma\left(\sum_{k=1}^p \alpha_k \operatorname{lenv}_\gamma\left(L_k \overset{\gamma}{\blacklozenge} g_k\right)\right) \\ &= \left(\left(\sum_{k=1}^p \alpha_k \operatorname{lenv}_\gamma\left(L_k \overset{\gamma}{\blacklozenge} g_k\right)\right)^* - \gamma\mathscr{Q}_{\mathcal{H}}\right)^* \\ &= \operatorname{pav}_\gamma\left(L_k \overset{\gamma}{\blacklozenge} g_k\right)_{1\leqslant k\leqslant p}. \end{aligned} \tag{5.29}$$

(iii): Since $\|\cdot\|_{\mathcal{G}}^2\colon \mathcal{G}\to\mathbb{R}\colon \boldsymbol{y}\mapsto \sum_{k=1}^p \alpha_k\|y_k\|_{\mathcal{G}_k}^2$, $\operatorname{lenv}_\gamma g\colon \mathcal{G}\to\mathbb{R}\colon \boldsymbol{y}\mapsto \sum_{k=1}^p \alpha_k(\operatorname{lenv}_\gamma g_k)(y_k)$. Therefore, $(\operatorname{lenv}_\gamma g)\circ L = \sum_{k=1}^p \alpha_k(\operatorname{lenv}_\gamma g_k)\circ L_k$, and the result follows from (i), items (ii) and (i) in Lemma 5.4, and (5.24). ☐

Proposition 5.10 *Let $\gamma\in\,]0,+\infty[$. Then the following hold:*

(i) $\operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p}\in\Gamma_0(\mathcal{H})$.

(ii) $\operatorname{prox}_{\gamma\operatorname{pcm}_\gamma(g_k,L_k)_{1\leqslant k\leqslant p}} = \operatorname{Id}_{\mathcal{H}} - \sum_{k=1}^p \alpha_k L_k^*\circ(\operatorname{Id}_{\mathcal{G}_k} - \operatorname{prox}_{\gamma g_k})\circ L_k$.

(iii) *Suppose that one of the following is satisfied:*

  (a) $\sum_{k=1}^p \alpha_k\|L_k\|^2 < 1$.

  (b) *For every $k\in\{1,\ldots,p\}$,* $\operatorname{dom} g_k = \mathcal{G}_k$.

  (c) $\sum_{k=1}^p \alpha_k = 1$, $\max_{1\leqslant k\leqslant p}\|L_k\|\leqslant 1$, *and there exists $\ell\in\{1,\ldots,p\}$ such that* $\operatorname{dom} g_\ell = \mathcal{G}_\ell$.

  Then $\operatorname{dom}\operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p} = \mathcal{H}$.

(iv) $\operatorname{lenv}_\gamma \operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p} = \sum_{k=1}^p \alpha_k(\operatorname{lenv}_\gamma g_k)\circ L_k$.

(v) $\operatorname{Argmin}\operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p} = \operatorname{Argmin}\sum_{k=1}^p \alpha_k(\operatorname{lenv}_\gamma g_k)\circ L_k$.

(vi) *Suppose that* $\bigcap_{k=1}^p L_k^{-1}(\operatorname{Argmin} g_k)\neq\varnothing$. *Then*

$$\operatorname{Argmin}\operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p} = \operatorname{Argmin}\operatorname{cav}(g_k, L_k)_{1\leqslant k\leqslant p} = \bigcap_{k=1}^p L_k^{-1}(\operatorname{Argmin} g_k). \tag{5.30}$$

Proof. Define $\mathcal{G}$, L, and g as in (5.26), and note that

$$(\forall x\in\mathcal{H})\quad \|Lx\|_{\mathcal{G}}^2 = \sum_{k=1}^p \alpha_k\|L_kx\|_{\mathcal{G}_k}^2 \leqslant \sum_{k=1}^p \alpha_k\|L_k\|^2\|x\|_{\mathcal{H}}^2, \tag{5.31}$$

which yields $\|L\|^2\leqslant\sum_{k=1}^p\alpha_k\|L_k\|^2\leqslant 1$ by Assumption 5.1. Further, by Proposition 5.9(iii),

$$\operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p} = L \overset{\gamma}{\blacklozenge} g. \tag{5.32}$$

(i): This follows from (5.32) and [17, Proposition 3.7(i)].

(ii): Note that $(\forall (y_k^*)_{1\leqslant k\leqslant p} \in \mathcal{G})\ L^*(y_k^*)_{1\leqslant k\leqslant p} = \sum_{k=1}^p \alpha_k L_k^* y_k^*$ and $\operatorname{prox}_{\gamma g}(y_k^*)_{1\leqslant k\leqslant p} = (\operatorname{prox}_{\gamma g_k} y_k^*)_{1\leqslant k\leqslant p}$. It therefore follows from (5.32) and [17, Proposition 3.10(ii)] that

$$\begin{aligned}\operatorname{prox}_{\gamma \operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p}} &= \operatorname{prox}_{\gamma(L \overset{\gamma}{\blacklozenge} g)}\\ &= \operatorname{Id}_{\mathcal{H}} - L^* \circ (\operatorname{Id}_{\mathcal{G}} - \operatorname{prox}_{\gamma g}) \circ L\\ &= \operatorname{Id}_{\mathcal{H}} - \sum_{k=1}^p \alpha_k L_k^* \circ (\operatorname{Id}_{\mathcal{G}_k} - \operatorname{prox}_{\gamma g_k}) \circ L_k. \end{aligned} \tag{5.33}$$

(iii)(a): In this case, $\|L\|^2 \leqslant \sum_{k=1}^p \alpha_k \|L_k\|^2 < 1$. Therefore, the assertion follows from (5.32) and [17, Proposition 3.2(iv)(a)].

(iii)(b): Since $\operatorname{dom} g = \times_{k=1}^p \operatorname{dom} g_k = \times_{k=1}^p \mathcal{G}_k = \mathcal{G}$, the assertion follows from (5.32) and [17, Proposition 3.2(iv)(b)].

(iii)(c): According to [17, Proposition 3.2(iv)(b)], $\operatorname{dom}(L_\ell \overset{\gamma}{\blacklozenge} g_\ell) = \mathcal{H}$. Hence, we derive from Proposition 5.9(ii) and [15, Remark 5.11(v)] that

$$\operatorname{dom} \operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p} = \operatorname{dom} \operatorname{pav}_\gamma \big(L_k \overset{\gamma}{\blacklozenge} g_k\big)_{1\leqslant k\leqslant p} = \sum_{k=1}^p \alpha_k \operatorname{dom}\big(L_k \overset{\gamma}{\blacklozenge} g_k\big) = \mathcal{H}. \tag{5.34}$$

(iv): Set $\varphi = \sum_{k=1}^p \alpha_k (\operatorname{lenv}_\gamma g_k) \circ L_k$. As seen after (5.27), $\nabla\varphi$ is γ^{-1}-Lipschitzian. On the other hand, Proposition 5.9(i) asserts that $\operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p} = \operatorname{uenv}_\gamma \varphi$. Altogether, appealing to the equivalence (i)⇔(vi) in Lemma 5.6, we conclude that $\operatorname{lenv}_\gamma \operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p} = \operatorname{lenv}_\gamma (\operatorname{uenv}_\gamma \varphi) = \varphi$.

(v): The result follows from (i), (iv), and the fact that the set of minimizers of a function in $\Gamma_0(\mathcal{H})$ coincides with that of its lower Moreau envelope [5, Proposition 17.5].

(vi): Since $\bigcap_{k=1}^p L_k^{-1}(\operatorname{Argmin} g_k) \neq \varnothing$, [5, Proposition 17.5] and (v) imply that

$$\begin{aligned}\operatorname{Argmin} \operatorname{cav}(g_k, L_k)_{1\leqslant k\leqslant p} &= \bigcap_{k=1}^p L_k^{-1}(\operatorname{Argmin} g_k)\\ &= \bigcap_{k=1}^p L_k^{-1}(\operatorname{Argmin} \operatorname{lenv}_\gamma g_k)\\ &= \operatorname{Argmin} \sum_{k=1}^p \alpha_k (\operatorname{lenv}_\gamma g_k) \circ L_k\\ &= \operatorname{Argmin} \operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p}, \end{aligned} \tag{5.35}$$

which completes the proof. ☐

Remark 5.11 Let us make a couple of observations about Proposition 5.10.

(i) Suppose that, in Assumption 5.1, $f = h = 0$ and, for every $k \in \{1, \dots, p\}$, $g_k = \iota_{D_k}$ for some nonempty closed convex set $D_k \subset \mathcal{G}_k$. Then Problem 5.2 amounts to finding $x \in \mathcal{H}$ such that, for every $k \in \{1, \dots, p\}$, $L_k x \in D_k$. On the other hand, it follows from Proposition 5.10(v) and [5, Example 12.21] that Problem 5.3 amounts to finding a minimizer of the least-squares function $x \mapsto \sum_{k=1}^{p} \alpha_k d_{D_k}^2(L_k x)$, which has been used in the relaxation of inconsistent feasibility problems [9, 14, 18].

(ii) In general, the solution sets of Problems 5.2 and 5.3 are distinct. However, it follows from Proposition 5.10(vi) that, if the set Z in (5.6) is nonempty, then it coincides with the solution set of both problems.

Next, we compare the standard composite average of $\operatorname{cav}(g_k, L_k)_{1 \leqslant k \leqslant p}$ of (5.22) with the proximal mixture $\operatorname{pcm}_\gamma(g_k, L_k)_{1 \leqslant k \leqslant p}$ of (5.23). Let us start with some basic inequalities.

Proposition 5.12 *Let* $\gamma \in \left]0, +\infty\right[$ *and*

$$\varphi\colon \mathcal{H} \to \left]-\infty, +\infty\right]\colon x \mapsto \inf_{\substack{1 \leqslant k \leqslant p \\ y_k^* \in \partial g_k(L_k x)}} \left(\sum_{k=1}^{p} \alpha_k \mathscr{Q}_{\mathcal{G}_k}(y_k^*) - \mathscr{Q}_{\mathcal{H}}\left(\sum_{k=1}^{p} \alpha_k L_k^* y_k^* \right) \right). \tag{5.36}$$

Then $\sum_{k=1}^{p} \alpha_k (\operatorname{lenv}_\gamma g_k) \circ L_k \leqslant \operatorname{pcm}_\gamma(g_k, L_k)_{1 \leqslant k \leqslant p} \leqslant \operatorname{cav}(g_k, L_k)_{1 \leqslant k \leqslant p} \leqslant \operatorname{pcm}_\gamma(g_k, L_k)_{1 \leqslant k \leqslant p} + \gamma\varphi$.

Proof. Define $\mathcal{G}$, L, and g as in (5.26), and recall from Proposition 5.9(iii) that $\operatorname{pcm}_\gamma(g_k, L_k)_{1 \leqslant k \leqslant p} = L \overset{\gamma}{\blacklozenge} g$. Further, $g \circ L = \operatorname{cav}(g_k, L_k)_{1 \leqslant k \leqslant p}$, and as seen after (5.31), $\|L\| \leqslant 1$. It follows from [17, Proposition 3.20(ii)] that

$$\sum_{k=1}^{p} \alpha_k (\operatorname{lenv}_\gamma g_k) \circ L_k = (\operatorname{lenv}_\gamma g) \circ L \leqslant L \overset{\gamma}{\blacklozenge} g \leqslant g \circ L, \tag{5.37}$$

where $L \overset{\gamma}{\blacklozenge} g = \operatorname{pcm}_\gamma(g_k, L_k)_{1 \leqslant k \leqslant p}$ and $g \circ L = \operatorname{cav}(g_k, L_k)_{1 \leqslant k \leqslant p}$. It remains to show the rightmost inequality. Let $x \in \mathcal{H}$. The result is clear if $x \notin \operatorname{dom} \varphi$. Now, suppose that $x \in \operatorname{dom} \varphi$ and let $\boldsymbol{y}^* = (y_k^*)_{1 \leqslant k \leqslant p} \in \times_{k=1}^{p} \partial g_k(L_k x)$. Then (5.25)–(5.26) yield $L^* \boldsymbol{y}^* = \sum_{k=1}^{p} \alpha_k L_k^* y_k^*$ and $\mathscr{Q}_{\mathcal{G}}(\boldsymbol{y}^*) = \sum_{k=1}^{p} \alpha_k \mathscr{Q}_{\mathcal{G}_k}(y_k^*)$. On the other hand, by [5, Proposition 16.9], $\boldsymbol{y}^* \in \times_{k=1}^{p} \partial g_k(L_k x) = \partial g(Lx)$ and we therefore deduce from [17, Proposition 3.23(i)] that

$$0 \leqslant \big(\operatorname{cav}(g_k, L_k)_{1 \leqslant k \leqslant p}\big)(x) - \big(\operatorname{pcm}_\gamma(g_k, L_k)_{1 \leqslant k \leqslant p}\big)(x) \leqslant \gamma\big(\mathscr{Q}_{\mathcal{G}}(\boldsymbol{y}^*) - \mathscr{Q}_{\mathcal{H}}(L^* \boldsymbol{y}^*)\big). \tag{5.38}$$

Hence, taking the infimum over $\boldsymbol{y}^* \in \times_{k=1}^{p} \partial g_k(L_k x)$ yields the claim. ☐

The following result establishes asymptotic relations between standard composite averages and proximal comixtures.

Theorem 5.13 *Let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $]0,+\infty[$ such that $\gamma_n \downarrow 0$ and let $x \in \mathcal{H}$. For brevity, write $\mathfrak{c} = \operatorname{cav}(g_k, L_k)_{1\leqslant k\leqslant p}$ and $(\forall n \in \mathbb{N})$ $\mathfrak{m}_{\gamma_n} = \operatorname{pcm}_{\gamma_n}(g_k, L_k)_{1\leqslant k\leqslant p}$. Then the following hold:*

(i) $\mathfrak{m}_{\gamma_n}(x) \uparrow \mathfrak{c}(x)$.

(ii) *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that $x_n \rightharpoonup x$. Then $\mathfrak{c}(x) \leqslant \underline{\lim}\, \mathfrak{m}_{\gamma_n}(x_n)$.*

(iii) *Suppose that $\mathfrak{c}$ is proper and let $\gamma \in]0,+\infty[$. Then $\operatorname{prox}_{\gamma\mathfrak{m}_{\gamma_n}} x \to \operatorname{prox}_{\gamma\mathfrak{c}} x$.*

(iv) *Let $(\varepsilon_n)_{n\in\mathbb{N}}$ be a sequence in $]0,+\infty[$ such that $\varepsilon_n \downarrow 0$, and suppose that there exists a bounded sequence $(z_n)_{n\in\mathbb{N}}$ such that $(\forall n \in \mathbb{N})$ $z_n \in \varepsilon_n$–$\operatorname{Argmin}(f + \mathfrak{m}_{\gamma_n} + h)$. Then the following hold:*

   (a) $\inf_{x\in\mathcal{H}}(f(x) + \mathfrak{m}_{\gamma_n}(x) + h(x)) \uparrow \min_{x\in\mathcal{H}}(f(x) + \mathfrak{c}(x) + h(x))$.

   (b) *Every weak sequential cluster point of $(z_n)_{n\in\mathbb{N}}$ is in $\operatorname{Argmin}(f + \mathfrak{c} + h)$.*

Proof. (i): Define L and g as in (5.26), and recall from Proposition 5.9(iii) that $(\forall n \in \mathbb{N})$ $\mathfrak{m}_{\gamma_n} = L \overset{\gamma_n}{\blacklozenge} g$. By items (ii) and (iv) in [17, Theorem 3.29], the function $\mathbb{N} \to]-\infty,+\infty] \colon n \mapsto \mathfrak{m}_{\gamma_n}(x)$ is increasing and $\lim_{n\to+\infty} \mathfrak{m}_{\gamma_n}(x) = g(Lx) = \mathfrak{c}(x)$.

(ii): The weak continuity of bounded linear operators [5, Lemma 2.41] yields $(\forall k \in \{1,\ldots,p\})$ $L_k x_n \rightharpoonup L_k x$. Thus, [36, Proposition 2.2(d)] implies that

$$(\forall k \in \{1,\ldots,p\}) \quad g_k(L_k x) \leqslant \underline{\lim}(\operatorname{lenv}_{\gamma_n} g_k)(L_k x_n). \tag{5.39}$$

On the other hand, recall that $\mathfrak{c} = \operatorname{cav}(g_k, L_k)_{1\leqslant k\leqslant p}$. Therefore, it follows from (5.39), [5, Lemma 1.16], and Proposition 5.12 that

$$\mathfrak{c}(x) \leqslant \sum_{k=1}^{p} \alpha_k \underline{\lim}(\operatorname{lenv}_{\gamma_n} g_k)(L_k x_n) \leqslant \underline{\lim} \sum_{k=1}^{p} \alpha_k (\operatorname{lenv}_{\gamma_n} g_k)(L_k x_n) \leqslant \underline{\lim}\, \mathfrak{m}_{\gamma_n}(x_n). \tag{5.40}$$

(iii): Recall from Proposition 5.10(i) that $(\forall n \in \mathbb{N})$ $\mathfrak{m}_{\gamma_n} \in \Gamma_0(\mathcal{H})$. Therefore, the result follows from (i), (ii), the equivalence (i)⇔(iii) in [3, Proposition 3.19], and the equivalence (i)⇔(ii) in [3, Theorem 3.26].

(iv): By Proposition 5.10(i) and (i), $(\mathfrak{m}_{\gamma_n})_{n\in\mathbb{N}}$ is an increasing sequence of functions in $\Gamma_0(\mathcal{H})$ with $\sup_{n\in\mathbb{N}} \mathfrak{m}_{\gamma_n} = \mathfrak{c}$. Therefore, the equivalence (iii)⇔(iv) in [5, Theorem 9.1] implies that $(f + \mathfrak{m}_{\gamma_n} + h)_{n\in\mathbb{N}}$ is an increasing sequence of weakly lower semicontinuous convex functions with supremum $f+\mathfrak{c}+h$. Since $(z_n)_{n\in\mathbb{N}}$ is bounded, [5, Lemma 2.45] asserts that $(z_n)_{n\in\mathbb{N}}$ admits a weakly convergent subsequence. Hence, the results follow from [3, Proposition 2.42] applied to the weak topology. ☐

The following result focuses on the case in which the functions $(g_k)_{1\leqslant k\leqslant p}$ are Lipschitzian.

Proposition 5.14 *Assume that, for every $k \in \{1,\ldots,p\}$, g_k is real-valued and μ_k-Lipschitzian for some $\mu_k \in]0,+\infty[$. Set $\mu = \sum_{k=1}^{p} \alpha_k \mu_k \|L_k\|$, set $\vartheta = (1/2)\sum_{k=1}^{p} \alpha_k \mu_k^2$, let $x \in \mathcal{H}$, and let $\gamma \in]0,+\infty[$. Then the following hold:*

(i) $\operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p}$ *is* μ*-Lipschitzian.*

(ii) $0 \leqslant \operatorname{cav}(g_k, L_k)_{1\leqslant k\leqslant p} - \operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p} \leqslant \gamma\vartheta$.

(iii) *There exists* $r \in B(0; 2\gamma\mu)$ *such that* $\operatorname{prox}_{\gamma\operatorname{pcm}_\gamma(g_k,L_k)_{1\leqslant k\leqslant p}} x = \operatorname{prox}_{\gamma\operatorname{cav}(g_k,L_k)_{1\leqslant k\leqslant p}}(x + r)$.

(iv) $\|\operatorname{prox}_{\gamma\operatorname{cav}(g_k,L_k)_{1\leqslant k\leqslant p}} x - \operatorname{prox}_{\gamma\operatorname{pcm}_\gamma(g_k,L_k)_{1\leqslant k\leqslant p}} x\| \leqslant \gamma \min\{2\mu, \sqrt{2\vartheta}\}$.

(v) *Let* $\sigma \in \left]0, +\infty\right[$ *and assume that* $h \in \Gamma_{-\sigma}(\mathcal{H})$. *Then*

$$\big\|\operatorname{argmin}\big(\operatorname{cav}(g_k, L_k)_{1\leqslant k\leqslant p} + h\big) - \operatorname{argmin}\big(\operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p} + h\big)\big\|_{\mathcal{H}} \leqslant \frac{2\mu}{\sigma}. \tag{5.41}$$

Proof. According to [5, Corollary 17.19], a lower semicontinuous convex function $\varphi\colon \mathcal{H} \to \mathbb{R}$ is μ-Lipschitzian if and only if $\operatorname{ran}\partial\varphi \subset B(0;\mu)$. Thus, the Lipschitz continuity of the functions $(g_k)_{1\leqslant k\leqslant p}$ implies that

$$(\forall k \in \{1,\ldots,p\}) \quad \operatorname{ran}\partial g_k \subset B(0;\mu_k). \tag{5.42}$$

Further, by [5, Proposition 16.27], $(\forall k \in \{1,\ldots,p\})$ $\operatorname{dom}\partial g_k = \mathcal{G}_k$.

(i): By virtue of Proposition 5.10(iii)(b), $\operatorname{dom}\operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p} = \mathcal{H}$ and therefore [5, Proposition 16.27] yields $\operatorname{dom}\partial\operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p} = \mathcal{H}$. Now, let $u \in \operatorname{ran}\partial\operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p}$. Then there exists $x \in \mathcal{H}$ such that $u \in (\partial\operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p})(x)$ which, by (5.18), is equivalent to $\operatorname{prox}_{\gamma\operatorname{pcm}_\gamma(g_k,L_k)_{1\leqslant k\leqslant p}}(x + \gamma u) = x$. Now, set $(\forall k \in \{1,\ldots,p\})$ $y_k^* = L_k(x + \gamma u) - \operatorname{prox}_{\gamma g_k}(L_k(x + \gamma u))$. By Proposition 5.10(ii),

$$x + \gamma u - \sum_{k=1}^{p} \alpha_k L_k^* y_k^* = x. \tag{5.43}$$

Further, it follows from (5.18) and (5.42) that

$$(\forall k \in \{1,\ldots,p\}) \quad \frac{1}{\gamma} y_k^* \in \partial g_k\big(L_k(x + \gamma u) - y_k^*\big) \subset \operatorname{ran}\partial g_k \subset B(0;\mu_k). \tag{5.44}$$

Therefore, it follows from (5.43) and (5.44) that

$$\|u\|_{\mathcal{H}} = \left\|\frac{1}{\gamma}\sum_{k=1}^{p}\alpha_k L_k^* y_k^*\right\|_{\mathcal{H}} \leqslant \sum_{k=1}^{p}\alpha_k\|L_k\|\left\|\frac{1}{\gamma}y_k^*\right\|_{\mathcal{G}_k} \leqslant \sum_{k=1}^{p}\alpha_k\|L_k\|\mu_k = \mu. \tag{5.45}$$

Altogether, $\operatorname{ran}\partial\operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p} \subset B(0;\mu)$, which yields the conclusion via [5, Corollary 17.19].

(ii): Recall that $(\forall k \in \{1,\ldots,p\})$ $\operatorname{dom}\partial g_k = \mathcal{G}_k$. For every $k \in \{1,\ldots,p\}$, let $y_k^* \in \partial g_k(L_k x)$. Thus, it follows from Proposition 5.12 and (5.42) that

$$0 \leqslant \big(\operatorname{cav}(g_k, L_k)_{1\leqslant k\leqslant p}\big)(x) - \big(\operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p}\big)(x) \leqslant \gamma\sum_{k=1}^{p}\alpha_k \mathscr{Q}_{\mathcal{G}_k}(y_k^*) \leqslant \frac{\gamma}{2}\sum_{k=1}^{p}\alpha_k\mu_k^2 = \gamma\vartheta. \tag{5.46}$$

(iii): Set $u_\gamma = \operatorname{prox}_{\gamma \operatorname{pcm}_\gamma(g_k,L_k)_{1\leqslant k\leqslant p}} x$. Then, by (5.18), $x-u_\gamma \in \gamma(\partial \operatorname{pcm}_\gamma(g_k,L_k)_{1\leqslant k\leqslant p})(u_\gamma)$. On the other hand, since $\operatorname{dom}\partial \operatorname{cav}(g_k,L_k)_{1\leqslant k\leqslant p} = \mathcal{H}$, there exists $r\in\mathcal{H}$ such that $x+r-u_\gamma \in \gamma(\partial \operatorname{cav}(g_k,L_k)_{1\leqslant k\leqslant p})(u_\gamma)$. Thus, by (5.18), $\operatorname{prox}_{\gamma \operatorname{cav}(g_k,L_k)_{1\leqslant k\leqslant p}}(x+r) = u_\gamma$. Further, by (i) and [5, Corollary 17.19],

$$\operatorname{ran}\partial \operatorname{cav}(g_k,L_k)_{1\leqslant k\leqslant p} \subset B(0;\mu) \quad \text{and} \quad \operatorname{ran}\partial \operatorname{pcm}_\gamma(g_k,L_k)_{1\leqslant k\leqslant p} \subset B(0;\mu). \tag{5.47}$$

Therefore, $r = (x+r-u_\gamma)-(x-u_\gamma) \in \gamma B(0;\mu)+\gamma B(0;\mu) = B(0;2\gamma\mu)$.

(iv): Set $u = \operatorname{prox}_{\gamma \operatorname{cav}(g_k,L_k)_{1\leqslant k\leqslant p}} x$, $u_\gamma = \operatorname{prox}_{\gamma \operatorname{pcm}_\gamma(g_k,L_k)_{1\leqslant k\leqslant p}} x$, and

$$(\forall k\in\{1,\ldots,p\})\quad y_k^* = L_k x - \operatorname{prox}_{\gamma g_k}(L_k x). \tag{5.48}$$

Further, set $z = \sum_{k=1}^p \alpha_k L_k^* y_k^*$. Then, Proposition 5.10(ii) yields $x-z = u_\gamma$, whereas (5.18), (5.48), and (5.42) imply that

$$(\forall k\in\{1,\ldots,p\})\quad y_k^* \in \gamma\partial g_k(L_k x - y_k^*) \subset B(0;\gamma\mu_k). \tag{5.49}$$

By [5, Corollaries 16.48(iii) and 16.53(i)], $x-u \in \gamma(\partial \operatorname{cav}(g_k,L_k)_{1\leqslant k\leqslant p})(u) = \gamma\sum_{k=1}^p \alpha_k L_k^*(\partial g_k(L_k u))$. Thus, for every $k\in\{1,\ldots,p\}$, there exists $w_k^* \in \gamma\partial g_k(L_k u)$ such that $x-u = \sum_{k=1}^p \alpha_k L_k^* w_k^*$. It follows from (5.49) and the monotonicity of the subdifferential operators $(\gamma\partial g_k)_{1\leqslant k\leqslant p}$ [5, Theorem 20.25] that

$$(\forall k\in\{1,\ldots,p\})\quad \langle L_k x - y_k^* - L_k u \mid y_k^* - w_k^*\rangle_{\mathcal{G}_k} \geqslant 0. \tag{5.50}$$

Since $\sum_{k=1}^p \alpha_k\|L_k\|^2 \leqslant 1$ by Assumption 5.1, (5.49) implies that

$$\sum_{k=1}^p \alpha_k\|y_k^* - L_k z\|_{\mathcal{G}_k}^2 = \sum_{k=1}^p \alpha_k\|y_k^*\|_{\mathcal{G}_k}^2 - 2\|z\|_{\mathcal{H}}^2 + \sum_{k=1}^p \alpha_k\|L_k z\|_{\mathcal{G}_k}^2 \leqslant \sum_{k=1}^p \alpha_k\|y_k^*\|_{\mathcal{G}_k}^2 \leqslant 2\gamma^2\vartheta, \tag{5.51}$$

whereas the Cauchy–Schwarz inequality yields

$$\|z\|_{\mathcal{H}}^2 \leqslant \left(\sum_{k=1}^p \alpha_k\|L_k\|\|y_k^*\|_{\mathcal{G}_k}\right)^2 \leqslant \left(\sum_{k=1}^p \alpha_k\|L_k\|^2\right)\left(\sum_{k=1}^p \alpha_k\|y_k^*\|_{\mathcal{G}_k}^2\right) \leqslant \sum_{k=1}^p \alpha_k\|y_k^*\|_{\mathcal{G}_k}^2. \tag{5.52}$$

Altogether, we deduce from $x = z+u_\gamma$, $\sum_{k=1}^p \alpha_k L_k^* y_k^* = z$, $\sum_{k=1}^p \alpha_k L_k^* w_k^* = x-u$, (5.50), the Cauchy–Schwarz inequality, (5.51), and (5.52) that

$$0 \leqslant \sum_{k=1}^p \alpha_k\langle L_k u_\gamma - L_k u + L_k z - y_k^* \mid y_k^* - w_k^*\rangle_{\mathcal{G}_k}$$

$$
\begin{aligned}
&= \langle u_\gamma - u \,|\, (x-u_\gamma)-(x-u)\rangle_{\mathcal{H}} + \sum_{k=1}^{p} \alpha_k \langle L_k z - y_k^* \,|\, y_k^* - w_k^*\rangle_{\mathcal{G}_k} \\
&= -\|u-u_\gamma\|_{\mathcal{H}}^2 + \sum_{k=1}^{p} \alpha_k \langle y_k^* - L_k z \,|\, w_k^*\rangle_{\mathcal{G}_k} - \sum_{k=1}^{p} \alpha_k \langle y_k^* - L_k z \,|\, y_k^*\rangle_{\mathcal{G}_k} \\
&\leqslant -\|u-u_\gamma\|_{\mathcal{H}}^2 + \sqrt{\sum_{k=1}^{p} \alpha_k \|y_k^* - L_k z\|_{\mathcal{G}_k}^2}\sqrt{\sum_{k=1}^{p} \alpha_k \|w_k^*\|_{\mathcal{G}_k}^2} + \left(\|z\|_{\mathcal{H}}^2 - \sum_{k=1}^{p} \alpha_k \|y_k^*\|_{\mathcal{G}_k}^2\right) \\
&\leqslant -\|u-u_\gamma\|_{\mathcal{H}}^2 + \sqrt{2\gamma^2\vartheta}\sqrt{\gamma^2 \sum_{k=1}^{p} \alpha_k \mu_k^2} \\
&= -\|u-u_\gamma\|_{\mathcal{H}}^2 + 2\gamma^2\vartheta.
\end{aligned} \tag{5.53}
$$

However, by (iii), there exists $r \in B(0;2\gamma\mu)$ such that $\operatorname{prox}_{\gamma \operatorname{pcm}_\gamma(g_k,L_k)_{1\leqslant k\leqslant p}} x = \operatorname{prox}_{\gamma \operatorname{cav}(g_k,L_k)_{1\leqslant k\leqslant p}}(x+r)$. Hence, the nonexpansiveness of the proximity operator [5, Proposition 12.28] implies that

$$
\|u-u_\gamma\|_{\mathcal{H}} = \big\|\operatorname{prox}_{\gamma \operatorname{cav}(g_k,L_k)_{1\leqslant k\leqslant p}} x - \operatorname{prox}_{\gamma \operatorname{cav}(g_k,L_k)_{1\leqslant k\leqslant p}}(x+r)\big\|_{\mathcal{H}} \leqslant \|r\|_{\mathcal{H}} \leqslant 2\gamma\mu. \tag{5.54}
$$

Finally, the result follows from (5.53) and (5.54).

(v): Set $x = \operatorname{argmin}(\operatorname{cav}(g_k,L_k)_{1\leqslant k\leqslant p} + h)$ and $x_\gamma = \operatorname{argmin}(\operatorname{pcm}_\gamma(g_k,L_k)_{1\leqslant k\leqslant p} + h)$. Then the equivalence (i)⇔(viii) in [5, Corollary 27.3] yields

$$
x = \operatorname{prox}_{\gamma \operatorname{cav}(g_k,L_k)_{1\leqslant k\leqslant p}}\big(x-\gamma\nabla h(x)\big) \quad \text{and} \quad x_\gamma = \operatorname{prox}_{\gamma \operatorname{pcm}_\gamma(g_k,L_k)_{1\leqslant k\leqslant p}}\big(x_\gamma-\gamma\nabla h(x_\gamma)\big). \tag{5.55}
$$

On the other hand, (iii) asserts that there exists $r \in B(0;2\gamma\mu)$ such that

$$
\operatorname{prox}_{\gamma \operatorname{pcm}_\gamma(g_k,L_k)_{1\leqslant k\leqslant p}}\big(x-\gamma\nabla h(x)\big) = \operatorname{prox}_{\gamma \operatorname{cav}(g_k,L_k)_{1\leqslant k\leqslant p}}\big(x-\gamma\nabla h(x)+r\big). \tag{5.56}
$$

Altogether, (5.55), (5.56), and the firm nonexpansiveness of $\operatorname{prox}_{\gamma \operatorname{cav}(g_k,L_k)_{1\leqslant k\leqslant p}}$ [5, Proposition 12.28] yield

$$
\begin{aligned}
\|x-x_\gamma\|_{\mathcal{H}}^2 &= \Big\|\operatorname{prox}_{\gamma \operatorname{cav}(g_k,L_k)_{1\leqslant k\leqslant p}}\big(x-\gamma\nabla h(x)\big) - \operatorname{prox}_{\gamma \operatorname{pcm}_\gamma(g_k,L_k)_{1\leqslant k\leqslant p}}\big(x_\gamma-\gamma\nabla h(x_\gamma)\big)\Big\|_{\mathcal{H}}^2 \\
&= \Big\|\operatorname{prox}_{\gamma \operatorname{cav}(g_k,L_k)_{1\leqslant k\leqslant p}}\big(x-\gamma\nabla h(x)\big) - \operatorname{prox}_{\gamma \operatorname{cav}(g_k,L_k)_{1\leqslant k\leqslant p}}\big(x_\gamma-\gamma\nabla h(x_\gamma)+r\big)\Big\|_{\mathcal{H}}^2 \\
&\leqslant \langle x-\gamma\nabla h(x) - x_\gamma + \gamma\nabla h(x_\gamma) - r \,|\, x-x_\gamma\rangle_{\mathcal{H}} \\
&= \|x-x_\gamma\|_{\mathcal{H}}^2 - \gamma\langle \nabla h(x) - \nabla h(x_\gamma) \,|\, x-x_\gamma\rangle_{\mathcal{H}} - \langle r \,|\, x-x_\gamma\rangle_{\mathcal{H}}.
\end{aligned} \tag{5.57}
$$

Therefore, it follows from [5, Example 22.4(iv)], the Cauchy–Schwarz inequality, and the inequality $\|r\|_{\mathcal{H}} \leqslant 2\gamma\mu$ that

$$\gamma\sigma\|x - x_\gamma\|_{\mathcal{H}}^2 \leqslant \gamma\langle\nabla h(x) - \nabla h(x_\gamma) \mid x - x_\gamma\rangle_{\mathcal{H}} \leqslant 2\gamma\mu\|x - x_\gamma\|_{\mathcal{H}}, \tag{5.58}$$

which implies (5.41). ☐

As a special case of the above results, we recover some properties of the proximal average [6, 42, 45].

Remark 5.15 (proximal average) Suppose that $\sum_{k=1}^p \alpha_k = 1$ and that, for every $k \in \{1, \ldots, p\}$, $\mathcal{G}_k = \mathcal{H}$ and $L_k = \mathrm{Id}_{\mathcal{H}}$. Then $\mathrm{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p} = \mathrm{pav}_\gamma(g_k)_{1\leqslant k\leqslant p}$ is the proximal average of (5.9), while $\mathrm{cav}(g_k, L_k)_{1\leqslant k\leqslant p} = \mathrm{ave}(g_k)_{1\leqslant k\leqslant p}$ is the standard average of (5.11). In this context, we recover the following results:

(i) Proposition 5.10(i) yields $\mathrm{pav}_\gamma(g_k)_{1\leqslant k\leqslant p} \in \Gamma_0(\mathcal{H})$ (see [6, Corollary 5.2]).

(ii) Proposition 5.10(ii) yields $\mathrm{prox}_{\gamma \mathrm{pav}_\gamma(g_k)_{1\leqslant k\leqslant p}} = \sum_{k=1}^p \alpha_k \,\mathrm{prox}_{\gamma g_k}$ (see [6, Theorem 6.7]).

(iii) Proposition 5.10(iv) yields $\mathrm{lenv}_\gamma\, \mathrm{pav}_\gamma(g_k)_{1\leqslant k\leqslant p} = \sum_{k=1}^p \alpha_k \mathrm{lenv}_\gamma\, g_k$ (see [6, Theorem 6.2(i)]).

(iv) Proposition 5.12 yields $\mathrm{pav}_\gamma(g_k)_{1\leqslant k\leqslant p} \leqslant \mathrm{ave}$ (see [6, Theorem 5.4]).

(v) Let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $]0, +\infty[$ such that $\gamma_n \downarrow 0$.

   (a) Theorem 5.13(i) yields $\mathrm{pav}_{\gamma_n}(g_k)_{1\leqslant k\leqslant p} \uparrow \mathrm{ave}(g_k)_{1\leqslant k\leqslant p}$ (see [6, Theorem 8.5]).

   (b) Theorem 5.13(i)–(ii) imply that $(\mathrm{pav}_{\gamma_n}(g_k)_{1\leqslant k\leqslant p})_{n\in\mathbb{N}}$ epi-converges to $\mathrm{ave}(g_k)_{1\leqslant k\leqslant p}$ (see [6, Corollary 9.6]).

(vi) Suppose that, for every $k \in \{1, \ldots, p\}$, g_k is real-valued and μ_k-Lipschitzian for some $\mu_k \in]0, +\infty[$. Set $\mu = \sum_{k=1}^p \alpha_k\mu_k$ and $\vartheta = (1/2)\sum_{k=1}^p \alpha_k\mu_k^2$.

   (a) Proposition 5.14(ii) yields $0 \leqslant \mathrm{ave}(g_k)_{1\leqslant k\leqslant p} - \mathrm{pav}_\gamma(g_k)_{1\leqslant k\leqslant p} \leqslant \gamma\vartheta$ (see [45, Proposition 4]).

   (b) Proposition 5.14(iv) yields $\|\,\mathrm{prox}_{\gamma \mathrm{pav}_\gamma(g_k)_{1\leqslant k\leqslant p}} - \mathrm{prox}_{\gamma \mathrm{ave}(g_k)_{1\leqslant k\leqslant p}}\|_{\mathcal{H}} \leqslant 2\gamma\mu$ (see [42, p. 851]).

In addition, we obtain the following new properties of the proximal average.

(vii) Let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $]0, +\infty[$ such that $\gamma_n \downarrow 0$. It follows from Theorem 5.13(iii) that $\mathrm{prox}_{\gamma \mathrm{pav}_{\gamma_n}(g_k)_{1\leqslant k\leqslant p}} x \to \mathrm{prox}_{\gamma \mathrm{ave}(g_k)_{1\leqslant k\leqslant p}} x$.

(viii) Suppose that, for every $k \in \{1, \ldots, p\}$, g_k is real-valued and μ_k-Lipschitzian for some $\mu_k \in]0, +\infty[$. Set $\mu = \sum_{k=1}^p \alpha_k\mu_k$. Then, by Proposition 5.14(iii), there exists $r \in B(0; 2\gamma\mu)$ such that $\mathrm{prox}_{\gamma \mathrm{pav}_\gamma(g_k)_{1\leqslant k\leqslant p}} x = \mathrm{prox}_{\gamma \mathrm{ave}(g_k)_{1\leqslant k\leqslant p}}(x + r)$.

Generalization of Problems 5.2 and 5.3 to arbitrary families of functions and linear operators can be formulated as follows.

Remark 5.16 (integral proximal comixture) Let $(\Omega, \mathcal{F}, \mu)$ be a complete σ-finite measure space and let H be a separable real Hilbert space. For every $\omega \in \Omega$, let G_ω be a real Hilbert space, let $\mathsf{g}_\omega \in \Gamma_0(\mathsf{G}_\omega)$, and let $\mathsf{L}_\omega \colon \mathsf{H} \to \mathsf{G}_\omega$ be a bounded linear operator. Assume that $0 < \int_\Omega \|\mathsf{L}_\omega\|^2 \mu(d\omega) \leqslant 1$. Under mild assumptions, the *integral proximal comixture* of $(\mathsf{g}_\omega)_{\omega\in\Omega}$ and $(\mathsf{L}_\omega)_{\omega\in\Omega}$ with parameter $\gamma \in \left]0, +\infty\right[$ is [8, Definition 4.2]

$$\overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{g}_\omega, \mathsf{L}_\omega)_{\omega\in\Omega} = \Big(\varphi^* - \frac{\gamma}{2}\|\cdot\|_{\mathsf{H}}^2\Big)^*, \quad \text{where} \quad (\forall \mathsf{x} \in \mathsf{H}) \quad \varphi(\mathsf{x}) = \int_\Omega (\operatorname{lenv}_\gamma \mathsf{g}_\omega)(\mathsf{L}_\omega \mathsf{x})\mu(d\omega). \tag{5.59}$$

Now let $\mathsf{f} \in \Gamma_0(\mathsf{H})$ and let $\mathsf{h} \colon \mathsf{H} \to \mathbb{R}$ be a convex and differentiable function with Lipschitzian gradient. Then a generalization of Problem 5.2 is

$$\underset{\mathsf{x}\in\mathsf{H}}{\text{minimize}} \;\; \mathsf{f}(\mathsf{x}) + \int_\Omega \mathsf{g}_\omega(\mathsf{L}_\omega \mathsf{x})\,\mu(d\omega) + \mathsf{h}(\mathsf{x}) \tag{5.60}$$

and a generalization of Problem 5.3 is

$$\underset{\mathsf{x}\in\mathsf{H}}{\text{minimize}} \;\; \mathsf{f}(\mathsf{x}) + \Big(\overset{\blacklozenge}{\mathsf{M}}_\gamma(\mathsf{g}_\omega, \mathsf{L}_\omega)_{\omega\in\Omega}\Big)(\mathsf{x}) + \mathsf{h}(\mathsf{x}). \tag{5.61}$$

Indeed, we recover (5.1) from (5.60) and (5.3) from (5.61) by taking $\Omega = \{1, \ldots, p\}$ and $\mathcal{F} = 2^\Omega$, and setting $(\forall k \in \{1, \ldots, p\})$ $\mu(\{k\}) = \alpha_k$. Most of the results of this section extend to this abstract framework using the tools of [8, 17]. If μ is a probability measure, we can regard (5.1) and (5.3) as empirical versions of (5.60) and (5.61), respectively.

5.2.4 Algorithms

Several splitting methods are available to solve Problems 5.2 and 5.3 [16]. Regarding Problem 5.2, we have compared several methods from [16, 20, 23, 44] and found that, overall, the following method based on [16, Example 8.50] performed the best in our experiments.

Proposition 5.17 *Consider the setting of Problem 5.2 and assume that*

$$0 \in \operatorname{ran}\Big(\partial f + \sum_{k=1}^{p} \alpha_k L_k^* \circ (\partial g_k) \circ L_k + \nabla h\Big). \tag{5.62}$$

Let $x_0 \in \mathcal{H}$ and, for every $k \in \{1, \ldots, p\}$, let $y_{k,0} \in \mathcal{G}_k$ and $v_{k,0}^* \in \mathcal{G}_k$. Set $\chi = 4\beta/(1 +$

$\sqrt{1+32\beta^2})$, *let* $\varepsilon \in \left]0, \chi/(\chi+1)\right[$, *and let* $(\eta_n)_{n\in\mathbb{N}}$ *be a sequence in* $[\varepsilon, (1-\varepsilon)\chi]$. *Iterate*

$$
\begin{array}{l}
\text{for } n = 0,1,\ldots\\
\left\lfloor
\begin{array}{l}
a_n = \operatorname{prox}_{\eta_n f}\Big(x_n + \eta_n\big(\sum_{k=1}^p \alpha_k L_k^* v_{k,n}^* - \nabla h(x_n)\big)\Big)\\
\text{for } k = 1,\ldots,p\\
\left\lfloor q_{k,n}^* = \eta_n(y_{k,n} - L_k x_n)\right.\\
x_{n+1} = a_n + \eta_n \sum_{k=1}^p \alpha_k L_k^* q_{k,n}^*\\
\text{for } k = 1,\ldots,p\\
\left\lfloor
\begin{array}{l}
b_{k,n} = \operatorname{prox}_{\eta_n g_k}\big(y_{k,n} + \eta_n v_{k,n}^*\big)\\
y_{k,n+1} = b_{k,n} - \eta_n q_{k,n}^*\\
v_{k,n+1}^* = v_{k,n}^* + \eta_n\big(L_k a_n - b_{k,n}\big).
\end{array}
\right.
\end{array}
\right.
\end{array}
\tag{5.63}
$$

Then $(x_n)_{n\in\mathbb{N}}$ *converges weakly to a solution to Problem* 5.2.

Proof. Apply [16, Example 8.50] in the Hilbert space $\mathcal{G}$ of (5.25), with g and L defined as in (5.26). ☐

We now turn to Problem 5.3, for which we found the following version of the method proposed in [25, Section 3.1] to perform best overall. Given $y_0 \in \mathcal{H}$, this algorithm iterates

$$
\begin{array}{l}
\text{for } n = 0,1,\ldots\\
\left\lfloor
\begin{array}{l}
x_n = \operatorname{prox}_{\gamma \operatorname{pcm}_\gamma (g_k, L_k)_{1\leqslant k\leqslant p}} y_n\\
z_n = \operatorname{prox}_{\gamma f}\big(2x_n - y_n - \gamma\nabla h(x_n)\big)\\
y_{n+1} = y_n + \lambda_n(z_n - x_n).
\end{array}
\right.
\end{array}
\tag{5.64}
$$

Hence, using Proposition 5.10(ii), we arrive at the following method.

Proposition 5.18 *Consider the setting of Problem* 5.3. *Assume that* $\gamma \in \left]0, 2\beta\right[$ *and that*

$$
0 \in \operatorname{ran}\big(\partial f + \partial \operatorname{pcm}_\gamma (g_k, L_k)_{1\leqslant k\leqslant p} + \nabla h\big). \tag{5.65}
$$

Set $\delta = 2 - \gamma/(2\beta)$, *let* $(\lambda_n)_{n\in\mathbb{N}}$ *be a sequence in* $\left]0, \delta\right[$ *such that* $\sum_{n\in\mathbb{N}} \lambda_n(\delta - \lambda_n) = +\infty$, *and let* $y_0 \in \mathcal{H}$. *Iterate*

$$
\begin{array}{l}
\text{for } n = 0,1,\ldots\\
\left\lfloor
\begin{array}{l}
x_n = y_n - \sum_{k=1}^p \alpha_k L_k^*\big(L_k y_n - \operatorname{prox}_{\gamma g_k}(L_k y_n)\big)\\
z_n = \operatorname{prox}_{\gamma f}\big(2x_n - y_n - \gamma\nabla h(x_n)\big)\\
y_{n+1} = y_n + \lambda_n(z_n - x_n).
\end{array}
\right.
\end{array}
\tag{5.66}
$$

Then $(x_n)_{n\in\mathbb{N}}$ *converges weakly to a solution to Problem* 5.3.

Heuristically, one can expect (5.66) to yield fast convergence than (5.63). Indeed, since the proximity operator of the aggregation $\operatorname{pcm}_\gamma (g_k, L_k)_{1\leqslant k\leqslant p}$ is computable in closed form via

Proposition 5.10(ii), algorithm (5.64) processes only the three functions f, $\operatorname{pcm}_\gamma(g_k, L_k)_{1\leqslant k\leqslant p}$, and h. It also requires minimum variable storage. By contrast, the standard composite average function $\operatorname{cav}(g_k, L_k)_{1\leqslant k\leqslant p}$ in Problem 5.2 has no explicit proximity operator and its processing in algorithm (5.63) requires the splitting of the p functions $(g_k)_{1\leqslant k\leqslant p}$ and the p linear operator $(L_k)_{1\leqslant k\leqslant p}$ (the same holds true for all splitting methods involving standard composite averages [16]). Additionally, algorithm (5.63) requires the storage of significantly more variables than (5.66).

5.2.5 Numerical experiments

We illustrate the proposed proximal comixture model of Problem 5.3 through several experiments. The algorithms are executed with all initial vectors set to 0 and they use parameters $(\eta_n)_{n\in\mathbb{N}}$ and $(\lambda_n)_{n\in\mathbb{N}}$ aiming at optimizing their performance.

5.2.5.1 Experiment 1: Proximal total variation denoising

The purpose of this experiment is to illustrate the asymptotic property of Theorem 5.13(iii) on a simple denoising problem. Let $\overline{x} \in \mathbb{R}^N$ $(N = 256)$ be the original 1-dimensional signal shown in Fig. 5.2(a) and let

$$z = \overline{x} + \frac{1}{2}w \tag{5.67}$$

be the noisy observation shown in Fig. 5.2(b), where $w \in \mathbb{R}^N$ is a realization of a normalized white Gaussian noise vector. We denote by $D\colon \mathbb{R}^N \to \mathbb{R}^N \colon (\xi_1,\ldots,\xi_N) \mapsto (\xi_2-\xi_1,\ldots,\xi_N-\xi_{N-1},\xi_1-\xi_N)/2$ the normalized discrete gradient operator.

Problem 5.19 Let $\rho = 3/2$ and let $\operatorname{tv} = \|\cdot\|_1 \circ D$ be the standard total variation loss. The task is to

$$\underset{x\in\mathbb{R}^N}{\text{minimize}} \;\; \operatorname{tv}(x) + \frac{1}{2\rho}\|x-z\|^2. \tag{5.68}$$

Problem 5.20 Let $\rho = 3/2$ and $\gamma \in \left]0,+\infty\right[$. Let us introduce the *proximal total variation* loss $\operatorname{ptv}_\gamma = D \overset{\gamma}{\blacklozenge} \|\cdot\|_1$. The task is to

$$\underset{x\in\mathbb{R}^N}{\text{minimize}} \;\; \operatorname{ptv}_\gamma(x) + \frac{1}{2\rho}\|x-z\|^2. \tag{5.69}$$

Problems 5.19 and 5.20 are particular instances of Problems 5.2 and 5.3, respectively, where $\mathcal{H} = \mathbb{R}^N$, $f = 0$, $h = \|\cdot - z\|^2/(2\rho)$, $p = 1$, $\alpha_1 = 1$, $g_1 = \|\cdot\|_1$, and $L_1 = D$. In view of Proposition 5.10(i) and (5.17), the unique solutions to Problems 5.19 and 5.20 are, respectively, $\operatorname{prox}_{\rho\operatorname{tv}} z$ and $\operatorname{prox}_{\rho\operatorname{ptv}_\gamma} z$. Therefore, Theorem 5.13(iii) asserts that the solution curve $(\operatorname{prox}_{\rho\operatorname{ptv}_\gamma} z)_{\gamma\in\left]0,+\infty\right[}$ in Problem 5.20 converges to the solution to Problem 5.19, to wit,

$$\operatorname{prox}_{\rho\operatorname{ptv}_\gamma} z \to \operatorname{prox}_{\rho\operatorname{tv}} z \quad \text{as} \quad \gamma \downarrow 0. \tag{5.70}$$

Since the proximity operator of tv is not known explicitly, algorithm (5.63) can be applied to obtain the solution to Problem 5.19. On the other hand, algorithm (5.66) can be used to obtain the solution to Problem 5.20. Our numerical experiments confirmed that, for $\gamma \leqslant 10^{-3}$, the solutions to Problems 5.19 and 5.20 were essentially identical, in conformity with (5.70); see Fig. 5.2(c) for the denoised signal.

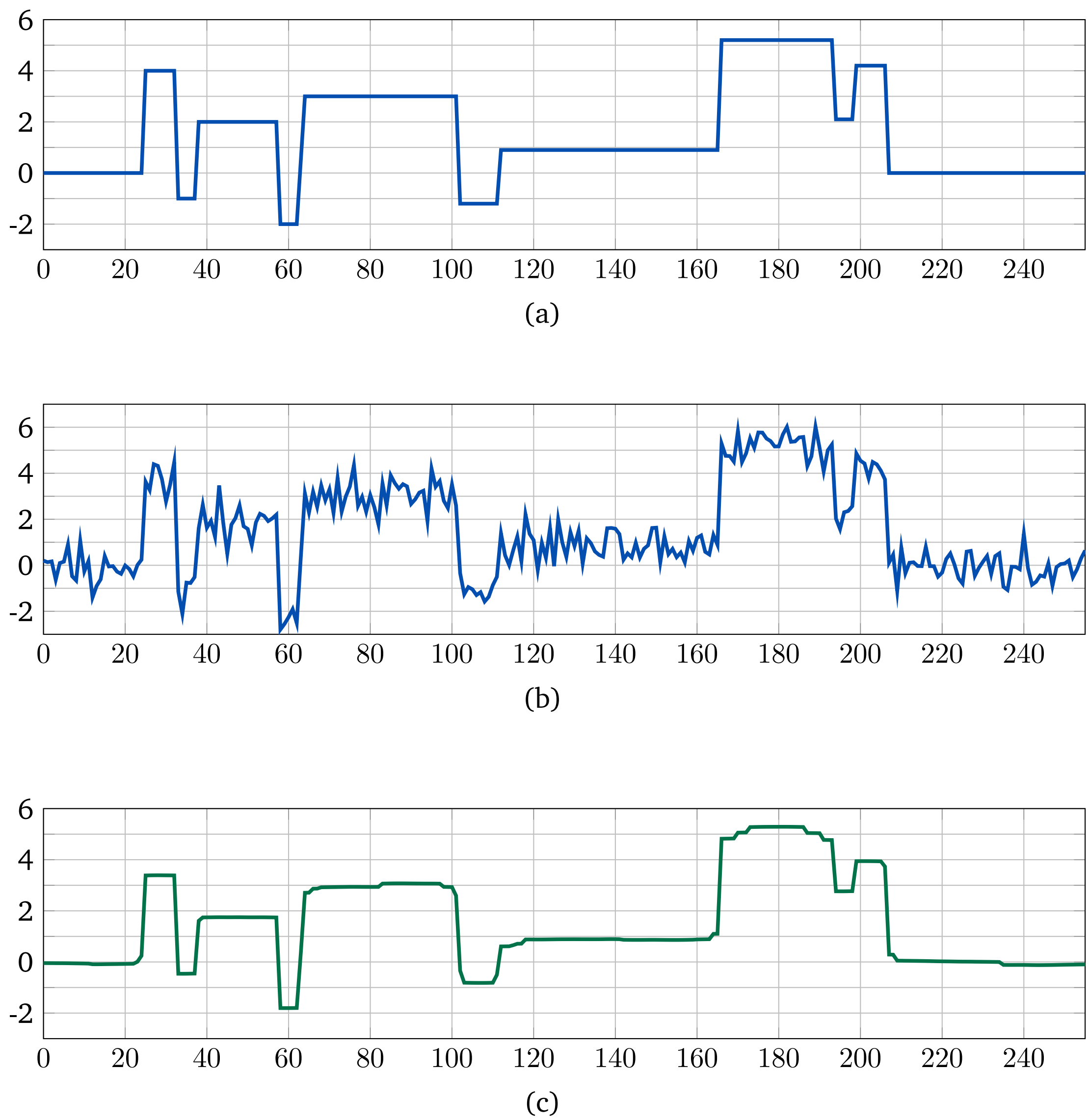


Figure 5.2 (a) Original signal $\overline{x}$. (b) Noisy observation z. (c) Solution to Problem 5.20 for $\gamma = 10^{-3}$ (the solution to Problem 5.19 is essentially identical).

5.2.5.2 Experiment 2: Multiview image reconstruction

We address the problem of reconstructing the original image $\overline{x} \in C = [0, 255]^N$ $(N = 512^2)$ of Fig. 5.3(a) from a partial observation of its possibly corrupted diffraction $r \approx \widehat{\overline{x}}$ over some frequency range R [39], where $\widehat{\overline{x}}$ denotes the two-dimensional discrete Fourier transform of $\overline{x}$. To exploit this imprecise information, we use the soft constraint distance penalty d_E, where

$$E = \big\{x \in \mathbb{R}^N \mid (\forall \nu \in R) \;\; \widehat{x}(\nu) = r(\nu)\big\}. \tag{5.71}$$

The set R contains the frequencies in $\{0, \ldots, 15\}^2$ as well as those resulting from the symmetry properties of the discrete Fourier transform. Further, two blurred noisy observations of $\overline{x}$ are available, namely (see Fig. 5.3(b)–(c))

$$z_1 = H_1\overline{x} + w_1 \quad \text{and} \quad z_2 = H_2\overline{x} + w_2, \tag{5.72}$$

where H_1 and H_2 model convolutional blurs with uniform rectangular kernels of sizes 14×18 and 20×5, respectively, while w_1 and w_2 represent zero-mean white Gaussian noise realizations of standard deviations of 2 and 3, respectively. The blurred image-to-noise ratios are 34.51 dB and 31.06 dB, respectively. Let

$$D \colon \mathbb{R}^N \to \mathbb{R}^N \times \mathbb{R}^N \colon x \mapsto (D_1x, D_2x), \tag{5.73}$$

where D_1 and D_2 denote, respectively, the horizontal and vertical first order discrete difference operators. Let $\|\cdot\|_{1,2} \colon \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R} \colon (\xi_i, \eta_i)_{1 \leqslant i \leqslant N} \mapsto \sum_{i=1}^N \sqrt{\xi_i^2 + \eta_i^2}$ and let

$$\mathfrak{h}_\rho \colon \mathbb{R} \to \mathbb{R} \colon \xi \mapsto \begin{cases} \rho|\xi| - \dfrac{\rho^2}{2}, & \text{if} \quad |\xi| > \rho; \\ \dfrac{|\xi|^2}{2}, & \text{if} \quad |\xi| \leqslant \rho \end{cases} \tag{5.74}$$

be the Huber function with parameter $\rho \in \left]0, +\infty\right[$. Our first formulation involves a standard composite average.

(a) (b)

(c)

Figure 5.3 (a) Original image $\overline{x}$. (b) Degraded image z_1. (c) Degraded image z_2.

Problem 5.21 Let $\rho_1 = 3000$ and $\rho_2 = 4000$. The task is to

$$\underset{x \in C}{\text{minimize}} \ \left(\frac{1}{2} d_E(x) + \frac{1}{2}\|Dx\|_{1,2}\right) + \left(\mathfrak{h}_{\rho_1}\big(\|H_1 x - z_1\|\big) + \mathfrak{h}_{\rho_2}\big(\|H_2 x - z_2\|\big)\right). \tag{5.75}$$

The second formulation involves a proximal comixture.

Problem 5.22 Let $\rho_1 = 3000$, $\rho_2 = 4000$, and $\gamma \in \left]0,1\right[$. The task is to

$$\underset{x\in C}{\text{minimize}} \; \Big(\text{pcm}_\gamma\big(d_E, \text{Id}; \sqrt{8}\|\cdot\|_{1,2}, D/\sqrt{8}\big)\Big)(x) + \Big(\mathfrak{h}_{\rho_1}\big(\|H_1x - z_1\|\big) + \mathfrak{h}_{\rho_2}\big(\|H_2x - z_2\|\big)\Big). \tag{5.76}$$

Problems 5.21 and 5.22 are particular instances of Problems 5.2 and 5.3, respectively, where $\mathcal{H} = \mathbb{R}^N$, $f = \iota_C$, $h = \mathfrak{h}_{\rho_1} \circ \|H_1 \cdot - z_1\| + \mathfrak{h}_{\rho_2} \circ \|H_2 \cdot - z_2\|$, $\beta = 1/2$, $p = 2$, $\alpha_1 = \alpha_2 = 1/2$, $\mathcal{G}_1 = \mathbb{R}^N$, $g_1 = d_E$, $L_1 = \text{Id}$, $\mathcal{G}_2 = \mathbb{R}^N \times \mathbb{R}^N$, $g_2 = \sqrt{8}\|\cdot\|_{1,2}$, and $L_2 = D/\sqrt{8}$. In this case, $\|L_1\|^2 = \|L_2\|^2 = 1$ and

$$\text{prox}_{\gamma f} \colon (\xi_i)_{1\leqslant i\leqslant N} \mapsto \big(\min\big\{\max\{\xi_i, 0\}, 255\big\}\big)_{1\leqslant i\leqslant N}. \tag{5.77}$$

Moreover, [5, Example 24.28] yields

$$\text{prox}_{\gamma g_1} \colon x \mapsto \begin{cases} x + \dfrac{\gamma}{d_E(x)}\big(\text{proj}_E\, x - x\big), & \text{if } d_E(x) > \gamma;\\ \text{proj}_E\, x, & \text{if } d_E(x) \leqslant \gamma, \end{cases} \tag{5.78}$$

and [5, Proposition 24.11] yields

$$\text{prox}_{\gamma g_2} \colon (\xi_i, \eta_i)_{1\leqslant i\leqslant N} \mapsto \big(\varrho_i\xi_i, \varrho_i\eta_i\big)_{1\leqslant i\leqslant N}, \tag{5.79}$$

where

$$\varrho_i = 1 - \frac{\sqrt{8}\gamma}{\max\{\sqrt{8}\gamma, \|(\xi_i, \eta_i)\|\}}. \tag{5.80}$$

At last, [18, Example 2.3] establishes that

$$\nabla h \colon x \mapsto \frac{\rho_1 H_1^*(H_1x - z_1)}{\max\{\rho_1, \|H_1x - z_1\|\}} + \frac{\rho_2 H_2^*(H_2x - z_2)}{\max\{\rho_2, \|H_2x - z_2\|\}}. \tag{5.81}$$

(a) Problem 5.21.

(b) Problem 5.22 ($\gamma = 0.1$).

(c) Problem 5.22 ($\gamma = 0.99$).

Figure 5.4 Images reconstructed by Problems 5.21 and 5.22.

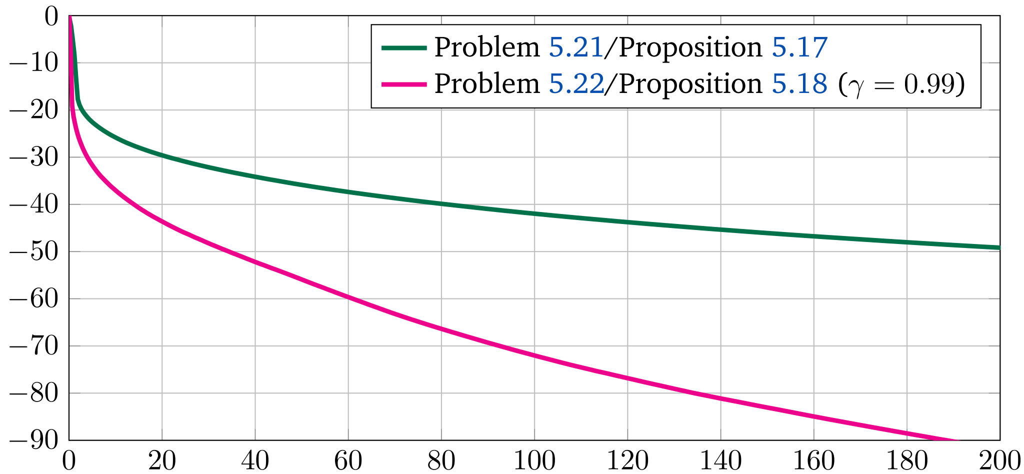


Figure 5.5 Normalized error $20\log_{10}(\|x_n - x_\infty\|/\|x_0 - x_\infty\|)$ (dB) versus time (s).

We construct the solution to Problem 5.21 shown in Fig. 5.4(a) via Proposition 5.17, where $\eta_n \equiv 0.49$. To apply Proposition 5.18 to Problem 5.22, let us verify that condition (5.65) is satisfied. By Proposition 5.10(iii)(c) and [5, Corollary 16.48(iii)],

$$\partial\Big(f + \operatorname{pcm}_\gamma\big(d_E, \mathrm{Id}; \sqrt{8}\|\cdot\|_{1,2}, D/\sqrt{8}\big) + h\Big) = \partial f + \partial \operatorname{pcm}_\gamma\big(d_E, \mathrm{Id}; \sqrt{8}\|\cdot\|_{1,2}, D/\sqrt{8}\big) + \nabla h, \quad (5.82)$$

which is a maximally monotone operator with domain C [5, Theorem 20.25]. Hence, (5.65) follows from [5, Corollary 21.25]. The image reconstructed by (5.66) for Problem 5.22 with $\gamma = 0.1$ and $\lambda_n \equiv 1.89 < 1.90 = 2 - \gamma/(2\beta)$ is shown in Fig. 5.4(b). As predicted by Theorem 5.13(iv)(b), since γ is small, this solution is similar to that produced by Problem 5.21 in Fig. 5.4(a). The solution produced by Problem 5.22 for $\gamma = 0.99$ with $\lambda_n \equiv 1 < 1.01 = 2 - \gamma/(2\beta)$ in (5.66) is shown in Fig. 5.4(c) to yield a slightly sharper reconstruction. Finally, Fig. 5.5 illustrates the faster convergence of algorithm (5.66) for the proximal comixture model (5.76).

5.2.5.3 Experiment 3: Image reconstruction from phase

We address a phase recovery problem considered in [18]. The goal is to recover the original image $\overline{x} \in C = [0, 255]^N$ ($N = 256^2$) shown in Fig. 5.6(a) from an imprecise observation of its Fourier phase $\theta \approx \angle\widehat{\overline{x}}$ [33]. The problem is modeled as a convex feasibility problem with the following constraint sets.

- Phase:

$$C_1 = \big\{x \in \mathbb{R}^N \mid \angle\,\widehat{x} = \theta\big\}. \tag{5.83}$$

- Mean pixel value:

$$C_2 = \big\{x \in \mathbb{R}^N \mid \langle x \mid 1\rangle = \eta\big\}. \tag{5.84}$$

- Proximity to the reference image r of Fig. 5.6(b):

$$C_3 = \big\{x \in \mathbb{R}^N \mid \|x - r\|_2 \leqslant \xi\big\}. \tag{5.85}$$

  The image r is a blurred and noise-corrupted version of $\overline{x}$, which is further degraded by saturation (the pixel values beyond 130 are clipped to 130).

- Upper bound on the norm of the gradient: $Dx/\sqrt{8} \in C_4$, where

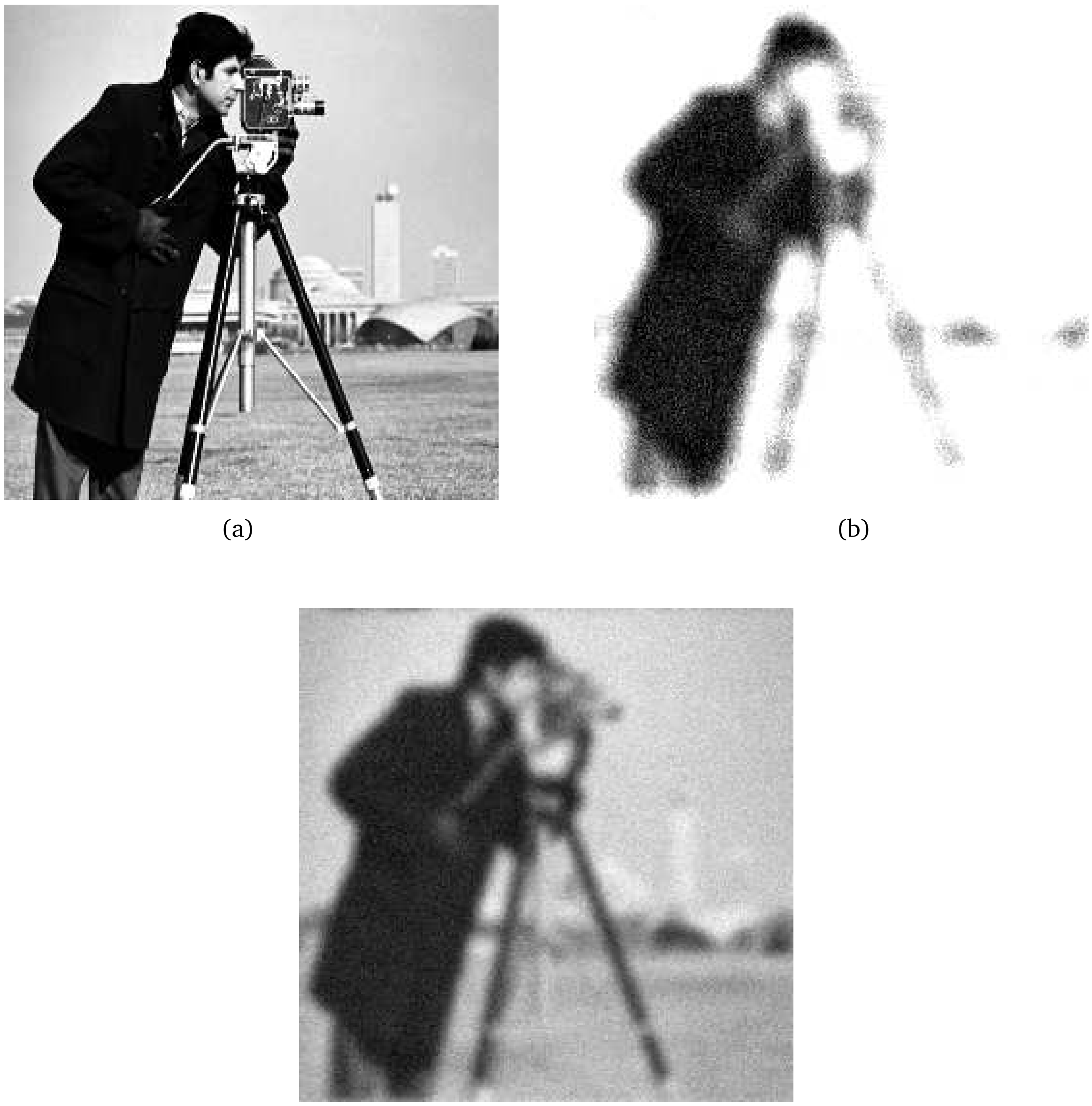
$$C_4 = \big\{y \in \mathbb{R}^N \times \mathbb{R}^N \mid \|y\|_2 \leqslant \rho\big\} \tag{5.86}$$

  and D is defined as in (5.73).

- A blurred observation of $\overline{x}$ is available, namely (see Fig. 5.6(c)) $z = H\overline{x} + w$, where H models a convolutional blur with a Gaussian kernel of size 23×13 and w is a white Gaussian noise realization.

Because of inaccuracies in the values θ, η, ρ, and ξ, the convex feasibility problem arising from the above constraints is inconsistent and we relax it using the Berhu function given by

$$\mathfrak{b} \colon \mathbb{R} \to \mathbb{R} \colon \xi \mapsto \begin{cases} \dfrac{\xi^2 + 1}{2}, & \text{if } |\xi| > 1; \\ |\xi|, & \text{if } |\xi| \leqslant 1. \end{cases} \tag{5.87}$$



(c)

Figure 5.6 (a) Original image $\overline{x}$. (b) Reference image r. (c) Corrupted observation z.

The first formulation employs a standard composite average.

Problem 5.23 The task is to

$$\underset{x \in C}{\text{minimize}} \ \ \frac{1}{4}\sum_{k=1}^{3} \mathfrak{b}\big(d_{C_k}(x)\big) + \frac{1}{4}\mathfrak{b}\left(d_{C_4}\left(\frac{1}{\sqrt{8}}Dx\right)\right) + \frac{1}{2}\|Hx - z\|^2. \tag{5.88}$$

Our second formulation employs a proximal comixture.

Problem 5.24 Let $\gamma \in \,]0,2[$. The task is to

$$\underset{x \in C}{\text{minimize}} \;\; \Big(\text{pcm}_\gamma\big(\mathfrak{b} \circ d_{C_1}, \text{Id}; \mathfrak{b} \circ d_{C_2}, \text{Id}; \mathfrak{b} \circ d_{C_3}, \text{Id}; \mathfrak{b} \circ d_{C_4}, D/\sqrt{8}\big)\Big)(x) + \frac{1}{2}\|Hx - z\|^2. \quad (5.89)$$

(a) Problem 5.23.

(b) Problem 5.24 ($\gamma = 0.1$).

(c) Problem 5.24 ($\gamma = 1.99$).

Figure 5.7 Images reconstructed by Problems 5.23 and 5.24.

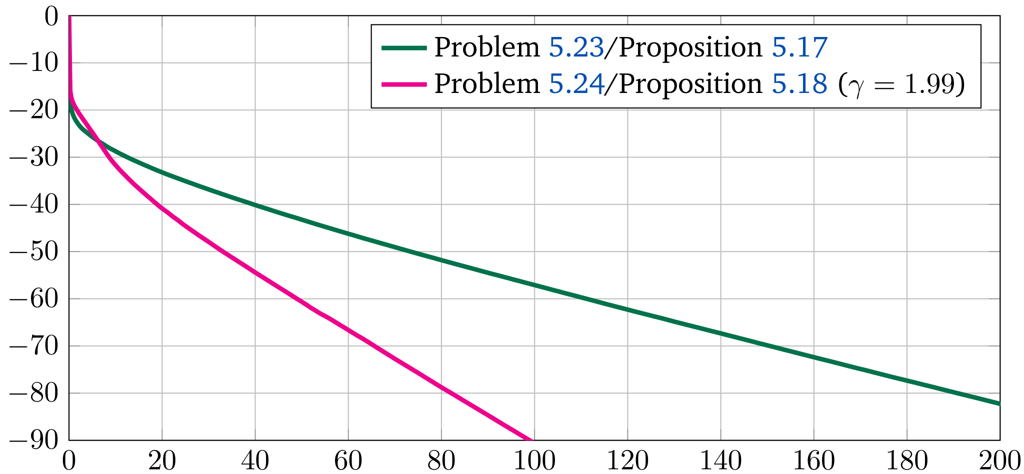


Figure 5.8 Normalized error $20\log_{10}(\|x_n - x_\infty\|/\|x_0 - x_\infty\|)$ (dB) versus time (s).

Problems 5.23 and 5.24 are particular instances of Problems 5.2 and 5.3, respectively, where $\mathcal{H} = \mathbb{R}^N$, $f = \iota_C$, $h = \|H \cdot -z\|^2/2$, $\beta = 1$, $p = 4$, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1/4$, and, for every $k \in \{1,2,3\}$, $\mathcal{G}_k = \mathbb{R}^N$, $g_k = \mathfrak{b} \circ d_{C_k}$, and $L_k = \mathrm{Id}$, whereas $\mathcal{G}_4 = \mathbb{R}^N \times \mathbb{R}^N$, $g_4 = \mathfrak{b} \circ d_{C_4}$, and $L_4 = D/\sqrt{8}$. We have $\|L_1\|^2 = \|L_2\|^2 = \|L_3\|^2 = \|L_4\|^2 = 1$. Further, $\mathrm{prox}_{\gamma f}$ is given as in (5.77), [5, Proposition 24.27] yields

$$(\forall k \in \{1,\ldots,4\}) \;\; \mathrm{prox}_{\gamma g_k} \colon x \mapsto \begin{cases} \mathrm{proj}_{C_k} x + \dfrac{1}{1+\gamma}\big(x - \mathrm{proj}_{C_k} x\big), & \text{if } d_{C_k}(x) > 1+\gamma; \\ x + \dfrac{\gamma}{d_{C_k}(x)}\big(\mathrm{proj}_{C_k} x - x\big), & \text{if } \gamma < d_{C_k}(x) \leqslant 1+\gamma; \\ \mathrm{proj}_{C_k} x, & \text{if } d_{C_k}(x) \leqslant \gamma, \end{cases} \tag{5.90}$$

and $\nabla h = H^* \circ (H \cdot -z)$.

We apply Proposition 5.17 to Problem 5.23 with $\eta_n \equiv 0.59$, which produces the reconstructed image shown in Fig. 5.7(a). Following the same argument used in Problem 5.22, we note that condition (5.65) is satisfied, and we apply Proposition 5.18 to Problem 5.24. The image reconstructed by (5.66) for $\gamma = 0.1$ and $\lambda_n \equiv 1.94 < 1.95 = 2 - \gamma/(2\beta)$ is shown in Fig. 5.7(b). This solution is similar to that of Fig. 5.7(a), which is consistent with Theorem 5.13(iv)(b) since γ is small. For Problem 5.24 with $\gamma = 1.99$, algorithm (5.66) with $\lambda_n \equiv 1 < 1.005 = 2-\gamma/(2\beta)$ yields the somewhat sharper reconstruction shown in Fig. 5.7(c). Fig. 5.8 depicts the faster convergence of algorithm (5.66) for the proximal comixture model (5.89).

5.2.5.4 Experiment 4: Linear regression

This experiment focuses on a linear regression model with an overlapping group structure on the inputs [12, 46]. We consider $p = 40$ groups of indices $(I_k)_{1\leqslant k\leqslant p}$ in $\{1,\ldots,N\}$ ($N = 90p + 10 = 3610$) of length 100 such that two consecutive groups overlap by 10 variables, i.e.,

$$I_1 = \{1,\ldots,100\}, I_2 = \{91,\ldots,190\},\ldots, I_p = \{N-99,\ldots,N\}. \tag{5.91}$$

We use $M = 5000$ samples. The entries of the input matrix $A \in \mathbb{R}^{M\times N}$ are i.i.d. samples from a $\mathcal{N}(0,1)$ distribution. The entries of the true regression coefficients $\overline{x} \in \mathbb{R}^N$ are also i.i.d. samples from a $\mathcal{N}(0,1)$ distribution, and the output data is generated by the noisy linear model $z = A\overline{x} + w$, where $w \in \mathbb{R}^M$ has entries that are i.i.d. samples from a $\mathcal{N}(0,1)$ distribution. For every $k \in \{1,\ldots,p\}$, we set $L_k\colon \mathbb{R}^N \to \mathbb{R}^{100}\colon x = (\xi_i)_{1\leqslant i\leqslant N} \mapsto (\xi_i)_{i\in I_k}$.

As before, we consider formulations based on a standard composite average and the proximal comixture.

Problem 5.25 The task is to

$$\underset{x\in\mathbb{R}^N}{\text{minimize}}\;\; \frac{1}{p}\|x\|_1 + \frac{1}{p}\sum_{k=1}^{p}\|L_k x\| + \frac{1}{2p^2}\|Ax - z\|^2. \tag{5.92}$$

Problem 5.26 Let $\gamma \in \left]0, 2p^2/\|A\|^2\right[$. The task is to

$$\underset{x\in\mathbb{R}^N}{\text{minimize}}\;\; \frac{1}{p}\|x\|_1 + \Big(\text{pcm}_\gamma\big(\|\cdot\|, L_k\big)_{1\leqslant k\leqslant p}\Big)(x) + \frac{1}{2p^2}\|Ax - z\|^2. \tag{5.93}$$

Problems 5.25 and 5.26 are particular instances of Problems 5.2 and 5.3, respectively, where $\mathcal{H} = \mathbb{R}^N$, $f = \|\cdot\|_1/p$, $h = \|A\cdot - z\|^2/(2p^2)$, $\beta = p^2/\|A\|^2$, and, for every $k \in \{1,\ldots,p\}$, $\alpha_k = 1/p$, $\mathcal{G}_k = \mathbb{R}^{100}$, $g_k = \|\cdot\|$, and $\|L_k\| = 1$. Additionally,

$$\text{prox}_{\gamma f}\colon (\xi_i)_{1\leqslant i\leqslant N} \mapsto \big(\text{sign}\,(\xi_i)\max\{|\xi_i| - \gamma/p, 0\}\big)_{1\leqslant i\leqslant N}. \tag{5.94}$$

Further, [5, Example 24.20] yields

$$(\forall k \in \{1,\ldots,p\})\;\; \text{prox}_{\gamma g_k}\colon x \mapsto \left(1 - \frac{\gamma}{\max\{\|x\|, \gamma\}}\right)x, \tag{5.95}$$

while $\nabla h = (1/p^2)A^* \circ (A\cdot - z)$. Next, let us verify that condition (5.65) is satisfied. Note that, by Proposition 5.12, $\text{pcm}_\gamma(\|\cdot\|, L_k)_{1\leqslant k\leqslant p} \geqslant 0$. Thus, the objective function of Problem 5.26 is coercive and the existence of minimizers is guaranteed by [5, Proposition 11.15(i)]. Therefore, Fermat's rule [5, Theorem 16.3] and [5, Corollary 16.48(iii)] guarantee that condition (5.65) holds. The solutions x_∞ to Problem 5.25 and to Problem 5.26 for $\gamma = 0.18$ are very close. In addition, $\|x_\infty - \overline{x}\|/\|\overline{x}\| \approx 0.058$. For Problem 5.25, we use algorithm (5.63) with $\eta_n \equiv 0.17$, while for Problem 5.26 for $\gamma = 0.18$, we use algorithm (5.66) with $\lambda_n \equiv 1 < 1.039 = 2 - \gamma/(2\beta)$. The faster convergence of algorithm (5.66) for the proximal comixture model (5.93) is shown in Fig. 5.9.

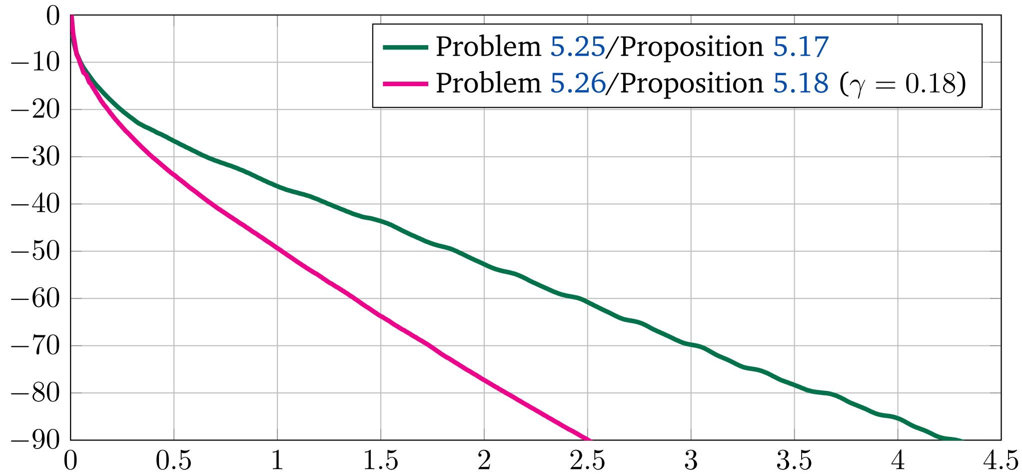


Figure 5.9 Normalized error $20\log_{10}(\|x_n - x_\infty\|/\|x_0 - x_\infty\|)$ (dB) versus time (s).

Acknowledgement

This work was supported by the National Science Foundation under grant CCF-2211123.

References

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

[2] A. Ang, W. Bin Hamed, and H. De Sterck, Sum-of-norms regularized nonnegative matrix factorization, *Neural Comput.*, vol. 38, pp. 228–255, 2026.

[3] H. Attouch, *Variational Convergence for Functions and Operators*. Pitman, Boston, MA, 1984.

[4] H. Attouch and D. Azé, Approximation and regularization of arbitrary functions in Hilbert spaces by the Lasry–Lions method, *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, vol. 10, pp. 289–312, 1993.

[5] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed. Springer, New York, 2017.

[6] H. H. Bauschke, R. Goebel, Y. Lucet, and X. Wang, The proximal average: Basic theory, *SIAM J. Optim.*, vol. 19, pp. 766–785, 2008.

[7] P. Bernard, Lasry–Lions regularization and lemma of Ilmanen, *Rend. Semin. Mat. Univ. Padova*, vol. 124, pp. 221–229, 2010.

[8] M. N. Bùi and P. L. Combettes, Integral resolvent and proximal mixtures, *J. Optim. Theory Appl.*, vol. 203, pp. 2328–2353, 2024.

[9] Y. Censor, T. Elfving, N. Kopf, and T. Bortfeld, The multiple-sets split feasibility problem and its applications for inverse problems, *Inverse Problems*, vol. 21, pp. 2071–2084, 2005.

[10] Y. Censor and M. Zaknoon, Algorithms and convergence results of projection methods for inconsistent feasibility problems: A review, *Pure Appl. Funct. Anal.*, vol. 3, pp. 565–586, 2018.

[11] A. Chambolle and T. Pock, An introduction to continuous optimization for imaging, *Acta Numer.*, vol. 25, pp. 161–319, 2016.

[12] X. Chen, Q. Lin, S. Kim, J. G. Carbonell, and E. P. Xing, Smoothing proximal gradient method for general structured sparse regression, *Ann. Appl. Stat.*, vol. 6, pp. 719–752, 2012.

[13] Y.-M. Cheung and J. Lou, Proximal average approximated incremental gradient descent for composite penalty regularized empirical risk minimization, *Mach. Learn.*, vol. 106, pp. 595–622, 2017.

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

[15] P. L. Combettes, Resolvent and proximal compositions, *Set-Valued Var. Anal.*, vol. 31, art. 22, 29 pp., 2023.

[16] P. L. Combettes, The geometry of monotone operator splitting methods, *Acta Numer.*, vol. 33, pp. 487–632, 2024.

[17] P. L. Combettes and D. J. Cornejo, Variational analysis of proximal compositions and integral proximal mixtures. *Evol. Equ. Control Theory*, vol. 17, pp. 106–139, 2026.

[18] P. L. Combettes and L. E. Glaudin, Proximal activation of smooth functions in splitting algorithms for convex image recovery, *SIAM J. Imaging Sci.*, vol. 12, pp. 1905–1935, 2019.

[19] P. L. Combettes and J.-C. Pesquet, Proximal splitting methods in signal processing, in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, pp. 185–212. Springer, New York, 2011.

[20] 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.

[21] P. L. Combettes and J.-C. Pesquet, Fixed point strategies in data science, *IEEE Trans. Signal Process.*, vol. 69, pp. 3878–3905, 2021.

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

[23] L. Condat, A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms, *J. Optim. Theory Appl.*, vol. 158, pp. 460–479, 2013.

[24] L. Condat, D. Kitahara, A. Contreras, and A. Hirabayashi, Proximal splitting algorithms for convex optimization: A tour of recent advances, with new twists, *SIAM Rev.*, vol. 65, pp. 375–435, 2023.

[25] D. Davis and W. Yin, A three-operator splitting scheme and its optimization applications, *Set-Valued Var. Anal.*, vol. 25, pp. 829–858, 2017.

[26] R. Glowinski, S. J. Osher, and W. Yin (eds.), *Splitting Methods in Communication, Imaging, Science, and Engineering*. Springer, New York, 2016.

[27] J.-B. Hiriart-Urruty, A general formula on the conjugate of the difference of functions, *Canad. Math. Bull.*, vol. 29, pp. 482–485, 1986.

[28] J.-B. Hiriart-Urruty, The deconvolution operation in convex analysis: An introduction, *Cybernet. Systems Anal.*, vol. 30, pp. 97–104, 1994.

[29] Z. Hu, K. Shaloudegi, G. Zhang, and Y. Yu, Federated learning meets multi-objective optimization, *IEEE Trans. Network Sci. Eng.*, vol. 9, pp. 2039–2051, 2022.

[30] A. B. Juditsky and A. S. Nemirovski, *Statistical Inference via Convex Optimization*. Princeton University Press, Princeton, NJ, 2020.

[31] J. M. Lasry and P. L. Lions, A remark on regularization in Hilbert spaces, *Israel J. Math.*, vol. 55, pp. 257–266, 1986.

[32] A. M. Legendre, *Nouvelles Méthodes pour la Détermination des Orbites des Comètes*. Firmin Didot, Paris, 1805.

[33] A. Levi and H. Stark, Signal reconstruction from phase by projection onto convex sets, *J. Opt. Soc. Am.*, vol. 73, pp. 810–822, 1983.

[34] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, Federated optimization in heterogeneous networks, *Proc. Mach. Learn. Syst.*, vol. 2, pp. 429–450, 2020.

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

[36] P. Pérez-Aros and E. Vilches, Moreau envelope of supremum functions with applications to infinite and stochastic programming, *SIAM J. Optim.*, vol. 31, pp. 1635–1657, 2021.

[37] D. L. Phillips, A technique for the numerical solution of certain integral equations of the first kind, *J. Assoc. Comput. Mach.*, vol. 9, pp. 84–97, 1962.

[38] R. T. Rockafellar and R. J. B. Wets, *Variational Analysis*, 3rd printing. Springer, New York, 2009.

[39] M. I. Sezan and H. Stark, Image restoration by convex projections in the presence of noise, *Appl. Opt.*, vol. 22, pp. 2781–2789, 1983.

[40] L. Shen, W. Liu, J. Huang, Y.-G. Jiang, and S. Ma, Adaptive proximal average approximation for composite convex minimization, *Proc. AAAI Conf. Artif. Intell.*, vol. 31, pp. 2513–2519, 2017.

[41] T. Strömberg, On regularization in Banach spaces, *Ark. Mat.*, vol. 34, pp. 383–406, 1996.

[42] Y. Sun, Z. Wu, X. Xu, B. Wohlberg, and U. S. Kamilov, Scalable plug-and-play ADMM with convergence guarantees, *IEEE Trans. Comput. Imaging*, vol. 7, pp. 849–863, 2021.

[43] S. Theodoridis, *Machine Learning: A Bayesian and Optimization Perspective*, 2nd ed. Elsevier, Amsterdam, 2020.

[44] B. C. Vũ, A splitting algorithm for dual monotone inclusions involving cocoercive operators, *Adv. Comput. Math.*, vol. 38, pp. 667–681, 2013.

[45] Y.-L. Yu, Better approximation and faster algorithm using the proximal average, *Proc. Conf. Adv. Neural Inform. Process. Syst.*, pp. 458–466, 2013.

[46] L. Yuan, J. Liu, and J. Ye, Efficient methods for overlapping group lasso, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, pp. 2104–2116, 2013.

Chapter 6

CONCLUSION

6.1 Summary

We have established new theoretical results that deepen the understanding of the parametrized resolvent and proximal compositions, and proposed a minimization model based on proximal comixtures. Our main contributions are as follows.

- In Chapter 2, we settled question **(Q1)** by establishing variational properties of the parametrized versions of the proximal composition and the integral proximal mixture. These include comparisons with standard compositions and results on asymptotic behavior, both pointwise and in the epi-convergence sense.
- In Chapter 3, we addressed question **(Q2)**, where we developed new interpretations and theoretical results for the parametrized resolvent compositions, including asymptotic results regarding operator convergence, specifically focusing on graph-convergence and the ρ-Hausdorff distance.
- Chapter 4 was devoted to answering question **(Q3)**, where we investigated the parametrized resolvent compositions for positive linear operators. The results include Löwner partial order relations, concavity, nonexpansiveness, asymptotic behavior in the operator norm topology, connections to a new form of geometric interpolation between the standard composite methods, and the study of nonlinear equations based on resolvent compositions.
- In Chapter 5, we answered question **(Q4)** by proposing a new minimization model based on proximal comixtures. We have analyzed the mathematical properties of these aggregation operations and presented comparisons with standard composite averages. Further, the benefits of the proximal comixtures compared to the standard composite averages in terms of modeling and algorithmic implementation were discussed. Finally, we have provided numerical illustrations in the context of image recovery and data analysis.

6.2 Future work

Direction 6.1 Resolvent and proximal compositions have been investigated in the Hilbert space setting. An interesting direction for future research is the analysis of these constructions in Banach spaces.

Direction 6.2 Proximal comixtures minimization models offer benefits from both modeling and algorithmic standpoints. Investigating block-iterative adaptive algorithms based on proximal comixtures should be explored.

Direction 6.3 Variants of the proximal average have been studied in the nonconvex setting [2,3]. An open question is to identify variants of the more general proximal compositions that can support a theoretical analysis in the nonconvex case.

Direction 6.4 The proposed proximal comixture minimization model can be viewed as an empirical mean of the more general concept of a proximal expectation [1]. In the case of standard expectations arising in risk minimization in statistics and machine learning, the asymptotic behavior of the empirical risk in the form of a finite average is well-understood. A natural open direction is to explore these statistical aspects in the case of estimators based on proximal comixtures.

Raleigh, March 6, 2026

References

[1] M. N. Bùi and P. L. Combettes, Integral resolvent and proximal mixtures, *J. Optim. Theory Appl.*, vol. 203, pp. 2328–2353, 2024.

[2] J. Chen, X. Wang, and C. Planiden, A proximal average for prox-bounded functions, *SIAM J. Optim.*, vol. 30, pp. 1366–1390, 2020.

[3] W. L. Hare, A proximal average for nonconvex functions: A proximal stability perspective, *SIAM J. Optim.*, vol. 20, pp. 650–666, 2009.