## ABSTRACT

WOODSTOCK, ZEV. Construction of Functions from Nonlinear Transformations. (Under the direction of Patrick L. Combettes.)

This thesis focuses on modeling, analyzing, and solving problems involving nonlinear equality constraints from the novel perspective of fixed point theory and monotone operator theory. It is shown that the class of nonlinearities involving firmly nonexpansive operators is broad enough to represent many equations arising in applications, even when the original equations feature non-Lipschitzian or even discontinuous operators. Adopting this model leads to feasibility and best approximation algorithms which are proven to converge to an exact solution of such equations from any initial point. Best approximation problems subjected to these nonlinear equations are solved with a new strongly convergent block-iterative algorithm that features an extrapolated relaxation scheme which exploits the presence of affine constraints. To address potentially inconsistent feasibility problems involving firmly nonexpansive equations, we propose a relaxation in the form of a variational inequality problem. Conditions for the existence of solutions to the relaxed problem are derived and a block-iterative algorithm is proposed for its solution. Next, block-iterative algorithms for fully nonsmooth convex minimization are analyzed, and their relative performance on concrete large-scale problems is assessed. Throughout, the theoretical and numerical aspects of this work are illustrated by applications to image processing, signal processing, and machine learning.



This thesis was supported by the National Science Foundation under grants DGE-1746939 and CCF-1715671.

Construction of Functions from Nonlinear Transformations

by
Zev Woodstock

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

2021

APPROVED BY:

C. Tim Kelley

Kazufumi Ito

Eric Chi

Ryan Murray

Patrick L. Combettes
Chair of Advisory Committee

## DEDICATION

To my cat, Jim.

## BIOGRAPHY

The author graduated *summa cum laude* with a B. S. in Mathematics from James Madison University in Harrisonburg, Virginia. He was admitted to North Carolina State University in 2016, and received his M.S. in Applied Mathematics from N.C. State in 2018. After being awarded the National Science Foundation Graduate Research Fellowship (NSF-GRFP) in 2018, he was admitted to doctoral candidacy in 2020.

## ACKNOWLEDGEMENTS

I would like to thank my advisor Patrick Combettes, who devoted a significant amount of time and effort to ensuring that I receive an excellent training. In addition to mathematical techniques, facts, and results, he also taught me about the philosophy of investigating, writing, and presenting mathematics. I would like to thank my committee members for their help and advice throughout this process. I would also like to thank my friends and colleagues Farid Benmouffok, Minh N. Bùi, Steven Gilmore, Lilian Glaudin, Ethan King, and Georgy Scholten for their helpful conversations and insights. Finally, I thank my family for their support and kindness.

This thesis was supported by the National Science Foundation under grants DGE-1746939 and CCF-1715671.

## TABLE OF CONTENTS

## LIST OF FIGURES

## NOTATION AND DEFINITIONS

The following notation are used throughout the thesis.

### General notation

- $\mathcal{H}, \mathcal{G}$: Real Hilbert spaces.
- $\langle \cdot \mid \cdot \rangle$: Scalar product of a real Hilbert space.
- $\| \cdot \|$: Norm of a Hilbert space.
- Let $(\mathcal{H}_i)_{i \in I}$ be real Hilbert spaces. $\bigoplus_{i \in I} \mathcal{H}_i = \left\{ (x_i)_{i \in I} \in \bigtimes_{i \in I} \mathcal{H}_i \;\middle|\; \sum_{i \in I} \|x_i\|^2 < +\infty \right\}$ is the Hilbert direct sum of $(\mathcal{H}_i)_{i \in I}$.
- $B(x; \rho)$: Closed ball with center $x \in \mathcal{H}$ and radius $\rho \in ]0, +\infty[$.
- Id : Identity operator.
- $2^{\mathcal{H}}$: The family of all subsets of $\mathcal{H}$.
- $\Gamma_0(\mathcal{H})$: The set of all proper, convex, lower semicontinuous functions from $\mathcal{H}$ to $]-\infty, +\infty]$.
- $\mathcal{B}(\mathcal{H}, \mathcal{G})$: The space of bounded linear operators from $\mathcal{H}$ to $\mathcal{G}$.
- $\mathcal{B}(\mathcal{H}) = \mathcal{B}(\mathcal{H}, \mathcal{H})$.
- $L^*$: The adjoint of $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.
- $\|L\| = \sup \left\{ \|Lx\| \;\middle|\; x \in B(0; 1) \right\}$: The norm of $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.
- $\to$ : Strong convergence.
- $\rightharpoonup$ : Weak convergence.

### Notation relative to a subset $C$ of $\mathcal{H}$

- $\iota_C$: Indicator function of $C$:

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

- $1_C$: Characteristic function of $C$:

$$1_C \colon \mathcal{H} \to [-\infty, +\infty] : x \mapsto \begin{cases} 1, & \text{if } x \in C; \\ 0, & \text{if } x \notin C. \end{cases}$$

- $\operatorname{cone} C = \big\{\lambda x \;\big|\; \lambda \in \left]0, +\infty\right[,\;\; x \in C\big\}$: Conical hull of $C$.
- $\operatorname{int} C$: Interior of $C$.
- $\overline{C}$: Closure of $C$.
- $\operatorname{ri} C = \big\{x \in C \;\big|\; \operatorname{cone}(C - x) = \operatorname{span}(C - x)\big\}$: Relative interior of $C$.
- $\operatorname{sri} C = \big\{x \in C \;\big|\; \operatorname{cone}(C - x) = \overline{\operatorname{span}}(C - x)\big\}$: Strong relative interior of $C$.
- $d_C \colon \mathcal{H} \to [0, +\infty] \colon x \mapsto \inf\limits_{p \in C} \|x - p\|$: Distance function to $C$.
- Suppose that $C$ is nonempty, closed, and convex. $\operatorname{proj}_C \colon \mathcal{H} \to \mathcal{H} \colon x \mapsto \operatorname*{argmin}\limits_{p \in C} \|x - p\|$: Projection operator onto $C$.

**Notation relative to an operator** $T \colon \mathcal{H} \to \mathcal{H}$

- $\operatorname{Fix} T = \big\{x \in \mathcal{H} \;\big|\; Tx = x\big\}$: Fixed points of $T$.
- Let $\beta \in \left]0, +\infty\right[$. $T$ is Lipschitz continuous with constant $\beta$, or $\beta$-Lipschitzian, if

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \|Tx - Ty\| \leqslant \beta \|x - y\|. \tag{1}$$

- $T$ is nonexpansive if it is Lipschitz continuous with constant $1$.
- Let $\alpha \in \left]0, 1\right]$. $T$ is $\alpha$-averaged if

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \|Tx - Ty\|^2 \leqslant \|x - y\|^2 - \frac{1 - \alpha}{\alpha} \|(\operatorname{Id} - T)x - (\operatorname{Id} - T)y\|^2.$$

- Let $\beta \in \left]0, +\infty\right[$. $T$ is $\beta$-cocoercive if

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

- $T$ is firmly nonexpansive if it is $1$-cocoercive, that is,

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

- Let $u \in \mathcal{H}$. $T$ is demi-closed at $u$ if, for every sequence $(x_n)_{n \in \mathbb{N}}$ in $\mathcal{H}$ and every $x \in \mathcal{H}$ such that $x_n \rightharpoonup x$ and $Tx_n \to u$, we have $Tx = u$. Furthermore, $T$ is demi-closed if it is demi-closed at every point in $\mathcal{H}$.

**Notation relative to a set-valued operator** $M\colon \mathcal{H} \to 2^{\mathcal{H}}$

- $\operatorname{dom} M = \big\{x \in \mathcal{H} \;\big|\; Mx \neq \varnothing\big\}$: Domain of $M$.
- $\operatorname{gra} M = \big\{(x,u) \in \mathcal{H} \times \mathcal{H} \;\big|\; u \in Mx\big\}$: Graph of $M$.
- $M^{-1}\colon \mathcal{H} \to 2^{\mathcal{H}}\colon u \mapsto \big\{x \in \mathcal{H} \;\big|\; u \in Mx\big\}$: Inverse of $M$.
- $\operatorname{zer} M = \big\{x \in \mathcal{H} \;\big|\; 0 \in Mx\big\}$: Zeros of $M$.
- $\operatorname{ran} M = \big\{u \in \mathcal{H} \;\big|\; (\exists x \in \mathcal{H})\ \ u \in Mx\big\}$: Range of $M$.
- $s\colon \operatorname{dom} M \to \mathcal{H}$ such that $(\forall x \in \operatorname{dom} M)\ s(x) \in Mx$: Selection of $M$.
- $M$ is monotone if

$$(\forall (x_1,u_1) \in \operatorname{gra} M)(\forall (x_2,u_2) \in \operatorname{gra} M) \quad \langle x_1 - x_2 \mid u_1 - u_2\rangle \geqslant 0.$$

- $M$ is maximally monotone if $M$ is monotone and there is no monotone operator from $\mathcal{H}$ to $2^{\mathcal{H}}$, the graph of which strictly contains $\operatorname{gra} M$.
- $J_M = (\operatorname{Id} + M)^{-1}$: Resolvent of $M$.

**Notation relative to a function** $f \in \Gamma_0(\mathcal{H})$

- $\operatorname{dom} f = \big\{x \in \mathcal{H} \;\big|\; f(x) < +\infty\big\}$: Domain of $f$.
- $\operatorname{epi} f = \big\{(x,\eta) \in \mathcal{H} \times \mathbb{R} \;\big|\; f(x) \leqslant \xi\big\}$: Epigraph of $f$.
- $\operatorname{epis} f = \big\{(x,\eta) \in \mathcal{H} \times \mathbb{R} \;\big|\; f(x) < \xi\big\}$: Strict epigraph of $f$.
- $\operatorname{lev}_{\leqslant \xi} f = \big\{x \in \mathcal{H} \;\big|\; f(x) \leqslant \xi\big\}$: Lower level set of $f$ at height $\xi \in \mathbb{R}$.
- $\operatorname{lev}_{< \xi} f = \big\{x \in \mathcal{H} \;\big|\; f(x) < \xi\big\}$: Strict lower level set of $f$ at height $\xi \in \mathbb{R}$.
- $\operatorname{dom} f = \big\{x \in \mathcal{H} \;\big|\; f(x) < +\infty\big\}$: Domain of $f$.
- $\operatorname{Argmin} f$: The set of minimizers of $f$ over $\mathcal{H}$.
- $f$ is convex if $\operatorname{epi} f$ is convex. Equivalently,

$$(\forall (x,y,\alpha) \in \mathcal{H} \times \mathcal{H} \times \left]0,1\right[) \quad f(\alpha x + (1-\alpha)y) \leqslant \alpha f(x) + (1-\alpha) f(y). \tag{3}$$

- $\bigoplus_{i\in I} f_i\colon \bigoplus_{i\in I} \mathcal{H}_i\colon (x_i)_{i\in I} \mapsto \sum_{i\in I} f_i(x_i)$ is the direct sum of finitely many functions $(f_i)_{i\in I}$ which respectively map from $\mathcal{H}_i$ to $\left]-\infty,+\infty\right]$.

- $f^*\colon \mathcal{H} \to ]-\infty,+\infty] : u \mapsto \sup_{x\in\mathcal{H}} \big(\langle x \mid u\rangle - f(x)\big)$: Conjugate of $f$.
- $\partial f\colon \mathcal{H} \to 2^{\mathcal{H}}\colon x \mapsto \big\{u \in \mathcal{H} \;\big|\; (\forall y \in \mathcal{H})\, \langle y - x \mid u\rangle + f(x) \leqslant f(y)\big\}$: Subdifferential of $f$.
- For every $x \in \mathcal{H}$, $\operatorname{prox}_f x$ is the unique minimizer of $f(\cdot) + \|x - \cdot\|^2/2$. The mapping $x \mapsto \operatorname{prox}_f x$ is the proximity operator of $f$.
- $f$ is smooth if $\operatorname{dom} f = \mathcal{H}$, $f$ is differentiable, and $\nabla f$ is Lipschitzian.
- Let $f\colon \mathcal{H} \to \mathbb{R}$ be continuous and convex, let $s$ be a selection of $\partial f$, and suppose that $\operatorname{lev}_{\leqslant 0} f \neq \varnothing$. The subgradient projector onto $\operatorname{lev}_{\leqslant 0} f$ is given by

$$(\forall x \in \mathcal{H})\;\; Tx = \operatorname{sproj}_C x = \begin{cases} x - \dfrac{f(x)}{\|s(x)\|^2} s(x), & \text{if } \; f(x) > 0; \\ x, & \text{if } \; f(x) \leqslant 0. \end{cases} \tag{4}$$

**Notation relating to Fourier analysis**

- Suppose that $\mathcal{H} = \mathbb{R}^N$ and let $x \in \mathcal{H}$. $\operatorname{DFT}(x)$ or $\widehat{x}$: Discrete Fourier transform of $x$.
- Suppose that $\mathcal{H} = \mathbb{R}^N$ and let $x \in \mathcal{H}$. $\operatorname{IDFT}(x)$: Inverse discrete Fourier transform of $x$.
- Let $\xi \in \mathbb{C}$. We write $\xi = |\xi| \exp(i(\angle \xi))$.

Chapter 1

# INTRODUCTION

## 1.1 Overview

This thesis focuses on modeling, analyzing, and solving problems involving nonlinear equality constraints from the novel perspective of fixed point theory and monotone operator theory. Nonlinear equations appear in classical and contemporary best approximation problems, as well as in a wide array of applications such as interpolation theory, optimization, machine learning, signal and image processing, and statistics [1, 2, 4, 5, 7, 18, 20, 21, 24, 25, 30, 35–39, 43, 45–47, 49]. Despite the great deal of effort devoted to this field, there are still many classes of nonlinear equations for which there are no provenly-convergent methods for their solution within the context of feasibility, best approximation, or optimization problems. This thesis contributes to this pursuit by identifying and analyzing one such class, which is broad enough to model many nonlinearities arising in applications, and simultaneously leads to algorithms which are guaranteed to converge to an exact solution from any initial point.

It is often the case that a function $\overline{x}$ in a real Hilbert space $\mathcal{H}$ is observed through a nonlinear process $F\colon \mathcal{H} \to \mathcal{H}$, say $F\overline{x} = p$ [19, 22, 23, 25, 26, 28, 29, 33, 42, 45, 46]. However, the task of recovering $\overline{x}$ from its nonlinear transformation $p$ is typically difficult or impossible. Therefore, in various scenarios, be they in synthesis or inverse problems, the constraint on a solution $x \in \mathcal{H}$ corresponding to this nonlinear transformation is given by

$$Fx = p. \tag{1.1}$$

In synthesis problems, this equation models a specification imposed on the solution, whereas in inverse problems it models an observation of the true function of interest. We propose modeling the underlying nonlinearity via a *firmly nonexpansive* operator, i.e., $F$ satisfies

$$(\forall (x, y) \in \mathcal{H} \times \mathcal{H}) \quad \|Fx - Fy\|^2 \leqslant \langle x - y \mid Fx - Fy \rangle. \tag{1.2}$$

As will be shown in Chapters 2, 3, and 4, this model covers many concrete situations. Even

though (1.2) requires that $F$ be Lipschitz continuous with constant 1, this work demonstrates that many constraints of the form (1.1) where $F$ is non-Lipschitzian and even discontinuous can be equivalently modeled via firmly nonexpansive operators [14–16]. While we frequently assume that the operator $F$ in (1.1) satisfies (1.2), we never assume that $F$ is smooth. This is because, even for simple examples in $\mathbb{R}$, e.g., when $F$ is the projector onto an interval, $F$ fails to be differentiable. In fact, even in $\mathbb{R}^2$, simple operators satisfying (1.2), such as the projector onto a closed convex set, may not even be directionally differentiable [44]. This means that common methods for solving smooth nonlinear equations such as Newton methods cannot be applied here [31]. Other methods for solving nonsmooth nonlinear equations have a different set of requirements [32] which are not assumed in this thesis.

In many applications, there are several operator-specification/observation pairs $(F_k, p_k)_{k\in K}$ as well as constraints in the form of closed convex subsets $(C_j)_{j\in J}$ of $\mathcal{H}$ which model *a priori* information on the desired solution [4, 10, 13, 43]. It is therefore natural to seek a vector $x \in \mathcal{H}$ which is consistent with all of the observations and constraints. This task is modeled as follows.

**Problem 1.1.1** Let $J$ and $K$ be at most countable sets such that $J \cap K = \varnothing$ and $J \cup K \neq \varnothing$. For every $j \in J$, let $C_j$ be a closed convex subset of $\mathcal{H}$ and, for every $k \in K$, let $p_k \in \mathcal{H}$ and let $F_k \colon \mathcal{H} \to \mathcal{H}$ be a firmly nonexpansive operator. The goal is to

$$\text{find} \;\; x \in \bigcap_{j\in J} C_j \;\; \text{such that} \;\; (\forall k \in K) \quad F_k x = p_k, \tag{1.3}$$

under the assumption that such a solution exists.

Note that if $K = \varnothing$, then Problem 1.1.1 reduces to the standard convex feasibility problem, which is well-understood and has many solution methods [10]. On the other hand, Problem 1.1.1 has a rich history for the setting when $(F_k)_{k\in K}$ are linear [3,13,27,34,36,37,40,49]. For instance, there are classical algorithms which solve Problem 1.1.1 when $(C_j)_{j\in J}$ are described by convex inequality constraints and the operators $(F_k)_{k\in K}$ in the equality constraints are linear [41].

On the other hand, when the transformations $(F_k)_{k\in K}$ are nonlinear, we are not aware of an algorithm which, under no further assumptions, possesses the ability to converge to an exact solution of Problem 1.1.1 from any initial point. The central challenge is enforcing constraints of the form (1.1) and, in particular, those in Problem 1.1.1. For instance, conventional approaches such as minimizing $\|F(\cdot) - p\|$ typically lead to nonconvex problem formulations, which make it difficult to guarantee that a numerical method will globally converge to a solution.

**Example 1.1.2** In signal declipping problems, the goal is to recover a function from its clipped values (see Figure 1.1). The nonlinearity in this problem models a sensor which records only values within an interval $D \subset \mathbb{R}$. This sensor is therefore the projector onto the set of functions whose range is contained in $D$. In the discrete setting, the nonlinearity is modeled via $F \colon \mathbb{R}^N \to \mathbb{R}^N \colon (\xi_i)_{1\leqslant i\leqslant N} \mapsto (\operatorname{proj}_D \xi_i)_{1\leqslant i\leqslant N}$ and the clipped signal by $p \in D^N$. Many approaches to solving

this task have been proposed; for instance, see the recent survey [50] and its references. Among them, many arrive at nonconvex problem formulations, do not guarantee convergence to a solution of Problem 1.1.1, rely on a parameterization of the signal, and/or make an assumption about the solution (e.g., sparsity or autoregressiveness) [2, 17, 19, 25]. In fact, to our knowledge, the only method which is proven to converge at all in this case [39] assumes that the underlying signal is sparse and converges to a minimizer of

$$x \mapsto \varphi(x) + d_C^2(x), \tag{1.4}$$

where $\varphi\colon \mathbb{R}^N \to \mathbb{R}$ promotes sparsity, and $d_C^2$ is the squared distance to $C = \big\{x \in \mathcal{H} \;\big|\; Fx = p\big\}$. It should be noted that a similar methodology also appears in quantization problems [42, 48]. The issue with this type of approach is that it requires access to the projection operator $\mathrm{proj}_C$, as it appears in the evaluation of $d_C^2$ and its gradient. When a nonlinearity applies univariate componentwise transformations, as is the case for clipping or quantization, $\mathrm{proj}_C$ can often be computed in closed form. On the other hand, outside of the univariate case, computing this operation is either impossible or numerically expensive.

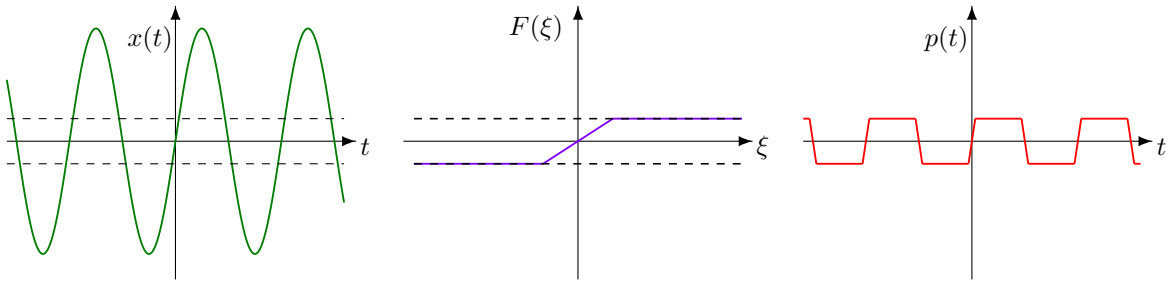



**Figure 1.1** (a): Original signal. (b): Clipping nonlinearity. (c): Clipped signal.

Many problems arising in approximation theory, harmonic analysis, signal processing, and optics also require the ability the select a particular point in the solution set of Problem 1.1.1 [3, 13, 27, 34, 36, 37, 40, 49]. For instance, to reduce oscillations in a recovery, one may seek the minimal norm solution. More generally, given a point $x_0 \in \mathcal{H}$, it is often desirable to find the solution which is closest to $x_0$, i.e., its *best approximation* from the solution set of Problem 1.1.1. This leads to the following formulation.

**Problem 1.1.3** Let $x_0 \in \mathcal{H}$ and let $J$ and $K$ be at most countable sets such that $J \cap K = \varnothing$ and $J \cup K \neq \varnothing$. For every $j \in J$, let $C_j$ be a closed convex subset of $\mathcal{H}$ and, for every $k \in K$, let $p_k \in \mathcal{H}$ and let $F_k\colon \mathcal{H} \to \mathcal{H}$ be a firmly nonexpansive operator. Suppose that there exists

$\overline{x} \in \bigcap_{j\in J} C_j$ such that $(\forall k \in K)\ F_k\overline{x} = p_k$. The task is to

$$\text{minimize } \|x - x_0\| \quad \text{subject to} \quad x \in \bigcap_{j\in J} C_j \quad \text{and} \quad (\forall k \in K) \quad F_k x = p_k. \tag{1.5}$$

The fully linear setting of Problem 1.1.3 has been studied in [13, 34]. While nonlinearities appear in the sets $(C_j)_{j\in J}$ in [8, 21, 24], in all of these settings, the $(F_k)_{k\in K}$ remain linear. To our knowledge, this problem has not been solved with a provenly-convergent algorithm in full generality, even for the case when $(F_j)_{j\in J}$ are all projections onto nonempty closed convex sets.

It may be the case that noise or inaccurate modeling can cause Problem 1.1.1 to have no solutions. Therefore, we are in need of a relaxed problem in the sense that, whenever Problem 1.1.1 is feasible, the relaxation is actually equivalent to the original problem. Furthermore, even if there is no solution to the original problem, we seek a problem which satisfies (1.1) in some approximate sense and is guaranteed to possess solutions under mild, verifiable conditions. To address this task, we consider the following variant of Problem 1.1.1 where the observation model is more structured.

**Problem 1.1.4** Let $C \subset \mathcal{H}$ be nonempty, closed, and convex. Let $(\mathcal{G}_i)_{i\in I}$ be a finite collection of real Hilbert spaces and, for every $i \in I$, let $L_i \in \mathcal{B}(\mathcal{H}, \mathcal{G}_i)$, let $p_i \in \mathcal{G}_i$, and let $F_i\colon \mathcal{H} \to \mathcal{H}$ be a firmly nonexpansive operator. The goal is to

$$\text{find } \ x \in C \ \text{ such that } \ (\forall i \in I) \quad F_i(L_i x) = p_i, \tag{1.6}$$

While Problem 1.1.4 may lack solutions, we show in Chapter 4 that the following relaxed problem frequently has solutions and, whenever solutions to Problem 1.1.4 exist, both problems are equivalent.

**Problem 1.1.5** Let $I$ be a nonempty finite set, let $(\omega_i)_{i\in I}$ real numbers in $]0, 1]$ such that $\sum_{i\in I} \omega_i = 1$, and let $C$ be a nonempty closed convex subset of a real Hilbert space $\mathcal{H}$. For every $i \in I$, let $\mathcal{G}_i$ be a real Hilbert space, let $p_i \in \mathcal{G}_i$, let $L_i\colon \mathcal{H} \to \mathcal{G}_i$ be a nonzero bounded linear operator, and let $F_i\colon \mathcal{G}_i \to \mathcal{G}_i$ be a firmly nonexpansive operator. The task is to

$$\text{Find } \ x \in C \ \text{ such that } \ (\forall y \in C) \ \sum_{i\in I} \omega_i \langle L_i(y - x) \mid F_i(L_i x) - p_i\rangle \geqslant 0. \tag{1.7}$$

It is shown in Chapter 4 that Problem 1.1.5 captures classical relaxations of inconsistent convex feasibility problems, e.g., the least-squares relaxation [9], as a special case.

Fixed point theory and monotone operator theory are used in a slightly different context in Chapter 5, where we transition from enforcing nonlinear equations to analyzing block-iterative algorithms for solving fully nonsmooth multicomponent convex minimization problems. While algorithms for optimization problems with at least one smooth term in the objective function are well-understood, few block-activated methods are available for the fully-nonsmooth case,

and their relative merits have not been thoroughly investigated. It is determined that only two block-activated algorithms [11, 12] possess a list of desirable features needed for large-scale applications. For the first time, the merits of both algorithms are analyzed and their numerical performance is compared.

## 1.2 Organization

Chapter 2 presents a method for solving Problem 1.1.1, under the assumptions that $J$ and $K$ are finite in the setting of a Euclidean space. This method is guaranteed to converge to an exact solution of Problem 1.1.1 from any initial point. Several examples in signal and image processing are used to demonstrate the process of *proxification*, by which we mean the replacement of a nonlinear equation with an equivalent one involving a firmly nonexpansive operator. It is shown that, even for transformations of the form (1.1) for which $F$ is discontinuous or non-Lipschitzian, there is frequently an equivalent formulation involving a firmly nonexpansive operator. A relaxation for Problem 1.1.1 is introduced as well.

While Chapter 2 has several examples of proxification, Chapter 3 develops in Section 3.2.2 generic strategies for constructing and modeling prescribed proximal points, i.e., extending the notion of "proxification" from Chapter 2. In Section 3.2.3, a new best approximation algorithm is proposed and proven to converge strongly; it is then used to solve Problem 1.1.3. Additionally, Remark 3.2.25(ii) provides an algorithm to solve Problem 1.1.1 in full generality.

Chapter 4 contains a detailed analysis of Problem 1.1.5 including common, mild, and verifiable guarantees of existence of its solutions. An efficient block-iterative algorithm for the solution of Problem 1.1.5 is presented. Finally, the theory of proxification from the previous chapters is extended with new results concerning matrix-valued and hypomonotone operators.

Chapter 5 concerns solving large-scale multicomponent nonsmooth optimization problems. It is determined that there are apparently only two algorithms possessing all of our required features for this class of problems. The merits of both algorithms are compared, and their relative numerical performance is evaluated in unexplored settings.

## 1.3 Main contributions

This work produced the articles [6, 14–16]. The main contributions and novelties are outlined below.

- A framework for modeling nonlinear equations with firmly nonexpansive operators. This class is shown to capture to many nonlinearities in applications, even when an observation operator fails to be continuous. (Chapters 2, 3, and 4)
- Provenly-convergent algorithms in the real Hilbert space setting for enforcing a countable number of closed convex set constraints and a countable number of firmly nonexpansive nonlinear equations of the form (1.1) in Problem 1.1.1 and Problem 1.1.3. (Chapter 3)

- A best approximation algorithm for solving Problem 1.1.3 which is block-iterative, strongly convergent, and whose extrapolation strategy is more general than other methods in the literature. This extrapolation strategy is shown numerically to significantly improve performance. (Chapter 3)
- Problem 1.1.5 is developed and shown to be a relaxation of Problem 1.1.4 in the sense that both problems are equivalent when Problem 1.1.4 has a solution. (Chapters 2 and 4)
- Existence theory showing that Problem 1.1.5 possess solutions under several mild and easily verifiable conditions. (4)
- A operator-theoretic methodology for promoting properties such as sparsity or smoothness in a solution. (Chapter 4)
- Block-activated fully proximal splitting algorithms for convex minimization are investigated in previously-unexplored settings. (Chapter 5)
- Throughout the thesis, theoretical and numerical aspects of this work are demonstrated by applications in bandlimited extrapolation, source separation, phase reconstruction, matrix completion, compression, sparse signal recovery, minimal-norm interpolation, classification problems in data science, image interpolation, image super-resolution, image deblurring, signal declipping, and multi-objective minimization.

## References


[1] N. Aronszajn, *Introduction to Theory of Hilbert Spaces*. Oklahoma A & M College Mathematical Monographs, Stillwater, OK, 1950.

[2] F. R. Ávila, M. P. Tcheou, and L. W. P. Biscainho, Audio soft declipping based on constrained weighted least squares, *IEEE Signal Process. Lett.*, vol. 24, pp. 1348–1352, 2017.

[3] A. L. Babierra and N. N. Reyes, A new characterization of the generalized inverse using projections on level sets, *J. Approx. Theory*, vol. 236, pp. 23–35, 2018.

[4] H. Boche, M. Guillemard, G. Kutyniok, and F. Philipp, Signal recovery from thresholded frame measurements, *Proc. 15th SPIE Wavelets Sparsity Conf.*, vol. 8858, pp. 80–86, 2013.

[5] L. M. Brègman, The method of successive projection for finding a common point of convex sets, *Soviet Math.-Dokl.*, vol. 6, pp. 688–692, 1965.

[6] M. N. Bùi, P. L. Combettes, and Z. C. Woodstock, Block-activated algorithms for multicomponent fully nonsmooth minimization, submitted.

[7] E. W. Cheney and A. A. Goldstein, Proximity maps for convex sets, *Proc. Amer. Math. Soc.*, vol. 10, pp. 448–450, 1959.

[8] C. K. Chui, F. Deutsch, and J. D. Ward, Constrained best approximation in Hilbert space, *Constr. Approx.*, vol. 6, pp. 35–64, 1990.

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

[10] P. L. Combettes, "The Convex Feasibility Problem in Image Recovery," vol. 95 of *Advances in Imaging and Electron Physics*, Academic Press, New York, 1996.

[11] P. L. Combettes and J. Eckstein, Asynchronous block-iterative primal-dual decomposition methods for monotone inclusions, *Math. Program.*, vol. B168, pp. 645–672, 2018.

[12] P. L. Combettes and J.-C. Pesquet, Stochastic quasi-Fejér block-coordinate fixed point iterations with random sweeping, *SIAM J. Optim.*, vol. 25, pp. 1221–1248, 2015.

[13] P. L. Combettes and N. N. Reyes, Functions with prescribed best linear approximations, *J. Approx. Theory*, vol. 162, pp. 1095–1116, 2010.

[14] P. L. Combettes and Z. C. Woodstock, A fixed point framework for recovering signals from nonlinear transformations, *Proc. Eur. Signal Process. Conf.*, pp. 2120–2124. Amsterdam, The Netherlands, January 18–22, 2021.

[15] P. L. Combettes and Z. C. Woodstock, A variational inequality model for the construction of signals from inconsistent nonlinear equations, submitted.

[16] P. L. Combettes and Z. C. Woodstock, Reconstruction of functions from prescribed proximal points, submitted.

[17] A. Dahimene, M. Noureddine, and A. Azrar, A simple algorithm for the restoration of clipped speech signal, *Informatica*, vol. 32, pp. 183–188, 2008.

[18] C. De Boor, On "best" interpolation, *J. Approx. Theory*, vol. 16, pp. 28–42, 1976.

[19] B. Defraene et al., Declipping of audio signals using perceptual compressed sensing, *IEEE Trans. Audio, Speech, and Lang. Process.*, vol. 21, pp. 2627–2637, 2013.

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

[21] F. Deutsch, W. Li, and J. D. Ward, Best approximation from the intersection of a closed convex set and a polyhedron in Hilbert space, weak Slater conditions, and the strong conical hull intersection property, *SIAM J. Optim.*, vol. 10, pp. 252–268, 1999.

[22] R. A. DeVore, B. Jawerth, and V. Popov, Compression of wavelet decompositions, *Amer. J. Math.*, vol. 114, pp. 737–785, 1992.

[23] S. J. Dilworth, N. J. Kalton, D. Kutzarova, and V. N. Temlyakov, The thresholding greedy algorithm, greedy bases, and duality, *Constr. Approx.*, vol. 19, pp. 575–597, 2003.

[24] J. Favard, Sur l'interpolation, *J. Math. Pures. Appl.*, vol. 19, pp. 281–306, 1940.

[25] S. Foucart and T. Needham, Sparse recovery from saturated measurements, *Inf. Inference*, vol. 6, pp. 196–212, 2017.

[26] S. Foucart and T. Needham, Sparse recovery from inaccurate saturated measurements, *Acta Appl. Math.*, vol. 158, pp. 49–66, 2018.

[27] L. Gosse, A Donoho-Stark criterion for stable signal recovery in discrete wavelet subspaces, *J. Comput. Appl. Math.*, vol. 235, pp 5024–5039, 2011.

[28] P. E. Greenwood, U. U. Müller, and L. M. Ward, Soft threshold stochastic resonance, *Phys. Rev. E*, vol. 70, pp. 51110–51120, 2004.

[29] R. Gribonval and M. Nielsen, Nonlinear approximation with dictionaries I. Direct estimates, *J. Fourier Anal. Appl.*, vol. 10, pp. 51–71, 2004.

[30] R. Gribonval and M. Nielsen, On approximation with Spline generated framelets, *Constr. Approx.*, vol. 20, pp. 207–232, 2004.

[31] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*. SIAM, Philadelphia, 1995.

[32] C. T. Kelley, Numerical methods for nonlinear equations, *Acta Numer.*, vol. 27, pp. 207–287, 2018.

[33] H. J. Landau and W. L. Miranker, The recovery of distorted band-limited signals, *J. Math. Anal. Appl.*, vol. 2, pp. 97–104, 1961.

[34] Z. Li, C. A. Micchelli, and Y. Xu, Fixed-point proximity algorithm for minimal norm interpolation, *Appl. Comput. Harm. Anal.*, vol. 49, pp. 328–342, 2020.

[35] A. Marmin, A. Jezierska, M. Castella, and J.-C. Pesquet, Global optimization for recovery of clipped signals corrupted with Poisson-Gaussian noise, *IEEE Signal Process. Lett.*, vol. 27, pp. 970–974, 2020.

[36] J. M. Melenk and G. Zimmermann, Functions with time and frequency gaps, *J. Fourier Anal. Appl.*, vol. 2, pp. 611–614, 1996.

[37] W. D. Montgomery, Optical applications of Von Neumann's alternating-projection theorem, *Optics Lett.*, vol. 7, pp. 1–3, 1982.

[38] A. Pinkus, On smoothest interpolants, *SIAM J. Math. Anal.*, vol. 19, pp. 1431–1441, 1988.

[39] L. Rencker, F. Bach, W. Wang, and M. D. Plumbley, Sparse recovery and dictionary learning from nonlinear compressive measurements, *IEEE Trans. Signal Process.*, vol. 67, pp. 5659–5670, 2019.

[40] N. N. Reyes and L. J. D. Vallejo, Global growth of band-limited local approximations, *J. Math. Anal. Appl.*, vol. 400, pp. 418–424, 2013.

[41] A. P. Ruszcyński, *Nonlinear Programming*. Princeton University Press, New Jersey, 2006.

[42] D. Rzepka, M. Miśkowicz, D. Kościelnik, and N. T. Thao, Reconstruction of signals from level-crossing samples using implicit information, *IEEE Access*, vol. 6, pp. 35001–35011, 2018.

[43] M. I. Sezan and H. Stark, Incorporation of a priori movement information into signal recovery and synthesis problems, *J. Math. Anal. Appl.*, vol. 122, pp. 172–186, 1987.

[44] A. Shapiro, Directionally nondifferentiable metric projection, *J. Optim. Theory Appl.*, vol. 81, pp. 203–204, 1994.

[45] M. Soltani and C. Hedge, Fast algorithms for demixing sparse signals from nonlinear observations, *IEEE Trans. Signal Process.*, vol. 65, pp. 4209–4222, 2017.

[46] V. N. Temlyakov, The best $m$-term approximation and greedy algorithms, *Adv. Comput. Math.*, vol. 8, pp. 249–265, 1998.

[47] T. Teshima, M. Xu, I. Sato, and M. Sugiyama, Clipped matrix completion: A remedy for ceiling effects, *Proc. AAAI Conf. Artif. Intell.*, pp. 5151–5158, 2019.

[48] N. T. Thao and M. Vetterli, Deterministic analysis of oversampled A/D conversion and decoding improvement based on consistent estimates, *IEEE Trans. Signal Process.*, vol. 42, pp. 519–531, 1994.

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

[50] P. Záviška, P. Rajmic, A. Ozerov, and L. Rencker, A survey and an extensive evaluation of popular audio declipping methods, *IEEE J. Sel. Top. Signal Process.*, vol. 15, pp. 5–24, 2021.

Chapter 2

# A FIXED POINT FRAMEWORK FOR RECOVERING SIGNALS FROM NONLINEAR TRANSFORMATIONS

## 2.1 Introduction and context

This chapter solves the feasibility Problem 1.1.1 and demonstrates its applications in concrete problems from signal processing. The methods introduced here are further generalized in the setting of best-approximation problems in Chapter 3. The inconsistent formulation presented in Section 2.2.4 is later generalized in Chapter 4.

This chapter presents the following article.

> P. L. Combettes and Z. C. Woodstock, A fixed point framework for recovering signals from nonlinear transformations, *Proceedings of the 2020 European Signal Processing Conference*, pp. 2120–2124. Amsterdam, The Netherlands, January 18–22, 2021.

## 2.2 Article: A fixed point framework for recovering signals from nonlinear transformations

**Abstract**. We consider the problem of recovering a signal from nonlinear transformations, under convex constraints modeling *a priori* information. Standard feasibility and optimization methods are ill-suited to tackle this problem due to the nonlinearities. We show that, in many common applications, the transformation model can be associated with fixed point equations involving firmly nonexpansive operators. In turn, the recovery problem is reduced to a tractable common fixed point formulation, which is solved efficiently by a provably convergent, block-iterative

algorithm. Applications to signal and image recovery are demonstrated. Inconsistent problems are also addressed.

### 2.2.1 Introduction

Under consideration is the general problem of recovering an original signal $\overline{x}$ in a Euclidean space $\mathcal{H}$ from a finite number of transformations $(r_k)_{k\in K}$ of the form

$$r_k = R_k\overline{x} \in \mathcal{G}_k, \tag{2.1}$$

where $R_k\colon \mathcal{H} \to \mathcal{G}_k$ is an operator mapping the solution space $\mathcal{H}$ to the Euclidean space $\mathcal{G}_k$. In addition to these transformations, some *a priori* constraints on $\overline{x}$ are available in the form of a finite family of closed convex subsets $(C_j)_{j\in J}$ of $\mathcal{H}$ [4, 14, 17, 18, 20]. Altogether, the recovery problem is to

$$\text{find } \; x \in \bigcap_{j\in J} C_j \; \text{ such that } \; (\forall k \in K) \quad R_k x = r_k. \tag{2.2}$$

One of the most classical instances of this formulation was proposed by Youla in [19], namely

$$\text{find } \; x \in V_1 \; \text{ such that } \; \operatorname{proj}_{V_2} x = r_2, \tag{2.3}$$

where $V_1$ and $V_2$ are vector subspaces of $\mathcal{H}$ and $\operatorname{proj}_{V_2}$ is the projection operator onto $V_2$. As shown in [19], (2.3) covers many basic signal processing problems, such as band-limited extrapolation or image reconstruction from diffraction data, and it can be solved with a simple alternating projection algorithm. The extension of (2.3) to recovery problems with several transformations modeled as linear projections $r_k = \operatorname{proj}_{V_k}\overline{x}$ is discussed in [9, 13]. More broadly, if the operators $(R_k)_{k\in K}$ are linear, reliable algorithms are available to solve (2.2). In particular, since the associated constraint set is an affine subspace with an explicit projection operator, standard feasibility algorithms can be used [4]. Alternatively, proximal splitting methods can be considered; see [6] and its references.

In the present paper we consider the general situation in which the operators $(R_k)_{k\in K}$ in (2.1) are not necessarily linear, a stark departure from common assumptions in signal recovery problems. Examples of such nonlinearly generated data $(r_k)_{k\in K}$ in (2.1) include hard-thresholded wavelet coefficients of $\overline{x}$, the positive part of the Fourier transform of $\overline{x}$, a mixture of best approximations of $\overline{x}$ from closed convex sets, a maximum a posteriori denoised version of $\overline{x}$, or measurements of $\overline{x}$ acquired through nonlinear sensors.

A significant difficulty one faces in the nonlinear context is that the constraint (2.1) is typically not representable by an exploitable convex constraint; see, e.g., [2, 3]. As a result, finding a solution to (2.2) with a provenly convergent and numerically efficient algorithm is a challenging task. In particular, standard convex feasibility algorithms are not applicable. Furthermore, variational relaxations involving a penalty of the type $\sum_{k\in K} \phi_k(\|R_k x - r_k\|)$

typically lead to nonconvex problems, even for choices as basic as $\phi_k = |\cdot|^2$ and $R_k$ taken as the projection operator onto a closed convex set.

Our strategy to solve (2.2) is to forego the feasibility and optimization approaches in favor of the flexible and unifying framework of fixed point theory. Our first contribution is to show that, while $R_k$ in (2.1) may be a very badly conditioned (possibly discontinuous) operator, common transformation models can be reformulated as fixed point equations with respect to an operator with much better properties, namely a firmly nonexpansive operator. Next, using a suitable modeling of the constraint sets $(C_j)_{j\in J}$, we rephrase (2.2) as an equivalent common fixed point problem and solve it with a reliable and efficient extrapolated block-iterative fixed point algorithm. This strategy is outlined in Section 2.2.2, where we also provide the algorithm. In Section 2.2.3, we present several numerical illustrations of the proposed framework to nonlinear signal and image recovery. Finally, inconsistent problems are addressed in Section 2.2.4.

### 2.2.2 Fixed point model and algorithm

For background on the tools from fixed point theory and convex analysis used in this section, we refer the reader to [1]. Let us first recall that an operator $T\colon \mathcal{H} \to \mathcal{H}$ is firmly nonexpansive if

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \|Tx - Ty\|^2 \leqslant \|x - y\|^2 - \|(\mathrm{Id}\ -T)x - (\mathrm{Id}\ -T)y\|^2, \tag{2.4}$$

and firmly quasinonexpansive if

$$(\forall x \in \mathcal{H})(\forall y \in \operatorname{Fix} T) \quad \langle y - Tx \mid x - Tx\rangle \leqslant 0, \tag{2.5}$$

where $\operatorname{Fix} T = \big\{x \in \mathcal{H} \;\big|\; Tx = x\big\}$. Finally, the subdifferential of a convex function $f\colon \mathcal{H} \to \mathbb{R}$ at $x \in \mathcal{H}$ is

$$\partial f(x) = \big\{u \in \mathcal{H} \,\big|\, (\forall y \in \mathcal{H})\ \langle y - x \mid u\rangle + f(x) \leqslant f(y)\big\}. \tag{2.6}$$

As discussed in Section 2.2.1, the transformation model (2.1) is too general to make finding a solution to (2.2) via a provenly convergent method possible. We therefore assume the following.

**Assumption 2.2.1** *The problem (2.2) has at least one solution, $J \cap K = \varnothing$, and the following hold:*

(i) *For every $k \in K$, $S_k\colon \mathcal{G}_k \to \mathcal{H}$ is an operator such that $S_k \circ R_k$ is firmly nonexpansive and*

$$\Big(\forall x \in \bigcap_{j\in J} C_j\Big)\ S_k(R_k x) = S_k r_k \ \Rightarrow\ R_k x = r_k. \tag{2.7}$$

(ii) *For every $j \in J_1 \subset J$, the operator $\operatorname{proj}_{C_j}$ is easily implementable.*

(iii) *For every $j \in J \smallsetminus J_1$, $f_j\colon \mathcal{H} \to \mathbb{R}$ is a convex function such that $C_j = \big\{x \in \mathcal{H} \;\big|\; f_j(x) \leqslant 0\big\}$.*

In view of Assumption 2.2.1(i), let us replace (2.2) by the equivalent problem

$$\text{find } x \in \bigcap_{j\in J} C_j \text{ such that } (\forall k \in K)\ S_k(R_k x) = S_k r_k. \tag{2.8}$$

Concrete examples of suitable operators $(S_k)_{k\in K}$ will be given in Section 2.2.3 (see also [10]). The motivation behind (2.8) is that it leads to a tractable fixed point formulation. To see this, set

$$(\forall k \in K) \quad T_k = S_k r_k + \mathrm{Id} - S_k \circ R_k \tag{2.9}$$

and let $x \in \bigcap_{j\in J} C_j$. Then, for every $k \in K$, (2.1) $\Leftrightarrow S_k(R_k x) = S_k r_k \Leftrightarrow x = S_k r_k + x - S_k(R_k x)$ $\Leftrightarrow x \in \operatorname{Fix} T_k$. A key observation at this point is that (2.4) implies that the operators $(T_k)_{k\in K}$ are firmly nonexpansive, hence firmly quasinonexpansive.

If $j \in J_1$, per Assumption 2.2.1(ii), the set $C_j$ will be activated in the algorithm through the use of the operator $T_j = \operatorname{proj}_{C_j}$, which is firmly nonexpansive [1, Proposition 4.16]. On the other hand, if $j \in J \smallsetminus J_1$, the convex inequality representation of Assumption 2.2.1(iii) will lead to an activation of $C_j$ through its subgradient projector. Recall that the subgradient projection of $x \in \mathcal{H}$ onto $C_j$ relative to $u_j \in \partial f_j(x)$ is

$$T_j x = \begin{cases} x - \dfrac{f_j(x)}{\|u_j\|^2} u_j, & \text{if } \ f_j(x) > 0; \\ x, & \text{if } \ f_j(x) \leqslant 0, \end{cases} \tag{2.10}$$

and that $T_j$ is firmly quasinonexpansive, with $\operatorname{Fix} T_j = C_j$ [1, Proposition 29.41]. The advantage of the subgradient projector onto $C_j$ is that, unlike the exact projector, it does not require solving a nonlinear best approximation problem, which makes it much easier to implement in the presence of convex inequality constraints [5]. Altogether, (2.2) is equivalent to the common fixed point problem

$$\text{find } \ x \in \bigcap_{i\in J\cup K} \operatorname{Fix} T_i, \tag{2.11}$$

where each $T_i$ is firmly quasinonexpansive. This allows us to solve (2.2) as follows.

**Theorem 2.2.2** [10] *Consider the setting of problem* (2.2) *under Assumption 2.2.1. Let* $x_0 \in \mathcal{H}$, *let* $0 < \varepsilon < 1/\operatorname{card}(J \cup K)$, *and set* $(\forall k \in K)\ p_k = S_k r_k$ *and* $F_k = S_k \circ R_k$. *Iterate*

$$
\begin{array}{l}
\textit{for}\ \ n = 0, 1, \ldots \\
\quad \varnothing \neq I_n \subset J \cup K \\
\quad \{\omega_{i,n}\}_{i \in I_n} \subset [\varepsilon, 1],\ \sum_{i \in I_n} \omega_{i,n} = 1 \\
\quad \textit{for every}\ \ i \in I_n \\
\quad\quad \textit{if}\ \ i \in J_1 \\
\quad\quad\quad y_{i,n} = \mathrm{proj}_{C_i} x_n - x_n \\
\quad\quad \textit{if}\ \ i \in J \smallsetminus J_1 \\
\quad\quad\quad u_{i,n} \in \partial f_i(x_n) \\
\quad\quad\quad y_{i,n} = \begin{cases} -\dfrac{f_i(x_n)}{\|u_{i,n}\|^2} u_{i,n} & \textit{if}\ f_i(x_n) > 0 \\ 0, & \textit{if}\ f_i(x_n) \leqslant 0 \end{cases} \\
\quad\quad \textit{else} \\
\quad\quad\quad y_{i,n} = p_i - F_i x_n \\
\quad\quad\quad \nu_{i,n} = \|y_{i,n}\| \\
\quad \nu_n = \sum_{i \in I_n} \omega_{i,n} \nu_{i,n}^2 \\
\quad \textit{if}\ \nu_n = 0 \\
\quad\quad x_{n+1} = x_n \\
\quad \textit{else} \\
\quad\quad y_n = \sum_{i \in I_n} \omega_{i,n} y_{i,n} \\
\quad\quad \Lambda_n = \nu_n / \|y_n\|^2 \\
\quad\quad \lambda_n \in [\varepsilon, (2 - \varepsilon)\Lambda_n] \\
\quad\quad x_{n+1} = x_n + \lambda_n y_n.
\end{array}
\tag{2.12}
$$

*Suppose that there exists an integer $M > 0$ such that*

$$
(\forall n \in \mathbb{N}) \quad \bigcup_{m=0}^{M-1} I_{n+m} = J \cup K. \tag{2.13}
$$

*Then $(x_n)_{n \in \mathbb{N}}$ converges to a solution to* (2.2).

When $K = \varnothing$, (2.12) coincides with the extrapolated method of parallel subgradient projections (EMOPSP) of [5]. It has in addition the ability to incorporate the constraints (2.1), while maintaining the attractive features of EMOPSP. First, it can process the operators in blocks of variable size. The control scheme (2.13) just imposes that every operator be activated at least once within any $M$ consecutive iterations. Second, because the extrapolation parameters $(\Lambda_n)_{n \in \mathbb{N}}$ can attain large values in $[1, +\infty[$, large steps are possible, which lead to fast convergence compared to standard relaxation schemes, where $\Lambda_n \equiv 1$.

### 2.2.3 Applications

We illustrate several instances of (2.2), develop tractable reformulations of the form (2.8), and solve them using (2.12), where $x_0 = 0$ and the relaxation strategy is that recommended

in [4, Chapter 5], namely

$$(\forall n \in \mathbb{N}) \quad \lambda_n = \begin{cases} \Lambda_n/2, & \text{if } \ n = 0 \mod 3; \\ 1.99\Lambda_n, & \text{otherwise.} \end{cases} \tag{2.14}$$

#### 2.2.3.1 Restoration from distorted signals

The goal is to recover the original form of the $N$-point ($N = 2048$) signal $\overline{x}$ from the following (see Fig. 2.1):

- A bound $\gamma_1$ on the energy of the finite differences of $\overline{x}$, namely $\|D\overline{x}\| \leqslant \gamma_1$, where $D\colon (\xi_i)_{i\in\{0,\dots,N-1\}} \mapsto (\xi_{i+1} - \xi_i)_{i\in\{0,\dots,N-2\}}$. The bound is given from prior information as $\gamma_1 = 1.17$.

- A distortion $r_2 = R_2\overline{x}$, where $R_2$ clips componentwise to $[-\gamma_2, \gamma_2]$ ($\gamma_2 = 0.1$) [16, Section 10.5].

- A distortion $r_3 = R_3\overline{x}$ of a low-pass version of $\overline{x}$, where $R_3 = Q_3 \circ L_3$. Here $L_3$ bandlimits by zeroing all but the $83$ lowest-frequency coefficients of the Discrete Fourier Transform, and $Q_3$ induces componentwise distortion via the operator [16, Section 10.6] $\theta_3 = (2/\pi)\arctan(\gamma_3\,\cdot)$, where $\gamma_3 = 10$ (see Fig. 2.2).

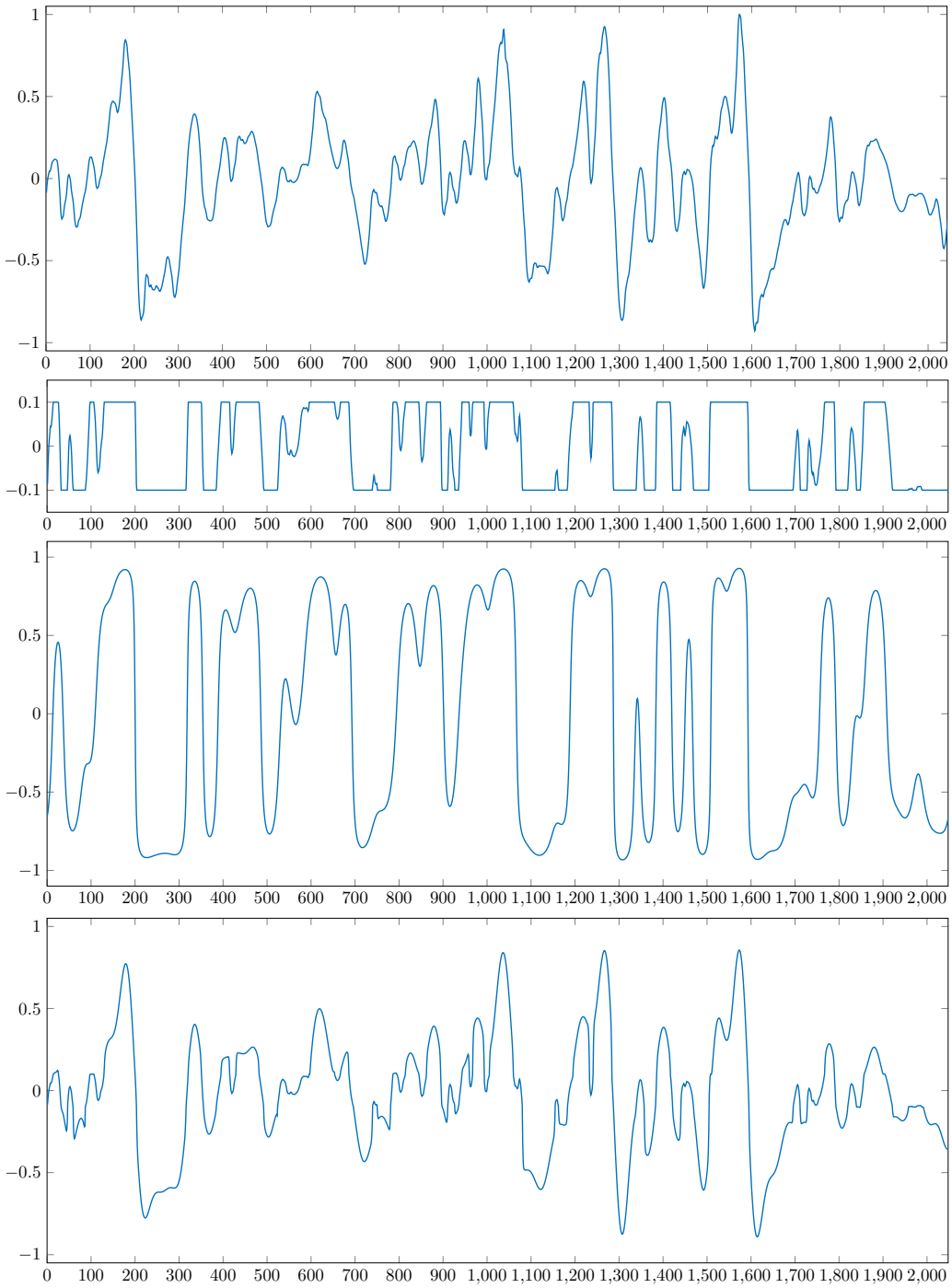



**Figure 2.1** Signals in Section 2.2.3.1. Top to bottom: original signal $\overline{x}$, distorted signal $r_2$, distorted signal $r_3$, recovered signal.

The solution space is the standard Euclidean space $\mathcal{H} = \mathbb{R}^N$. To formulate the recovery problem as an instance of (2.2), set $J = \{1\}$, $J_1 = \varnothing$, $K = \{2, 3\}$, and $C_1 = \big\{x \in \mathcal{H} \;\big|\; f_1(x) \leqslant 0\big\}$, where

$f_1 = \|D \cdot\| - \gamma_1$. Then the objective is to

$$\text{find} \;\; x \in C_1 \;\; \text{such that} \;\; R_2 x = r_2 \;\; \text{and} \;\; R_3 x = r_3. \tag{2.15}$$

Next, let us verify that Assumption 2.2.1(i) is satisfied. On the one hand, since $R_2$ is the projection onto the closed convex set $[-\gamma_2, \gamma_2]^N$, it is firmly nonexpansive, so we set $S_2 = \mathrm{Id}$. On the other hand, if we set $S_3 = \gamma_3^{-1} L_3$, then $S_3 \circ R_3$ is firmly nonexpansive and satisfies (2.7) [10]. We thus obtain an instance of (2.8), to which we apply (2.12) with (2.14) and $(\forall n \in \mathbb{N})\; I_n = J \cup K$ and $(\forall i \in I_n)\; \omega_{i,n} = 1/3$. The recovered signal shown in Fig. 2.1 effectively incorporates the information from the prior constraint and the nonlinear distortions.

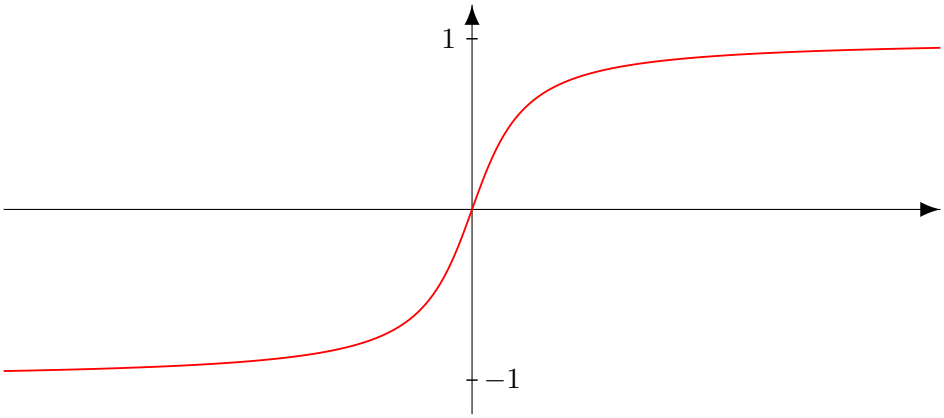



**Figure 2.2** Distortion operator $\theta_3$ in Section 2.2.3.1.

#### 2.2.3.2 Reconstruction from thresholded scalar products

The goal is to recover the original form of the $N$-point ($N = 1024$) signal $\overline{x}$ shown in Fig. 2.3 from thresholded scalar products $(r_k)_{k \in K}$ given by

$$(\forall k \in K) \quad r_k = R_k \overline{x}, \quad \text{with} R_k \colon \mathcal{H} \to \mathbb{R} \colon x \mapsto Q_\gamma \langle x \mid e_k \rangle, \tag{2.16}$$

where

- $(e_k)_{k \in K}$ is a collection of normalized vectors in $\mathbb{R}^N$ with zero-mean i.i.d. entries.
- $Q_\gamma$ ($\gamma = 0.05$) is the thresholding operator

$$Q_\gamma \colon \xi \mapsto \begin{cases} \operatorname{sign}(\xi) \sqrt{\xi^2 - \gamma^2}, & \text{if } \; |\xi| > \gamma; \\ 0, & \text{if } \; |\xi| \leqslant \gamma \end{cases} \tag{2.17}$$

  of [15] (see Fig. 2.4).

- $K = \{1, \dots, m\}$, where $m = 1200$.

The solution space $\mathcal{H}$ is the standard Euclidean space $\mathbb{R}^N$, and (2.16) gives rise to the special case of (2.2)

$$\text{find} \quad x \in \mathcal{H} \quad \text{such that} \quad (\forall k \in K) \quad r_k = Q_\gamma \langle x \mid e_k \rangle, \tag{2.18}$$

in which $J = \varnothing$.

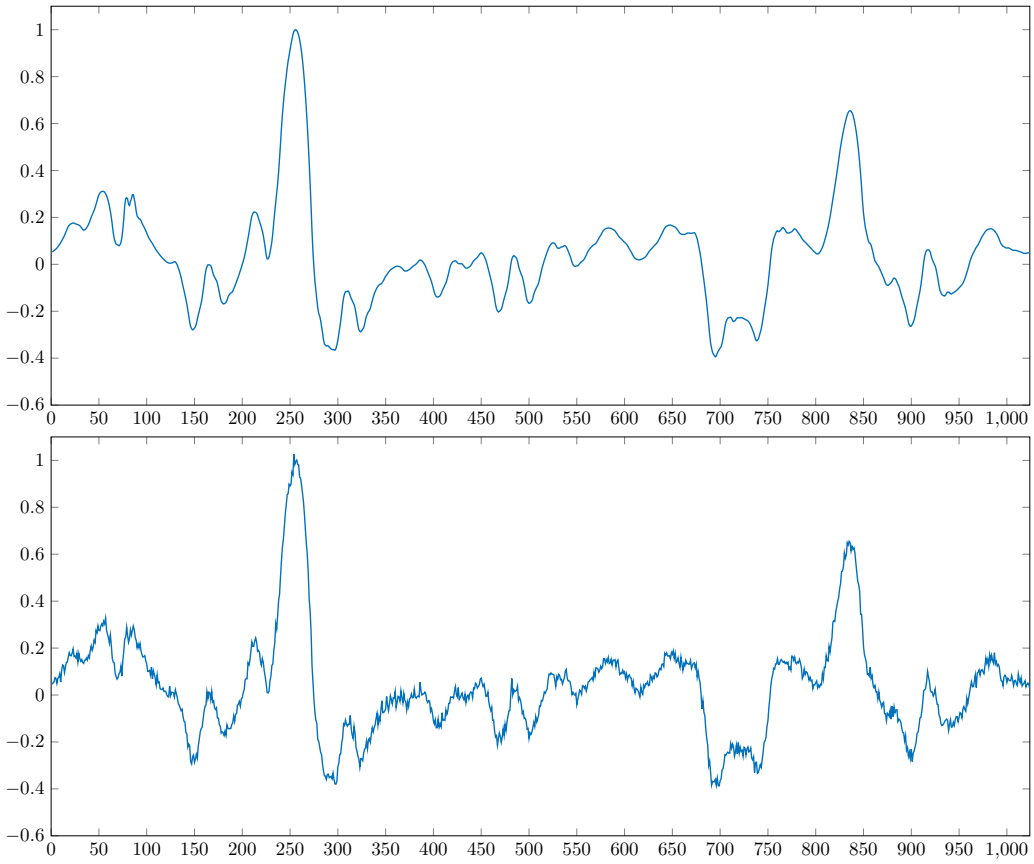



**Figure 2.3** Original signal $\overline{x}$ (top) and recovery (bottom) in Section 2.2.3.2.

Note that the standard soft-thresholder on $[-\gamma, \gamma]$ can be written as

$$\text{soft}_\gamma : \xi \mapsto \text{sign}\,(Q_\gamma \xi) \left( \sqrt{(Q_\gamma \xi)^2 + \gamma^2} - \gamma \right). \tag{2.19}$$

To formulate (2.8) we set, for every $k \in K$,

$$S_k \colon \mathbb{R} \to \mathcal{H} \colon \xi \mapsto \text{sign}\,(\xi) \left( \sqrt{\xi^2 + \gamma^2} - \gamma \right) e_k, \tag{2.20}$$

which fulfills Assumption 2.2.1(i) and yields $S_k \circ R_k = (\,\text{soft}_\gamma \langle \cdot \mid e_k \rangle) e_k$ [10]. We apply (2.12)

with (2.14) and the following control scheme. We split $K$ into 12 blocks of 100 consecutive indices, and select $I_n$ by periodically sweeping through the blocks, hence satisfying (2.13) with $M = 12$. Moreover, $\omega_{i,n} \equiv 1/100$. The reconstructed signal shown in Fig. 2.3 illustrates the ability of the proposed approach to effectively exploit nonlinearly generated data.

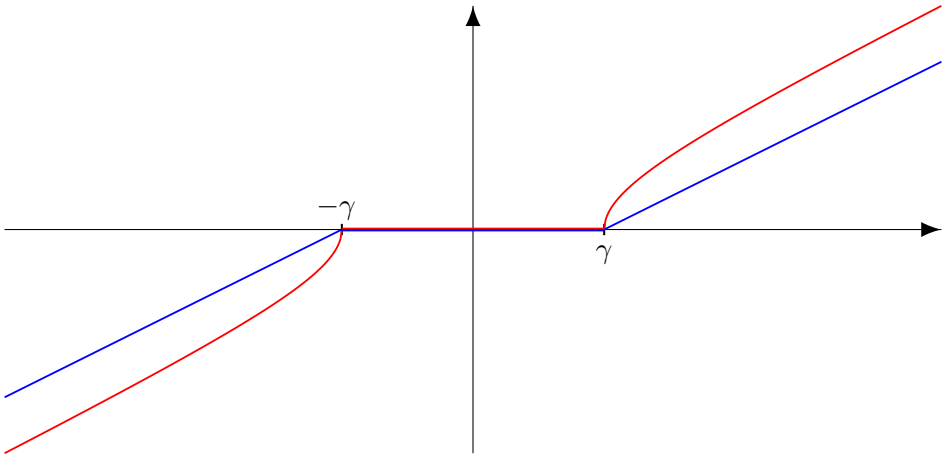



**Figure 2.4** The thresholder (2.17) of [15] (red) and the soft thresholder (blue) used in Section 2.2.3.2.

#### 2.2.3.3 Image recovery

The goal is to recover the $N \times N$ ($N = 256$) image $\overline{x}$ from the following (see Fig. 2.5):

- The Fourier phase $\angle\mathrm{DFT}\,(\overline{x})$ ($\mathrm{DFT}\,(\overline{x})$ denotes the 2D Discrete Fourier Transform of $\overline{x}$).
- The pixel values of $\overline{x}$ reside in $[0, 255]$.
- An upper bound $\gamma_3$ on the total variation $\mathrm{tv}(\overline{x})$ [8]. In this experiment, $\gamma_3 = 1.2\,\mathrm{tv}(\overline{x}) = 1.10 \times 10^6$.
- A compressed representation $r_4 = R_4\overline{x}$. Here, $R_4 = Q_4 \circ W$, where $W$ is the 2D Haar wavelet transform and $Q_4$ performs componentwise hard-thresholding via ($\rho = 325$)

$$(\forall \xi \in \mathbb{R}) \quad \mathrm{hard}_\rho\, \xi = \begin{cases} \xi, & \text{if } |\xi| > \rho; \\ 0, & \text{if } |\xi| \leqslant \rho. \end{cases} \tag{2.21}$$

- A down-sampled blurred image $r_5 = R_5\overline{x}$. Here $R_5 = Q_5 \circ H_5$, where the linear operator $H_5 \colon \mathbb{R}^{N\times N} \to \mathbb{R}^{N\times N}$ convolves with a $5 \times 5$ Gaussian kernel with variance 1, and $Q_5 \colon \mathbb{R}^{N\times N} \to \mathbb{R}^{8\times 8}$ maps the average of each of the 64 disjoint $32 \times 32$ blocks of an $N \times N$ image to a representative pixel in an $8 \times 8$ image [12].

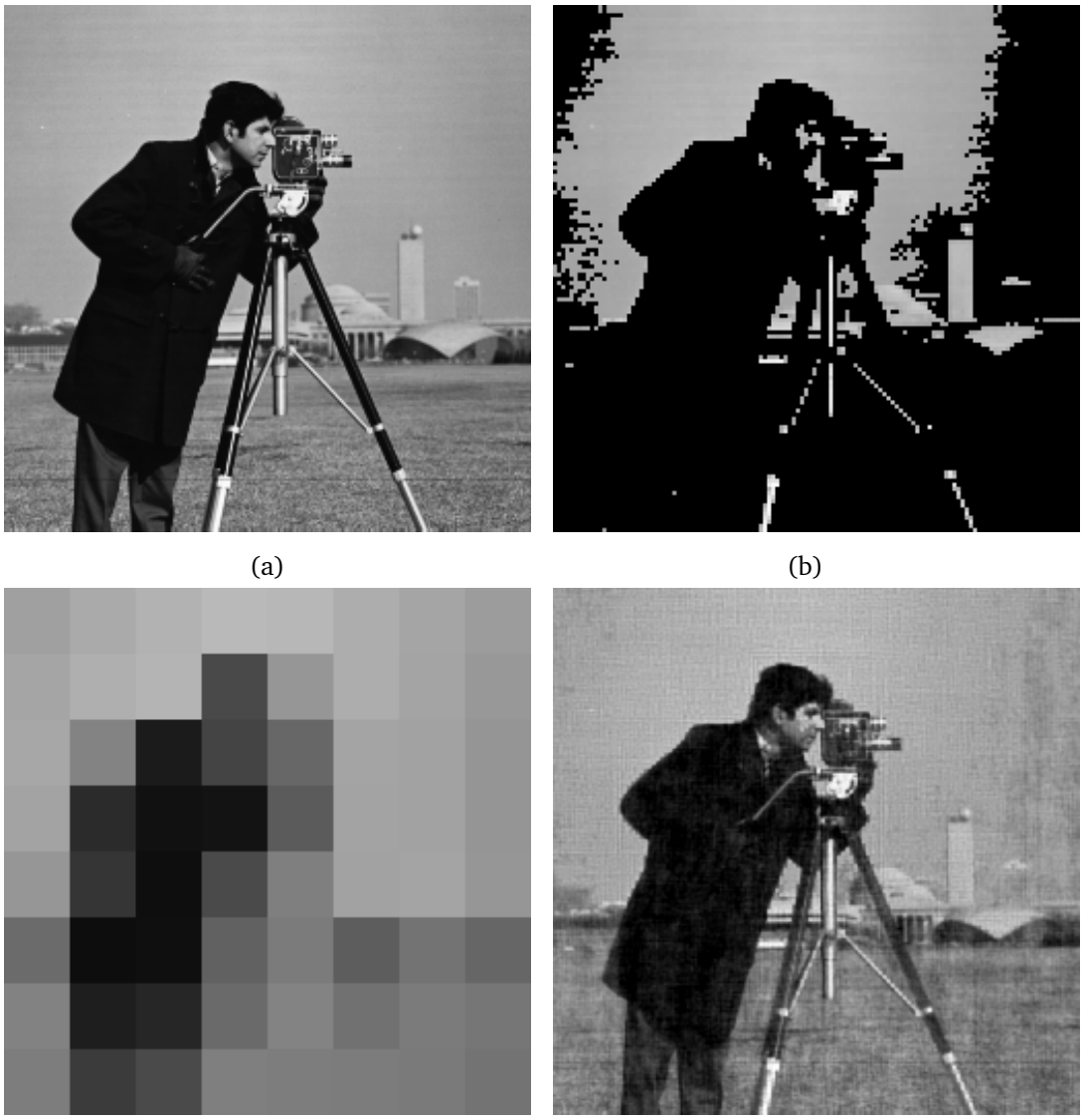
The solution space is $\mathcal{H} = \mathbb{R}^{N\times N}$ equipped with the Frobenius norm $\|\cdot\|$. To cast the recovery task as an instance of (2.2), we set $J = \{1, 2, 3\}$, $J_1 = \{1, 2\}$, $K = \{4, 5\}$, $C_1 =$

$\big\{x \in \mathcal{H} \;\big|\; \angle\mathrm{DFT}\,(x) = \angle\mathrm{DFT}\,(\overline{x})\big\}$, $C_2 = [0,255]^{N\times N}$, $f_3 = \mathrm{tv} - \gamma_3$, and $C_3 = \big\{x \in \mathcal{H} \;\big|\; f_3(x) \leqslant 0\big\}$. Expressions for $\mathrm{proj}_{C_1}$ and $\partial f_3$ are provided in [11] and [8], respectively. The objective is to

$$\text{find} \;\; x \in \bigcap_{j=1}^{3} C_j \;\; \text{such that} \;\; \begin{cases} R_4 x = r_4; \\ R_5 x = r_5. \end{cases} \tag{2.22}$$

Let us verify that Assumption 2.2.1(i) holds. For every $\xi \in \mathbb{R}$,

$$\mathrm{soft}_\rho\,\xi = \mathrm{hard}_\rho\,\xi + \begin{cases} -\rho, & \text{if} \;\; \mathrm{hard}_\rho\,\xi > \rho; \\ 0, & \text{if} \;\; -\rho \leqslant \mathrm{hard}_\rho\,\xi \leqslant \rho; \\ \rho, & \text{if} \;\; \mathrm{hard}_\rho\,\xi < -\rho. \end{cases} \tag{2.23}$$



(a) (b)

(c) (d)

**Figure 2.5** Images from Section 2.2.3.3. (a) Original image $\overline{x}$. (b) Compressed image $W^* r_4$. (c) Downsampled $8 \times 8$ image $r_5$. (d) Recovered image.

We construct $S_4$ such that $S_4 \circ R_4 = W^{-1} \circ T \circ W$, where $T$ applies $\text{soft}_\rho$ componentwise. In turn, recalling that $r_4$ is the result of hard-thresholding, $S_4 r_4$ is built by first adding the quantity on the right-hand side of (2.23) to $r_4$ componentwise, and then applying the inverse Haar transform. This guarantees that $S_4$ satisfies Assumption 2.2.1(i) [10]. Next, we let $D_5 \subset \mathcal{H}$ be the subspace of $32 \times 32$-block-constant matrices and construct an operator $S_5$ satisfying Assumption 2.2.1(i) and the identity $S_5 \circ R_5 = H_5 \circ \text{proj}_{D_5} \circ H_5$ [10]. In turn, $S_5 r_5 = H_5 s_5$, where $s_5 \in D_5$ is built by repeating each pixel value of $r_5$ in the block it represents. We thus arrive at an instance of

(2.8), which we solve using (2.12) with (2.14) and

$$(\forall n \in \mathbb{N}) \quad I_n = J \cup K \text{ and } (\forall i \in I_n) \quad \omega_{i,n} = 1/5. \tag{2.24}$$

The resulting image displayed in Fig. 2.5(d) shows that our framework makes it possible to exploit the information from the three prior constraints and from the transformations $r_4$ and $r_5$ to obtain a quality recovery.

### 2.2.4 Inconsistent problems

Inaccuracies and unmodeled dynamics may cause (2.2) to admit no solution. In such instances, we propose the following relaxation for (2.2) [10].

**Assumption 2.2.3** *For every $j \in J$, the operator $\text{proj}_{C_j}$ is easily implementable and, for every $k \in K$, Assumption 2.2.1(i) holds. In addition, $\{\omega_j\}_{j\in J} \subset ]0,1]$ and $\{\omega_k\}_{k\in K} \subset ]0,1]$ satisfy $\sum_{j\in J} \omega_j + \sum_{k\in K} \omega_k = 1$.*

Under Assumption 2.2.3, the goal is to

$$\text{find} \;\; x \in \mathcal{H} \;\; \text{such that} \sum_{j\in J} \omega_j (x - \text{proj}_{C_j} x) + \sum_{k\in K} \omega_k (S_k R_k x - S_k r_k) = 0. \tag{2.25}$$

When $K = \varnothing$, the solutions of (2.25) are the minimizers of the least squared-distance proximity function $\sum_{j\in J} \omega_j d^2_{C_j}$ [4]. If (2.2) does have solutions, then it is equivalent to (2.25). The algorithm of [7] can be used to solve (2.25) block-iteratively.

## References


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

[2] T. Blumensath, "Compressed sensing with nonlinear observations and related nonlinear optimization problems," *IEEE Trans. Inform. Theory,* vol. 59, pp. 3466–3474, 2013.

[3] M. Castella, J.-C. Pesquet, and A. Marmin, "Rational optimization for nonlinear reconstruction with approximate $\ell_0$ penalization," *IEEE Trans. Signal Process.,* vol. 67, pp. 1407–1417, 2019.

[4] P. L. Combettes, "The Convex Feasibility Problem in Image Recovery," in *Advances in Imaging and Electron Physics,* vol. 95, pp. 155–270. New York, NY: Academic Press, 1996.

[5] P. L. Combettes, "Convex set theoretic image recovery by extrapolated iterations of parallel subgradient projections," *IEEE Trans. Image Process.,* vol. 6, pp. 493–506, 1997.

[6] P. L. Combettes and J. Eckstein, "Asynchronous block-iterative primal-dual decomposition methods for monotone inclusions," *Math. Program.,* vol. B168, pp. 645–672, 2018.

[7] P. L. Combettes and L. E. Glaudin, "Solving composite fixed point problems with block updates," preprint, 2020.

[8] P. L. Combettes and J.-C. Pesquet, "Image restoration subject to a total variation constraint," *IEEE Trans. Image Process.,* vol. 13, pp. 1213–1222, 2004.

[9] P. L. Combettes and N. N. Reyes, "Functions with prescribed best linear approximations," *J. Approx. Theory,* vol. 162, pp. 1095–1116, 2010.

[10] P. L. Combettes and Z. C. Woodstock, "Reconstruction of functions from prescribed proximal points," preprint, 2020.

[11] A. Levi and H. Stark, "Signal reconstruction from phase by projection onto convex sets," *J. Opt. Soc. Amer.,* vol. 73, pp. 810–822, 1983.

[12] K. Nasrollahi and T. B. Moeslund, "Super-resolution: a comprehensive survey," *Mach. Vis. Appl.,* vol. 25, pp. 1423–1468, 2014.

[13] N. N. Reyes and L. J. D. Vallejo, "Global growth of band-limited local approximations," *J. Math. Anal. Appl.,* vol. 400, pp. 418–424, 2013.

[14] D. Rzepka, M. Miśkowicz, D. Kościelnik, and N. T. Thao, "Reconstruction of signals from level-crossing samples using implicit information," *IEEE Access,* vol. 6, pp. 35001–35011, 2018.

[15] T. Tao and B. Vidakovic, “Almost everywhere behavior of general wavelet shrinkage operators,” *Appl. Comput. Harmon. Anal.,* vol. 9, pp. 72–82, 2000.

[16] E. Tarr, *Hack Audio: An Introduction to Computer Programming and Digital Signal Processing in MATLAB.* New York, NY: Routledge, 2018.

[17] M. Tofighi, O. Yorulmaz, K. Köse, D. C. Yıldırım, R. Çetin-Atalay, and A. E. Çetin, “Phase and TV based convex sets for blind deconvolution of microscopic images,” *IEEE J. Selected Topics Signal Process.,* vol. 10, pp. 81–91, 2016.

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

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

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

Chapter 3

# RECONSTRUCTION OF FUNCTIONS FROM PRESCRIBED PROXIMAL POINTS

## 3.1 Introduction and context

The main contribution in this chapter is developing constructive techniques for modeling nonlinear equations with firmly nonexpansive operators. This approach is shown to capture many applications and leads to provenly-convergent solution algorithms. These techniques capture the specific instances of proxification in Chapter 2 as a special case. We also design a new algorithm for solving the best approximation Problem 1.1.3. As will be seen in Remark 3.2.25(ii), this chapter also extends the work of Chapter 2 by presenting a method for solving Problem 1.1.1 in the Hilbert space setting with a countable number of requirements.

This chapter presents the following article.

> P. L. Combettes and Z. C. Woodstock, Reconstruction of functions from prescribed proximal points, *Journal of Approximation Theory*, resubmitted with minor revisions.

## 3.2 Article: Reconstruction of functions from prescribed proximal points

*Dedicated to the memory of Noli N. Reyes (1963–2020)*

**Abstract**. Under investigation is the problem of finding the best approximation of a function in a Hilbert space subject to convex constraints and prescribed nonlinear transformations. We show that in many instances these prescriptions can be represented using firmly nonexpansive

operators, even when the original observation process is discontinuous. The proposed framework thus captures a large body of classical and contemporary best approximation problems arising in areas such as harmonic analysis, statistics, interpolation theory, and signal processing. The resulting problem is recast in terms of a common fixed point problem and solved with a new block-iterative algorithm that features approximate projections onto the individual sets as well as an extrapolated relaxation scheme that exploits the possible presence of affine constraints. A numerical application to signal recovery is demonstrated.

### 3.2.1 Introduction

Let $\mathcal{H}$ be a real Hilbert space with scalar product $\langle \cdot \mid \cdot \rangle$ and associated norm $\| \cdot \|$, let $x_0 \in \mathcal{H}$, let $U$ and $V$ be closed vector subspaces of $\mathcal{H}$ with projection operators $\mathrm{proj}_U$ and $\mathrm{proj}_V$, respectively, and let $p \in V$. The basic best approximation problem

$$\text{minimize } \|x - x_0\| \quad \text{subject to} \quad x \in U \quad \text{and} \quad \mathrm{proj}_V x = p \tag{3.1}$$

covers a wide range of scenarios in areas such as harmonic analysis, signal processing, interpolation theory, and optics [3, 22, 32, 35, 38, 40, 43, 52, 59]. In this setting, a function of interest $\overline{x} \in \mathcal{H}$ is known to lie in the subspace $U$ and its projection $p$ onto the subspace $V$ is known. The goal of (3.1) is then to find the best approximation to $x_0$ that is compatible with these two pieces of information. For example, band-limited extrapolation [49] aims at recovering a minimum energy band-limited function $\overline{x} \in \mathcal{H} = L^2(\mathbb{R})$ from the knowledge of its values on an interval $A$. This corresponds to the instance of (3.1) in which $x_0 = 0$, $V$ is the subspace of functions vanishing outside of $A$, $U$ is the subspace of functions with Fourier transform supported by a compact interval around the origin, and $p = 1_A \overline{x}$, where $1_A$ denotes the characteristic function of $A$. As shown in [59], if (3.1) is feasible (see [22] for necessary and sufficient conditions), then the sequence $(x_n)_{n \in \mathbb{N}}$ constructed by iterating

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = p + \mathrm{proj}_U x_n - \mathrm{proj}_V(\mathrm{proj}_U x_n) \tag{3.2}$$

converges strongly to its solution. The extension of (3.1) to finitely many vector subspaces $(U_j)_{j \in J}$ and $(V_k)_{k \in K}$ investigated in [22] is to

$$\text{minimize } \|x - x_0\| \quad \text{subject to } x \in \bigcap_{j \in J} U_j \text{ and } (\forall k \in K)\ \mathrm{proj}_{V_k} x = p_k, \text{ where } p_k \in V_k, \tag{3.3}$$

and it can be solved using affine projection methods. In many applications, the constraint sets [12–14, 17, 27, 30, 41, 48] or the operators yielding the prescribed values $(p_k)_{k \in K}$ [2, 7, 31, 39, 51, 57, 58] may not be linear. Our objective is to extend the linear formulation (3.3) by employing closed convex constraint subsets $(C_j)_{j \in J}$, together with prescriptions $(p_k)_{k \in K}$

resulting from nonlinear operators $(F_k)_{k\in K}$, i.e.,

$$\text{minimize } \|x - x_0\| \quad \text{subject to} \quad x \in \bigcap_{j\in J} C_j \quad \text{and} \quad (\forall k \in K) \quad F_k x = p_k. \tag{3.4}$$

In view of (3.3), projection operators onto closed convex sets constitute a natural class of candidates for the operators $(F_k)_{k\in K}$. For instance, in [51, 54, 58], $F_k$ is the projection operator onto a hypercube. However, many prescriptions $(p_k)_{k\in K}$ found in the literature, in particular those of [7, 31, 39, 57], do not reduce to best approximations from closed convex sets, and a more general formalism must be considered to represent them. A generalization of the notion of a best approximation was proposed by Moreau [44], who called the *proximal point* of $\overline{x} \in \mathcal{H}$ relative to a proper lower semicontinuous convex function $f_k \colon \mathcal{H} \to ]-\infty, +\infty]$ the unique minimizer $p_k \in \mathcal{H}$ of the function

$$y \mapsto f_k(y) + \frac{1}{2}\|\overline{x} - y\|^2, \tag{3.5}$$

and wrote $p_k = \operatorname{prox}_{f_k} \overline{x}$. This mechanism defines the proximity operator $\operatorname{prox}_{f_k} \colon \mathcal{H} \to \mathcal{H}$ of $f_k$. The case of a projector onto a nonempty closed convex set $D_k \subset \mathcal{H}$ is recovered by letting $f_k = \iota_{D_k}$, where

$$(\forall x \in \mathcal{H}) \quad \iota_{D_k}(x) = \begin{cases} 0, & \text{if } x \in D_k; \\ +\infty, & \text{if } x \notin D_k \end{cases} \tag{3.6}$$

is the indicator function of $D_k$. Proximity operators were initially motivated by applications in mechanics [9, 45, 47] and have become a central tool in the analysis and the numerical solution of numerous data processing tasks [21, 23]. We shall see later that they also model various nonlinear observation processes. The properties of proximity operators are detailed in [5, Chapter 24], among which is the fact that the operator $\operatorname{prox}_{f_k}$ can be expressed as the resolvent of the subdifferential of $f_k$, that is, $\operatorname{prox}_{f_k} = (\mathrm{Id} + \partial f_k)^{-1}$, where

$$(\forall x \in \mathcal{H}) \quad \partial f_k(x) = \big\{ u \in \mathcal{H} \;\big|\; (\forall y \in \mathcal{H}) \;\; \langle y - x \mid u \rangle + f_k(x) \leqslant f_k(y) \big\}. \tag{3.7}$$

As shown by Moreau [46], the set-valued operator $A_k = \partial f_k$ is maximally monotone, i.e.,

$$(\forall x \in \mathcal{H})(\forall u \in \mathcal{H}) \quad \big[\, u \in A_k x \quad \Leftrightarrow \quad (\forall y \in \mathcal{H})(\forall v \in A_k y) \quad \langle x - y \mid u - v \rangle \geqslant 0 \,\big]. \tag{3.8}$$

This property prompted Rockafellar [53] to generalize the notion of a proximal point as follows: given a maximally monotone set-valued operator $A_k \colon \mathcal{H} \to 2^{\mathcal{H}}$, the proximal point of $\overline{x} \in \mathcal{H}$ relative to $A_k$ is the unique point $p_k \in \mathcal{H}$ such that $\overline{x} - p_k \in A_k p_k$, i.e., $p_k = J_{A_k} \overline{x}$, where $J_{A_k} = (\mathrm{Id} + A_k)^{-1} \colon \mathcal{H} \to \mathcal{H}$ is the resolvent of $A_k$. As stated in [5, Corollary 23.9], a remarkable consequence of Minty's theorem [42] is that an operator $F_k \colon \mathcal{H} \to \mathcal{H}$ is the resolvent of a maximally monotone operator $A_k \colon \mathcal{H} \to 2^{\mathcal{H}}$ if and only if it is *firmly nonexpansive*, meaning that

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \|F_k x - F_k y\|^2 + \|(\mathrm{Id} - F_k)x - (\mathrm{Id} - F_k)y\|^2 \leqslant \|x - y\|^2. \tag{3.9}$$

In view of this equivalence, we call $p_k$ a *proximal point* of $\overline{x} \in \mathcal{H}$ relative to a firmly nonexpansive operator $F_k \colon \mathcal{H} \to \mathcal{H}$ if $p_k = F_k\overline{x}$. As we shall show in Section 3.2.2, firmly nonexpansive operators constitute a powerful device to represent a variety of nonlinear processes to generate the prescriptions $(p_k)_{k\in K}$ in (3.4). In light of these considerations, we propose to investigate the following nonlinear best approximation framework.

**Problem 3.2.1** Let $x_0 \in \mathcal{H}$ and let $J$ and $K$ be at most countable sets such that $J \cap K = \varnothing$ and $J \cup K \neq \varnothing$. For every $j \in J$, let $C_j$ be a closed convex subset of $\mathcal{H}$ and, for every $k \in K$, let $p_k \in \mathcal{H}$ and let $F_k \colon \mathcal{H} \to \mathcal{H}$ be a firmly nonexpansive operator. Suppose that there exists $\overline{x} \in \bigcap_{j\in J} C_j$ such that $(\forall k \in K)\ F_k\overline{x} = p_k$. The task is to

$$\text{minimize } \|x - x_0\| \quad \text{subject to} \quad x \in \bigcap_{j\in J} C_j \quad \text{and} \quad (\forall k \in K) \quad F_k x = p_k. \tag{3.10}$$

In Problem 3.2.1, the function of interest lies in the intersection of the sets $(C_j)_{j\in J}$, and its proximal points $(p_k)_{k\in K}$ relative to firmly nonexpansive operators $(F_k)_{k\in K}$ are prescribed. The objective is to obtain the best approximation to a function $x_0 \in \mathcal{H}$ from the set of functions which satisfy these properties.

As noted above, the numerical solution of the linear problem (3.3) is rather straightforward with existing projection techniques, while characterizing the existence of solutions for any choices of the prescribed values $(p_k)_{k\in K}$ – the so-called inverse best approximation property – is a more challenging task that was carried out in [22]. In the nonlinear setting, this property is of limited interest since it fails in simple scenarios [22, Remark 1.2]. Our objectives in the present paper are to demonstrate the far reach and the versatility of Problem 3.2.1, and to devise an efficient and flexible numerical method to solve it.

The remainder of the paper consists of four sections. In Section 3.2.2, we show the ability of our proximal point modeling to capture a variety of observation processes arising in practice, including some which result from discontinuous operators. In Section 3.2.3, we propose a new block-iterative algorithm to construct the best approximation to a reference point from a countable intersection of closed convex sets. The algorithm features approximate projections onto the individual sets as well as an extrapolated relaxation scheme that exploits the possible presence of affine subspaces in the constraint sets $(C_j)_{j\in J}$. In Section 3.2.4, Problem 3.2.1 is rephrased in terms of a common fixed point problem and the algorithm of Section 3.2.3 is used to solve it. A numerical illustration of our framework is presented in Section 3.2.5.

**Notation.** $\mathcal{H}$ is a real Hilbert space with scalar product $\langle \cdot \mid \cdot \rangle$, associated norm $\|\cdot\|$, and identity operator $\mathrm{Id}$. The family of all subsets of $\mathcal{H}$ is denoted by $2^{\mathcal{H}}$. The expressions $x_n \rightharpoonup x$ and $x_n \to x$ denote, respectively, the weak and the strong convergence of a sequence $(x_n)_{n\in\mathbb{N}}$ to $x$ in $\mathcal{H}$. The distance function to a subset $C$ of $\mathcal{H}$ is denoted by $d_C$. $\Gamma_0(\mathcal{H})$ is the class of all lower semicontinuous convex functions from $\mathcal{H}$ to $]-\infty, +\infty]$ which are proper in the sense that they are not identically $+\infty$. The conjugate of $f \in \Gamma_0(\mathcal{H})$ is denoted by $f^*$ and the infimal convolution operation by $\square$. The set of fixed points of an operator $T \colon \mathcal{H} \to \mathcal{H}$ is $\mathrm{Fix}\, T = \big\{x \in \mathcal{H} \;\big|\; Tx = x\big\}$.

The Hilbert direct sum of a family of real Hilbert spaces $(\mathsf{H}_i)_{i\in\mathbb{I}}$ is denoted by $\bigoplus_{i\in\mathbb{I}}\mathsf{H}_i$. For background on convex and nonlinear analysis, see [5].

### 3.2.2 Prescribed values as proximal points

We illustrate the fact that the proximal model adopted in Problem 3.2.1 captures a wealth of scenarios encountered in various areas to represent information on the ideal underlying function $\overline{x}\in\mathcal{H}$ obtained through some observation process. We discuss firmly nonexpansive observation processes in Section 3.2.2.1 and cocoercive ones in Section 3.2.2.2. In Section 3.2.2.3, we move to more general models in which the operators need not be Lipschitzian or even continuous.

#### 3.2.2.1 Prescriptions derived from firmly nonexpansive operators

We start with an instance of a proximal point prescription arising in a decomposition setting.

**Proposition 3.2.2** *Let $(\mathsf{H}_i)_{i\in\mathbb{I}}$ be an at most countable family of real Hilbert spaces, let $\mathcal{H}=\bigoplus_{i\in\mathbb{I}}\mathsf{H}_i$, let $\overline{x}\in\mathcal{H}$, and let $(\overline{\mathsf{x}}_i)_{i\in\mathbb{I}}$ be its decomposition, i.e., $(\forall i\in\mathbb{I})\ \overline{\mathsf{x}}_i\in\mathsf{H}_i$. For every $i\in\mathbb{I}$, let $\mathsf{F}_i\colon\mathsf{H}_i\to\mathsf{H}_i$ be a firmly nonexpansive operator. If $\mathbb{I}$ is infinite, suppose that there exists $z=(\mathsf{z}_i)_{i\in\mathbb{I}}\in\mathcal{H}$ such that $\sum_{i\in\mathbb{I}}\|\mathsf{F}_i\mathsf{z}_i-\mathsf{z}_i\|^2<+\infty$. Set $F\colon\mathcal{H}\to\mathcal{H}\colon x=(\mathsf{x}_i)_{i\in\mathbb{I}}\mapsto(\mathsf{F}_i\mathsf{x}_i)_{i\in\mathbb{I}}$ and $p=(\mathsf{F}_i\overline{\mathsf{x}}_i)_{i\in\mathbb{I}}$. Then $p$ is the proximal point of $\overline{x}$ relative to $F$.*

*Proof*. If $\mathbb{I}$ is infinite, we have

$$\begin{aligned}(\forall x\in\mathcal{H})\quad \frac{1}{3}\sum_{i\in\mathbb{I}}\|\mathsf{F}_i\mathsf{x}_i\|^2 &\leqslant \sum_{i\in\mathbb{I}}\|\mathsf{F}_i\mathsf{x}_i-\mathsf{F}_i\mathsf{z}_i\|^2+\sum_{i\in\mathbb{I}}\|\mathsf{F}_i\mathsf{z}_i-\mathsf{z}_i\|^2+\sum_{i\in\mathbb{I}}\|\mathsf{z}_i\|^2\\ &\leqslant \sum_{i\in\mathbb{I}}\|\mathsf{x}_i-\mathsf{z}_i\|^2+\sum_{i\in\mathbb{I}}\|\mathsf{F}_i\mathsf{z}_i-\mathsf{z}_i\|^2+\|z\|^2\\ &= \|x-z\|^2+\sum_{i\in\mathbb{I}}\|\mathsf{F}_i\mathsf{z}_i-\mathsf{z}_i\|^2+\|z\|^2\\ &< +\infty. \end{aligned}\tag{3.11}$$

This shows that, in all cases, $F$ is well defined and $p\in\mathcal{H}$. Furthermore,

$$\begin{aligned}(\forall x\in\mathcal{H})(\forall y\in\mathcal{H})\quad \|Fx-Fy\|^2 &= \sum_{i\in\mathbb{I}}\|\mathsf{F}_i\mathsf{x}_i-\mathsf{F}_i\mathsf{y}_i\|^2\\ &\leqslant \sum_{i\in\mathbb{I}}\|\mathsf{x}_i-\mathsf{y}_i\|^2-\sum_{i\in\mathbb{I}}\|(\mathsf{Id}-\mathsf{F}_i)\mathsf{x}_i-(\mathsf{Id}-\mathsf{F}_i)\mathsf{y}_i\|^2\\ &= \|x-y\|^2-\|(\mathrm{Id}-F)x-(\mathrm{Id}-F)y\|^2.\end{aligned}\tag{3.12}$$

Thus, $F$ is firmly nonexpansive. □

**Corollary 3.2.3** *Let $(\mathsf{H}_i)_{i\in\mathbb{I}}$ be an at most countable family of real Hilbert spaces, let $\mathcal{H}=\bigoplus_{i\in\mathbb{I}}\mathsf{H}_i$, let $\overline{x}\in\mathcal{H}$, and let $(\overline{\mathsf{x}}_i)_{i\in\mathbb{I}}$ be its decomposition. For every $i\in\mathbb{I}$, let $\mathsf{f}_i\in\Gamma_0(\mathsf{H}_i)$ and, if $\mathbb{I}$ is infinite,*

*suppose that* $\mathsf{f}_i \geqslant 0 = \mathsf{f}_i(0)$. *Then* $p = (\mathrm{prox}_{\mathsf{f}_i}\overline{\mathsf{x}}_i)_{i\in\mathbb{I}}$ *is a proximal point of* $\overline{x}$, *namely,* $p = \mathrm{prox}_f\overline{x}$, *where* $f\colon \mathcal{H} \to \left]-\infty,+\infty\right] : x = (\mathsf{x}_i)_{i\in\mathbb{I}} \mapsto \sum_{i\in\mathbb{I}} \mathsf{f}_i(\mathsf{x}_i)$.

*Proof*. We first note that $f$ is proper since the functions $(\mathsf{f}_i)_{i\in\mathbb{I}}$ are. Furthermore, we observe that, for every $i \in \mathbb{I}$, the function $f_i\colon \mathcal{H} \to \left]-\infty,+\infty\right] : x \mapsto \mathsf{f}_i(\mathsf{x}_i)$ lies in $\Gamma_0(\mathcal{H})$. We therefore derive from [5, Corollary 9.4] that $f = \sum_{i\in\mathbb{I}} f_i$ is lower semicontinuous and convex. This shows that $f \in \Gamma_0(\mathcal{H})$ and consequently that $\mathrm{prox}_f$ is well defined. For every $i \in \mathbb{I}$, let us introduce the firmly nonexpansive operator $\mathsf{F}_i = \mathrm{prox}_{\mathsf{f}_i}$. If $\mathbb{I}$ is infinite, since $\mathsf{0}$ is a minimizer of each of the functions $(\mathsf{f}_i)_{i\in\mathbb{I}}$, we derive from [5, Proposition 12.29] that $(\forall i \in \mathbb{I})\ \mathrm{prox}_{\mathsf{f}_i}\mathsf{0} = \mathsf{0}$. In turn, the condition $\sum_{i\in\mathbb{I}} \|\mathsf{F}_i\mathsf{z}_i - \mathsf{z}_i\|^2 < +\infty$ holds with $(\forall i \in \mathbb{I})\ \mathsf{z}_i = \mathsf{0}$. In view of Proposition 3.2.2, $p$ is the proximal point of $\overline{x}$ relative to $F\colon \mathcal{H} \to \mathcal{H}\colon x \mapsto (\mathrm{prox}_{\mathsf{f}_i}\mathsf{x}_i)_{i\in\mathbb{I}}$. Finally, since

$$\begin{aligned}
f(\mathrm{prox}_f\overline{x}) + \frac{1}{2}\|\overline{x} - \mathrm{prox}_f\overline{x}\|^2 &= \min_{y\in\mathcal{H}} \left(f(y) + \frac{1}{2}\|\overline{x} - y\|^2\right)\\
&= \min_{y\in\mathcal{H}} \sum_{i\in\mathbb{I}} \left(\mathsf{f}_i(\mathsf{y}_i) + \frac{1}{2}\|\overline{\mathsf{x}}_i - \mathsf{y}_i\|^2\right)\\
&= \sum_{i\in\mathbb{I}} \min_{\mathsf{y}_i\in\mathsf{H}_i} \left(\mathsf{f}_i(\mathsf{y}_i) + \frac{1}{2}\|\overline{\mathsf{x}}_i - \mathsf{y}_i\|^2\right)\\
&= \sum_{i\in\mathbb{I}} \left(\mathsf{f}_i(\mathrm{prox}_{\mathsf{f}_i}\overline{\mathsf{x}}_i) + \frac{1}{2}\|\overline{\mathsf{x}}_i - \mathrm{prox}_{\mathsf{f}_i}\overline{\mathsf{x}}_i\|^2\right)\\
&= f(p) + \frac{1}{2}\|\overline{x} - p\|^2,
\end{aligned} \tag{3.13}$$

we conclude that $p = \mathrm{prox}_f\overline{x}$. □

**Corollary 3.2.4** *Suppose that* $\mathcal{H}$ *is separable, let* $(e_i)_{i\in\mathbb{I}}$ *be an orthonormal basis of* $\mathcal{H}$, *and let* $\overline{x} \in \mathcal{H}$. *For every* $i \in \mathbb{I}$, *let* $\beta_i \in \left]0,+\infty\right[$ *and let* $\varrho_i\colon \mathbb{R} \to \mathbb{R}$ *be increasing and* $1/\beta_i$*-Lipschitzian. If* $\mathbb{I}$ *is infinite, suppose that* $(\forall i \in \mathbb{I})\ \varrho_i(0) = 0$. *Then* $p = \sum_{i\in\mathbb{I}} \beta_i\varrho_i(\langle \overline{x} \mid e_i\rangle)e_i$ *is a proximal point of* $\overline{x}$.

*Proof*. For every $i \in \mathbb{I}$, $\beta_i\varrho_i$ is increasing and nonexpansive, hence firmly nonexpansive. We then deduce from Proposition 3.2.2 that $\Phi\colon \ell^2(\mathbb{I}) \to \ell^2(\mathbb{I})\colon (\xi_i)_{i\in\mathbb{I}} \mapsto (\beta_i\varrho_i(\xi_i))_{i\in\mathbb{I}}$ is firmly nonexpansive. Now set $L\colon \mathcal{H} \to \ell^2(\mathbb{I})\colon x \mapsto (\langle x \mid e_i\rangle)_{i\in\mathbb{I}}$ and $F = L^* \circ \Phi \circ L$. Since $\|L\| = 1$, it follows from [5, Corollary 4.13] that $F$ is firmly nonexpansive. This shows that $p = L^*(\Phi(L\overline{x}))$ is the proximal point of $\overline{x}$ relative to $F$. □

**Example 3.2.5** In the context of Corollary 3.2.4, for every $i \in \mathbb{I}$, let $\omega_i \in [0,1]$, let $\eta_i \in \left]0,+\infty\right[$, let $\delta_i \in \left]0,+\infty\right[$, and set $\varrho_i\colon \xi \mapsto (2\omega_i/\pi)\mathrm{arctan}(\eta_i\,\xi) + (1-\omega_i)\mathrm{sign}\,(\xi)(1-\exp(-\delta_i|\xi|))$. Then, for every $i \in \mathbb{I}$, $\varrho_i$ is increasing and $(2\omega_i\eta_i/\pi + (1-\omega_i)\delta_i)$-Lipschitzian with $\varrho_i(0) = 0$. The resulting proximal point

$$p = \sum_{i\in\mathbb{I}} \frac{\varrho_i(\langle \overline{x} \mid e_i\rangle)}{(2\omega_i\eta_i/\pi + (1-\omega_i)\delta_i}e_i \tag{3.14}$$

models a parallel distortion of the original signal $\overline{x}$ [56, Sections 10.6 & 13.5].

**Example 3.2.6 (shrinkage)** In signal processing and statistics, a powerful idea is to decompose a function $\overline{x} \in \mathcal{H}$ in an orthonormal basis $(e_i)_{i\in\mathbb{I}}$ and to transform the coefficients of the decomposition to construct nonlinear approximations with certain attributes such as sparsity [11, 20, 23, 25, 26, 28, 55]. As noted in [23], a broad model in this context is

$$p = \sum_{i\in\mathbb{I}} \big(\operatorname{prox}_{\phi_i}\langle \overline{x} \mid e_i\rangle\big)e_i \tag{3.15}$$

where, for every $i \in \mathbb{I}$, the function $\phi_i \in \Gamma_0(\mathbb{R})$ satisfies $\phi_i \geqslant 0 = \phi_i(0)$ and models prior information on the coefficient $\langle \overline{x} \mid e_i\rangle$. The problem is then to reconstruct $\overline{x}$ given its shrunk version $p$. For instance, in the classical work of [28], $(e_i)_{i\in\mathbb{I}}$ is a wavelet basis and $(\forall i \in \mathbb{I})$ $\phi_i = \omega|\cdot|$, with $\omega \in \left]0, +\infty\right[$. This yields $p = \sum_{i\in\mathbb{I}}(\operatorname{sign}(\langle \overline{x} \mid e_i\rangle)\max\{|\langle \overline{x} \mid e_i\rangle| - \omega, 0\})e_i$. In general, to see that $p$ in (3.15) is a proximal point of $\overline{x}$, it suffices to apply Corollary 3.2.4 with, for every $i \in \mathbb{I}$, $\beta_i = 1$ and $\varrho_i = \operatorname{prox}_{\phi_i}$, whence $\varrho_i(0) = 0$ by [5, Proposition 12.29]. More precisely, [5, Proposition 24.16] entails that $p$ is the proximal point of $\overline{x}$ relative to the function $f\colon \mathcal{H} \to \left]-\infty, +\infty\right] : x \mapsto \sum_{i\in\mathbb{I}} \phi_i(\langle x \mid e_i\rangle)$.

**Example 3.2.7 (partitioning)** Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and let $(\Omega_i)_{i\in\mathbb{I}}$ be an at most countable $\mathcal{F}$-partition of $\Omega$. Let us consider the instantiation of Proposition 3.2.2 in which $\mathcal{H} = L^2(\Omega, \mathcal{F}, \mu)$ and, for every $i \in \mathbb{I}$, $\mathsf{H}_i = L^2(\Omega_i, \mathcal{F}_i, \mu)$, where $\mathcal{F}_i = \big\{\Omega_i \cap S \;\big|\; S \in \mathcal{F}\big\}$. Let $\overline{x} \in \mathcal{H}$ and $(\forall i \in \mathbb{I})$ $\overline{\mathsf{x}}_i = \overline{x}|_{\Omega_i}$. Moreover, for every $i \in \mathbb{I}$, $\phi_i$ is an even function in $\Gamma_0(\mathbb{R})$ such that $\phi_i(0) = 0$ and $\phi_i \neq \iota_{\{0\}}$, and we set $\rho_i = \max \partial\phi_i(0)$. Then we derive from Corollary 3.2.3 and [8, Proposition 2.1] that the proximal point of $\overline{x}$ relative to $f : x \mapsto \sum_{i\in\mathbb{I}} \phi_i(\|\mathsf{x}_i\|)$ is

$$p = \Big(\big(\operatorname{prox}_{\phi_i}\|\overline{\mathsf{x}}_i\|\big)\mathsf{u}_{\rho_i}(\overline{\mathsf{x}}_i)\Big)_{i\in\mathbb{I}}, \quad \text{where} \quad \mathsf{u}_{\rho_i}\colon \mathsf{H}_i \to \mathsf{H}_i\colon \mathsf{x}_i \mapsto \begin{cases} \mathsf{x}_i/\|\mathsf{x}_i\|, & \text{if } \|\mathsf{x}_i\| > \rho_i; \\ 0, & \text{if } \|\mathsf{x}_i\| \leqslant \rho_i. \end{cases} \tag{3.16}$$

For each $i \in \mathbb{I}$, this process eliminates the $i$th block $\overline{\mathsf{x}}_i$ if its norm is less than $\rho_i \in \left]0, +\infty\right[$.

**Example 3.2.8 (group shrinkage)** In Example 3.2.7, suppose that $\Omega = \{1, \ldots, N\}$, $\mathcal{F} = 2^\Omega$, and $\mu$ is the counting measure. Then $\mathcal{H}$ is the standard Euclidean space $\mathbb{R}^N$, which is decomposed in $m$ factors as $\mathbb{R}^N = \mathbb{R}^{N_1} \times \cdots \times \mathbb{R}^{N_m}$, where $\sum_{i=1}^m N_i = N$. Now suppose that $(\forall i \in \mathbb{I} = \{1, \ldots, m\})$ $\phi_i = \rho_i|\cdot|$, where $\rho_i \in \left]0, +\infty\right[$. Then it follows from [5, Example 14.5] that the proximal point $p$ of (3.16) is obtained by group-soft thresholding the vector $\overline{x} = (\overline{\mathsf{x}}_1, \ldots, \overline{\mathsf{x}}_m) \in \mathbb{R}^N$, that is [60],

$$p = \left(\left(1 - \frac{\rho_1}{\max\{\|\overline{\mathsf{x}}_1\|, \rho_1\}}\right)\overline{\mathsf{x}}_1, \ldots, \left(1 - \frac{\rho_m}{\max\{\|\overline{\mathsf{x}}_m\|, \rho_m\}}\right)\overline{\mathsf{x}}_m\right). \tag{3.17}$$

#### 3.2.2.2 Prescriptions derived from cocoercive operators

Let us first recall that, given a real Hilbert space $\mathcal{G}$ and $\beta \in ]0, +\infty[$, an operator $Q\colon \mathcal{G} \to \mathcal{G}$ is *$\beta$-cocoercive* if

$$(\forall x \in \mathcal{G})(\forall y \in \mathcal{G}) \quad \langle x - y \mid Qx - Qy \rangle \geqslant \beta \|Qx - Qy\|^2, \tag{3.18}$$

which means that $\beta Q$ is firmly nonexpansive [5, Section 4.2]. In the following proposition, a proximal point is constructed from a finite family of nonlinear observations $(q_i)_{i\in\mathbb{I}}$ of linear transformations of the function $\overline{x} \in \mathcal{H}$, where the nonlinearities are modeled via cocoercive operators. Item (ii) below shows that this proximal point contains the same information as the observations $(q_i)_{i\in\mathbb{I}}$.

**Proposition 3.2.9** *Let $(\mathcal{G}_i)_{i\in\mathbb{I}}$ be a finite family of real Hilbert spaces and let $\overline{x} \in \mathcal{H}$. For every $i \in \mathbb{I}$, let $\beta_i \in ]0, +\infty[$, let $Q_i\colon \mathcal{G}_i \to \mathcal{G}_i$ be $\beta_i$-cocoercive, let $L_i\colon \mathcal{H} \to \mathcal{G}_i$ be a nonzero bounded linear operator, and define $q_i = Q_i(L_i\overline{x})$. Set*

$$\beta = \frac{1}{\displaystyle\sum_{i\in\mathbb{I}} \frac{\|L_i\|^2}{\beta_i}}, \quad p = \beta \sum_{i\in\mathbb{I}} L_i^* q_i, \quad \text{and} \quad F = \beta \sum_{i\in\mathbb{I}} L_i^* \circ Q_i \circ L_i. \tag{3.19}$$

*Then the following hold:*

(i) *$p$ is the proximal point of $\overline{x}$ relative to $F$.*

(ii) $(\forall x \in \mathcal{H})\ Fx = p \Leftrightarrow (\forall i \in \mathbb{I})\ Q_i(L_i x) = q_i$.

*Proof.* (i): It is clear that $p = F\overline{x}$. In addition, the firm nonexpansiveness of $F$ follows from [5, Proposition 4.12].

(ii): Take $x \in \mathcal{H}$ such that $Fx = p$. Then $Fx = F\overline{x}$ and (4.11) yields

$$\begin{aligned} 0 &= \frac{\langle Fx - F\overline{x} \mid x - \overline{x} \rangle}{\beta} \\ &= \sum_{i\in\mathbb{I}} \langle Q_i(L_i x) - Q_i(L_i\overline{x}) \mid L_i x - L_i\overline{x} \rangle \\ &\geqslant \sum_{i\in\mathbb{I}} \beta_i \|Q_i(L_i x) - Q_i(L_i\overline{x})\|^2 \\ &= \sum_{i\in\mathbb{I}} \beta_i \|Q_i(L_i x) - q_i\|^2, \end{aligned} \tag{3.20}$$

and therefore $(\forall i \in \mathbb{I})\ Q_i(L_i x) = q_i$. The reverse implication is clear. □

Next, we consider the case when the observations $(q_i)_{i\in\mathbb{I}}$ in Proposition 3.2.9 are obtained through proximity operators.

**Proposition 3.2.10** *Let $(\mathcal{G}_i)_{i\in\mathbb{I}}$ be a finite family of real Hilbert spaces and let $\overline{x} \in \mathcal{H}$. For every $i \in \mathbb{I}$, let $g_i \in \Gamma_0(\mathcal{G}_i)$, let $L_i\colon \mathcal{H} \to \mathcal{G}_i$ be a nonzero bounded linear operator, and define $q_i = \operatorname{prox}_{g_i}(L_i\overline{x})$.*

*Suppose that* $\beta = 1/(\sum_{i\in\mathbb{I}}\|L_i\|^2)$*, and set* $p = \beta\sum_{i\in\mathbb{I}} L_i^* q_i$ *and* $F = \beta\sum_{i\in\mathbb{I}} L_i^*\circ\operatorname{prox}_{g_i}\circ L_i$*. Then the following hold:*

(i) $p$ *is the proximal point of* $\overline{x}$ *relative to* $F$.

(ii) $(\forall x\in\mathcal{H})\ Fx = p \Leftrightarrow (\forall i\in\mathbb{I})\ \operatorname{prox}_{g_i}(L_i x) = q_i$.

(iii) *If* $\beta \geqslant 1$*, then*

$$F = \beta\operatorname{prox}_f, \quad \textit{where} \quad f = \left(\sum_{i\in\mathbb{I}}\left(g_i^*\,\square\,\frac{\|\cdot\|_{\mathcal{G}_i}^2}{2}\right)\circ L_i\right)^* - \frac{\|\cdot\|_{\mathcal{H}}^2}{2}. \tag{3.21}$$

*Proof.* (i)–(ii): Apply Proposition 3.2.9 with $(\forall i\in\mathbb{I})\ Q_i = \operatorname{prox}_{g_i}$ and $\beta_i = 1$.
(iii): This follows from [18, Proposition 3.9]. □

**Example 3.2.11 (scalar observations)** We specialize the setting of Proposition 3.2.10 by assuming that, for some $i\in\mathbb{I}$, $\mathcal{G}_i = \mathbb{R}$ and $L_i = \langle\cdot\mid a_i\rangle$, where $0\neq a_i\in\mathcal{H}$. Let us denote by $\chi_i = \operatorname{prox}_{g_i}\langle\overline{x}\mid a_i\rangle$ the resulting observation. This scenario allows us to recover various nonlinear observation processes used in the literature.

(i) Set $g_i = \iota_D$, where $D$ is a nonempty closed interval in $\mathbb{R}$ with $\underline{\delta} = \inf D\in[-\infty,+\infty[$ and $\overline{\delta} = \sup D\in\,]-\infty,+\infty]$. Then we obtain the hard clipping process

$$\chi_i = \operatorname{proj}_D\langle\overline{x}\mid a_i\rangle = \begin{cases}\overline{\delta}, & \text{if } \langle\overline{x}\mid a_i\rangle > \overline{\delta};\\ \langle\overline{x}\mid a_i\rangle, & \text{if } \langle\overline{x}\mid a_i\rangle\in D;\\ \underline{\delta}, & \text{if } \langle\overline{x}\mid a_i\rangle < \underline{\delta},\end{cases} \tag{3.22}$$

which shows up in several nonlinear data collection processes; see for instance [2, 31, 54, 58]. It models the inability of the sensors to record values above $\overline{\delta}$ and below $\underline{\delta}$.

(ii) Let $\Omega$ be a nonempty closed interval of $\mathbb{R}$ and let $\operatorname{soft}_\Omega$ be the associated soft thresholder, i.e.,

$$\operatorname{soft}_\Omega\colon\mathbb{R}\to\mathbb{R}\colon\xi\mapsto\begin{cases}\xi-\overline{\omega}, & \text{if } \xi>\overline{\omega};\\ 0, & \text{if } \xi\in\Omega;\\ \xi-\underline{\omega}, & \text{if } \xi<\underline{\omega},\end{cases}\qquad\text{with}\quad\begin{cases}\overline{\omega} = \sup\Omega\\ \underline{\omega} = \inf\Omega.\end{cases} \tag{3.23}$$

Further, let $\psi\in\Gamma_0(\mathbb{R})$ be differentiable at $0$ with $\psi'(0) = 0$, and set $g_i = \psi+\sigma_\Omega$, where $\sigma_\Omega$ is the support function of $\Omega$. Then it follows from [20, Proposition 3.6] that

$$\chi_i = \operatorname{prox}_\psi\big(\operatorname{soft}_\Omega\langle\overline{x}\mid a_i\rangle\big) = \begin{cases}\operatorname{prox}_\psi(\langle\overline{x}\mid a_i\rangle-\overline{\omega}), & \text{if } \langle\overline{x}\mid a_i\rangle>\overline{\omega};\\ 0, & \text{if } \langle\overline{x}\mid a_i\rangle\in\Omega;\\ \operatorname{prox}_\psi(\langle\overline{x}\mid a_i\rangle-\underline{\omega}), & \text{if } \langle\overline{x}\mid a_i\rangle<\underline{\omega}.\end{cases} \tag{3.24}$$

In particular, if $\Omega = [-\omega, \omega]$ and $\psi = 0$, we obtain the standard soft thresholding operation

$$\chi_i = \operatorname{sign}\left(\langle \overline{x} \mid a_i \rangle\right) \max\{|\langle \overline{x} \mid a_i \rangle| - \omega, 0\} \tag{3.25}$$

of [28]. On the other hand, if $\Omega = ]-\infty, \omega]$ and $\psi = 0$, we obtain a nonlinear sensor model from [37].

(iii) In (ii) suppose that $\psi = \iota_D$, where $D$ is as in (i) and contains $0$ in its interior. Then (3.24) becomes

$$\chi_i = \begin{cases} \overline{\delta}, & \text{if } \langle \overline{x} \mid a_i \rangle \geqslant \overline{\delta} + \overline{\omega}; \\ \langle \overline{x} \mid a_i \rangle - \overline{\omega}, & \text{if } \overline{\omega} < \langle \overline{x} \mid a_i \rangle < \overline{\delta} + \overline{\omega}; \\ 0, & \text{if } \langle \overline{x} \mid a_i \rangle \in \Omega; \\ \langle \overline{x} \mid a_i \rangle - \underline{\omega}, & \text{if } \underline{\delta} + \underline{\omega} < \langle \overline{x} \mid a_i \rangle < \underline{\omega}; \\ \underline{\delta}, & \text{if } \langle \overline{x} \mid a_i \rangle \leqslant \underline{\delta} + \underline{\omega}. \end{cases} \tag{3.26}$$

This operation combines hard clipping and soft thresholding.

(iv) Set

$$g_i \colon \xi \mapsto \begin{cases} \dfrac{(1+\xi)\ln(1+\xi) + (1-\xi)\ln(1-\xi) - \xi^2}{2}, & \text{if } |\xi| < 1; \\ \ln(2) - 1/2, & \text{if } |\xi| = 1; \\ +\infty, & \text{if } |\xi| > 1. \end{cases} \tag{3.27}$$

Then it follows from [21, Example 2.12] that $\chi_i = \tanh(\langle \overline{x} \mid a_i \rangle)$. This soft clipping model is used in [2, 29].

(v) Set

$$g_i \colon \xi \mapsto \begin{cases} -\dfrac{2}{\pi} \ln\left(\cos\left(\dfrac{\pi\xi}{2}\right)\right) - \dfrac{\xi^2}{2}, & \text{if } |\xi| < 1; \\ +\infty, & \text{if } |\xi| \geqslant 1. \end{cases} \tag{3.28}$$

Then it follows from [21, Example 2.11] that $\chi_i = (2/\pi)\arctan(\langle \overline{x} \mid a_i \rangle)$. This soft clipping model appears in [2].

(vi) Set

$$g_i \colon \xi \mapsto \begin{cases} -|\xi| - \ln(1-|\xi|) - \xi^2/2, & \text{if } |\xi| < 1; \\ +\infty, & \text{if } |\xi| \geqslant 1. \end{cases} \tag{3.29}$$

Then it follows from [21, Example 2.15] that $\chi_i = \langle \overline{x} \mid a_i \rangle / (1 + |\langle \overline{x} \mid a_i \rangle|)$. This soft clipping model is found in [29, 39].

(vii) Set

$$g_i\colon \xi \mapsto \begin{cases} |\xi| + (1-|\xi|)\ln\big|1-|\xi|\big| - \xi^2/2, & \text{if } |\xi| < 1; \\ 1/2, & \text{if } |\xi| = 1; \\ +\infty, & \text{if } |\xi| > 1. \end{cases} \tag{3.30}$$

For every $\xi \in \left]-1,1\right[ = \operatorname{dom} g_i' = \operatorname{ran}\operatorname{prox}_{g_i}$, we have $\xi + g_i'(\xi) = -\operatorname{sign}(\xi)\ln(1-|\xi|)$. Hence,

$$\big(\operatorname{Id} + g_i'\big)^{-1} = \operatorname{prox}_{g_i}\colon \xi \mapsto \operatorname{sign}(\xi)\big(1-\exp(-|\xi|)\big) \tag{3.31}$$

and, therefore, $\chi_i = \operatorname{sign}(\langle \overline{x} \mid a_i\rangle)(1-\exp(-|\langle \overline{x} \mid a_i\rangle|))$. This distortion model is found in [56, Section 10.6.3].

(viii) Let $\eta_i \in \left]0,+\infty\right[$ and set

$$g_i\colon \xi \mapsto \eta_i\xi + \begin{cases} \xi\ln(\xi) + (1-\xi)\ln(1-\xi) - \xi^2/2, & \text{if } \xi \in \left]0,1\right[; \\ 0, & \text{if } \xi = 0; \\ -1/2, & \text{if } \xi = 1; \\ +\infty, & \text{if } \xi \in \mathbb{R}\smallsetminus[0,1]. \end{cases} \tag{3.32}$$

Proceeding as in (vii), we obtain

$$\chi_i = \frac{1}{1+\exp(\eta_i - \langle \overline{x} \mid a_i\rangle)}, \tag{3.33}$$

which is an encoding scheme used in [36].

**Example 3.2.12** In Proposition 3.2.10 suppose that, for some $i \in \mathbb{I}$, $g_i = \phi_i \circ d_{D_i}$, where $\phi_i \in \Gamma_0(\mathbb{R})$ is even with $\phi_i(0) = 0$, and $D_i \subset \mathcal{G}_i$ is nonempty, closed, and convex. Then it follows from [8, Proposition 2.1] that $q_i$ is the nonlinear observation defined as follows:

(i) Suppose that $\phi_i = \iota_{\{0\}}$. Then

$$q_i = \operatorname{proj}_{D_i}(L_i\overline{x}) \tag{3.34}$$

captures several applications. Thus, if $\mathcal{H} = \mathbb{R}^N$ and $D_i = \big\{(\xi_i)_{1\leqslant i\leqslant N} \in \mathbb{R}^N \;\big|\; \xi_1 \leqslant \cdots \leqslant \xi_N\big\}$, then $q_i$ is the best isotonic approximation to $L_i\overline{x}$ [24]. On the other hand, if $D_i$ is the closed ball with center $0$ and radius $\rho_i \in \left]0,+\infty\right[$, then (3.34) reduces to the hard saturation process

$$q_i = \begin{cases} \dfrac{\rho_i}{\|L_i\overline{x}\|}L_i\overline{x}, & \text{if } \|L_i\overline{x}\| > \rho_i; \\ L_i\overline{x}, & \text{if } \|L_i\overline{x}\| \leqslant \rho_i, \end{cases} \tag{3.35}$$

which can be viewed as an infinite dimensional version of Example 3.2.11(i).

(ii) Suppose that $\phi_i \neq \iota_{\{0\}}$ and set $\rho_i = \max \partial\phi_i(0)$. Then

$$q_i = \begin{cases} L_i\overline{x} + \dfrac{\operatorname{prox}_{\phi_i^*} d_{D_i}(L_i\overline{x})}{d_{D_i}(L_i\overline{x})}\big(\operatorname{proj}_{D_i}(L_i\overline{x}) - L_i\overline{x}\big), & \text{if } d_{D_i}(L_i\overline{x}) > \rho_i; \\ \operatorname{proj}_{D_i}(L_i\overline{x}), & \text{if } d_{D_i}(L_i\overline{x}) \leqslant \rho_i. \end{cases} \tag{3.36}$$

In particular, assume that $D_i = \{0\}$. Then (3.36) reduces to the abstract soft thresholding process

$$q_i = \begin{cases} L_i\overline{x} - \dfrac{\operatorname{prox}_{\phi_i^*}\|L_i\overline{x}\|}{\|L_i\overline{x}\|}\, L_i\overline{x}, & \text{if } \|L_i\overline{x}\| > \rho_i; \\ 0, & \text{if } \|L_i\overline{x}\| \leqslant \rho_i, \end{cases} \tag{3.37}$$

which cannot record inputs with norm below a certain value. Let us further specialize to the setting in which $\phi_i = \rho_i|\cdot|$ with $\rho_i \in \left]0, +\infty\right[$. Then $\phi_i^* = \iota_{[-\rho_i,\rho_i]}$, $\partial\phi_i(0) = [-\rho_i, \rho_i]$, and (3.37) becomes

$$q_i = \begin{cases} \left(1 - \dfrac{\rho_i}{\|L_i\overline{x}\|}\right) L_i\overline{x}, & \text{if } \|L_i\overline{x}\| > \rho_i; \\ 0, & \text{if } \|L_i\overline{x}\| \leqslant \rho_i, \end{cases} \tag{3.38}$$

which can be viewed as an infinite dimensional version of (3.25).

#### 3.2.2.3 Prescriptions derived from non-cocoercive operators

Here, we exemplify observation processes which are not cocoercive, and possibly not even continuous, but that can still be represented by proximal points relative to some firmly nonexpansive operator, as required in Problem 3.2.1. The results in this section constructively provide the proximal points and phrase the evaluation of each firmly nonexpansive operator in terms of the nonlinearity in the observation process.

**Example 3.2.13** In the spirit of the shrinkage ideas of Corollary 3.2.4 and Example 3.2.6, a prescription involving more general transformations $(\varrho_i)_{i\in\mathbb{I}}$ can be used to derive an equivalent prescribed proximal point. Let us adopt the setting of Corollary 3.2.4, except that $(\varrho_i)_{i\in\mathbb{I}}$ are now arbitrary operators from $\mathbb{R}$ to $\mathbb{R}$ such that, for some $\delta \in \left]0, +\infty\right[$, $\sup_{i\in\mathbb{I}} |\varrho_i| \leqslant \delta|\cdot|$. Since

$$\sum_{i\in\mathbb{I}} \big|\varrho_i(\langle \overline{x} \mid e_i\rangle)\big|^2 \leqslant \delta^2 \sum_{i\in\mathbb{I}} |\langle \overline{x} \mid e_i\rangle|^2 = \delta^2\|\overline{x}\|^2 < +\infty, \tag{3.39}$$

the prescription $q = \sum_{i\in\mathbb{I}} \varrho_i(\langle \overline{x} \mid e_i\rangle)e_i$ is well defined. While $q$ is not a proximal point in general, an equivalent proximal point $p$ can be constructed from it in certain instances. To illustrate this process, let us first compute $(\forall i \in \mathbb{I})\ \chi_i = \langle q \mid e_i\rangle = \varrho_i(\langle \overline{x} \mid e_i\rangle)$. In both examples to follow, for every $i \in \mathbb{I}$, we construct an operator $\sigma_i\colon \mathbb{R} \to \mathbb{R}$ such that $\varphi_i = \sigma_i \circ \varrho_i$ is firmly nonexpansive, $\varphi_i(0) = 0$, and no information is lost when $\sigma_i$ is applied to the prescription $\chi_i = \varrho_i(\langle \overline{x} \mid e_i\rangle)$ in

the sense that

$$(\forall \xi \in \mathbb{R}) \quad \big[\, \chi_i = \varrho_i(\xi) \quad \Leftrightarrow \quad \sigma_i(\chi_i) = \sigma_i\big(\varrho_i(\xi)\big) = \varphi_i(\xi) \,\big]. \tag{3.40}$$

Using Corollary 3.2.4 with the firmly nonexpansive operators $(\varphi_i)_{i\in\mathbb{I}}$, this implies that $p = \sum_{i\in\mathbb{I}} \sigma_i(\chi_i)e_i$ is a proximal point of $\overline{x}$.

(i) Let $i \in \mathbb{I}$, let $\omega_i \in \left]0, +\infty\right[$, and consider the non-Lipschitzian sampling operator [1, 55]

$$\varrho_i \colon \xi \mapsto \begin{cases} \operatorname{sign}(\xi)\sqrt{\xi^2 - \omega_i^2}, & \text{if } |\xi| > \omega_i; \\ 0, & \text{if } |\xi| \leqslant \omega_i. \end{cases} \tag{3.41}$$

It is straightforward to verify that (3.40) holds with

$$\sigma_i \colon \xi \mapsto \operatorname{sign}(\xi)\left(\sqrt{\xi^2 + \omega_i^2} - \omega_i\right), \tag{3.42}$$

in which case $\varphi_i = \sigma_i \circ \varrho_i$ is the soft thresholder on $[-\omega_i, \omega_i]$ of (3.23).

(ii) Let $i \in \mathbb{I}$, let $\omega_i \in \left]0, +\infty\right[$, and consider the discontinuous sampling operator [55]

$$\varrho_i = \operatorname{hard}_{[-\omega_i,\omega_i]} : \xi \mapsto \begin{cases} \xi, & \text{if } |\xi| > \omega_i; \\ 0, & \text{if } |\xi| \leqslant \omega_i, \end{cases} \tag{3.43}$$

which is also known as the hard thresholder on $[-\omega_i, \omega_i]$. This operator is used as a sensing model in [7] and as a compression model in [57]. Then (3.40) is satisfied with

$$\sigma_i \colon \xi \mapsto \xi - \omega_i \operatorname{sign}(\xi), \tag{3.44}$$

in which case $\varphi_i = \sigma_i \circ \varrho_i$ turns out to be the soft thresholder on $[-\omega_i, \omega_i]$ of (3.23).

Next, we revisit Proposition 3.2.2 by relaxing the firm nonexpansiveness of the observation operators and constructing an equivalent proximal point via some transformation. This equivalence is expressed in (iii) below.

**Proposition 3.2.14** *Let $(\mathsf{H}_i)_{i\in\mathbb{I}}$ be an at most countable family of real Hilbert spaces, let $\mathcal{H} = \bigoplus_{i\in\mathbb{I}} \mathsf{H}_i$, let $\overline{x} \in \mathcal{H}$, and let $(\overline{\mathsf{x}}_i)_{i\in\mathbb{I}}$ be its decomposition, i.e., $(\forall i \in \mathbb{I})\ \overline{\mathsf{x}}_i \in \mathsf{H}_i$. In addition, for every $i \in \mathbb{I}$, let $\mathsf{Q}_i \colon \mathsf{H}_i \to \mathsf{H}_i$ and let $\mathsf{q}_i = \mathsf{Q}_i\overline{\mathsf{x}}_i$. Suppose that there exist operators $(\mathsf{S}_i)_{i\in\mathbb{I}}$ from $\mathsf{H}_i$ to $\mathsf{H}_i$ such that the operators $(\mathsf{F}_i)_{i\in\mathbb{I}} = (\mathsf{S}_i \circ \mathsf{Q}_i)_{i\in\mathbb{I}}$ satisfy the following:*

(i) *The operators $(\mathsf{F}_i)_{i\in\mathbb{I}}$ are firmly nonexpansive.*

(ii) *If $\mathbb{I}$ is infinite, there exists $(\mathsf{z}_i)_{i\in\mathbb{I}} \in \mathcal{H}$ such that $\sum_{i\in\mathbb{I}} \|\mathsf{F}_i\mathsf{z}_i - \mathsf{z}_i\|^2 < +\infty$.*

(iii) $(\forall i \in \mathbb{I})(\forall \mathsf{x}_i \in \mathsf{H}_i)\ \big[\, \mathsf{F}_i\mathsf{x}_i = \mathsf{S}_i\mathsf{q}_i \Leftrightarrow \mathsf{Q}_i\mathsf{x}_i = \mathsf{q}_i \,\big]$.

*Then* $p = (\mathsf{S}_i\mathsf{q}_i)_{i\in\mathbb{I}}$ *is the proximal point of* $\overline{x}$ *relative to* $F\colon \mathcal{H} \to \mathcal{H}\colon (\mathsf{x}_i)_{i\in\mathbb{I}} \mapsto (\mathsf{F}_i\mathsf{x}_i)_{i\in\mathbb{I}}$.

*Proof*. This follows from Proposition 3.2.2. □

The following result illustrates the process described in Proposition 3.2.14, through a generalization of the discontinuous hard thresholding operator of Example 3.2.13(ii), which corresponds to the case when $\mathsf{H}_i = \mathbb{R}$ and $\mathsf{C}_i = \{0\}$ in (3.45) below.

**Proposition 3.2.15** *Let* $(\mathsf{H}_i)_{i\in\mathbb{I}}$ *be an at most countable family of real Hilbert spaces, let* $\mathcal{H} = \bigoplus_{i\in\mathbb{I}} \mathsf{H}_i$, *let* $\overline{x} \in \mathcal{H}$, *and let* $(\overline{\mathsf{x}}_i)_{i\in\mathbb{I}}$ *be its decomposition. For every* $i \in \mathbb{I}$, *let* $\omega_i \in \left]0, +\infty\right[$, *let* $\mathsf{C}_i$ *be a nonempty closed convex subset of* $\mathsf{H}_i$, *set*

$$\mathsf{Q}_i\colon \mathsf{H}_i \to \mathsf{H}_i\colon \mathsf{x}_i \mapsto \begin{cases} \mathsf{x}_i, & \text{if } d_{\mathsf{C}_i}(\mathsf{x}_i) > \omega_i; \\ \operatorname{proj}_{\mathsf{C}_i}\mathsf{x}_i, & \text{if } d_{\mathsf{C}_i}(\mathsf{x}_i) \leqslant \omega_i, \end{cases} \tag{3.45}$$

*and let* $\mathsf{q}_i = \mathsf{Q}_i\overline{\mathsf{x}}_i$ *be the associated prescription. If* $\mathbb{I}$ *is infinite, suppose that* $(\forall i \in \mathbb{I})\ 0 \in \mathsf{C}_i$. *Further, for every* $i \in \mathbb{I}$, *set*

$$\mathsf{S}_i\colon \mathsf{H}_i \to \mathsf{H}_i\colon \mathsf{x}_i \mapsto \begin{cases} \mathsf{x}_i + \dfrac{\omega_i}{d_{\mathsf{C}_i}(\mathsf{x}_i)}(\operatorname{proj}_{\mathsf{C}_i}\mathsf{x}_i - \mathsf{x}_i), & \text{if } \mathsf{x}_i \notin \mathsf{C}_i; \\ \mathsf{x}_i, & \text{if } \mathsf{x}_i \in \mathsf{C}_i \end{cases} \quad \text{and} \quad \begin{cases} \mathsf{F}_i = \mathsf{S}_i \circ \mathsf{Q}_i \\ \mathsf{p}_i = \mathsf{S}_i\mathsf{q}_i. \end{cases} \tag{3.46}$$

*Finally, set* $p = (\mathsf{p}_i)_{i\in\mathbb{I}}$ *and* $f\colon \mathcal{H} \to \left]-\infty, +\infty\right] : (\mathsf{x}_i)_{i\in\mathbb{I}} \mapsto \sum_{i\in\mathbb{I}} \omega_i d_{\mathsf{C}_i}(\mathsf{x}_i)$. *Then the following hold:*

(i) *For every* $i \in \mathbb{I}$, $\mathsf{F}_i = \operatorname{prox}_{\omega_i d_{\mathsf{C}_i}}$.

(ii) $p$ *is the proximal point of* $\overline{x}$ *relative to* $f$.

(iii) *Let* $x = (\mathsf{x}_i)_{i\in\mathbb{I}} \in \mathcal{H}$. *Then* $\big[\,(\forall i \in \mathbb{I})\ \mathsf{Q}_i\mathsf{x}_i = \mathsf{q}_i\,\big] \Leftrightarrow \operatorname{prox}_f x = p$.

*Proof*. We derive from (3.45), (3.46), and [5, Proposition 3.21] that

$$(\forall i \in \mathbb{I})(\forall \mathsf{x}_i \in \mathsf{H}_i) \quad \mathsf{F}_i\mathsf{x}_i = \begin{cases} \operatorname{proj}_{\mathsf{C}_i}\mathsf{x}_i + \left(1 - \dfrac{\omega_i}{d_{\mathsf{C}_i}(\mathsf{x}_i)}\right)(\mathsf{x}_i - \operatorname{proj}_{\mathsf{C}_i}\mathsf{x}_i) \notin \mathsf{C}_i, & \text{if } d_{\mathsf{C}_i}(\mathsf{x}_i) > \omega_i; \\ \operatorname{proj}_{\mathsf{C}_i}\mathsf{x}_i \in \mathsf{C}_i, & \text{if } d_{\mathsf{C}_i}(\mathsf{x}_i) \leqslant \omega_i. \end{cases} \tag{3.47}$$

(i): This is a consequence of (3.47) and [5, Example 24.28].

(ii): If $\mathbb{I}$ is infinite, $0 \in \mathsf{C}_i \Rightarrow d_{\mathsf{C}_i}(0) = 0 \Rightarrow \mathsf{F}_i(0) = 0$ by (3.47). In turn, the claim follows from Corollary 3.2.3 and (i).

(iii): We first note that Corollary 3.2.3 and (i) imply that

$$(\mathsf{F}_i\mathsf{x}_i)_{i\in\mathbb{I}} = (\operatorname{prox}_{\omega_i d_{\mathsf{C}_i}}\mathsf{x}_i)_{i\in\mathbb{I}} = \operatorname{prox}_f x. \tag{3.48}$$

Now, suppose that $(\forall i \in \mathbb{I})\ \mathsf{Q}_i\mathsf{x}_i = \mathsf{q}_i$. Then $(\forall i \in \mathbb{I})\ \mathsf{F}_i\mathsf{x}_i = \mathsf{S}_i(\mathsf{Q}_i\mathsf{x}_i) = \mathsf{S}_i\mathsf{q}_i = \mathsf{p}_i$. In turn, (3.48) yields $\operatorname{prox}_f x = (\mathsf{F}_i\mathsf{x}_i)_{i\in\mathbb{I}} = p$. Conversely, suppose that $\operatorname{prox}_f x = p$ and fix $i \in \mathbb{I}$. We derive from

(3.48) and (3.46) that

$$\mathsf{F}_i\mathsf{x}_i = \mathsf{p}_i = \mathsf{S}_i\mathsf{q}_i = \mathsf{S}_i(\mathsf{Q}_i\overline{\mathsf{x}}_i) = \mathsf{F}_i\overline{\mathsf{x}}_i. \tag{3.49}$$

We must show that $\mathsf{Q}_i\mathsf{x}_i = \mathsf{q}_i$. It follows from (3.45), (3.47), and (3.49) that

$$d_{\mathsf{C}_i}(\mathsf{x}_i) \leqslant \omega_i \;\;\Leftrightarrow\;\; \mathsf{Q}_i\mathsf{x}_i = \operatorname{proj}_{\mathsf{C}_i}\mathsf{x}_i = \mathsf{F}_i\mathsf{x}_i = \mathsf{F}_i\overline{\mathsf{x}}_i \in \mathsf{C}_i \;\;\Rightarrow\;\; \begin{cases} d_{\mathsf{C}_i}(\overline{\mathsf{x}}_i) \leqslant \omega_i \\ \mathsf{Q}_i\mathsf{x}_i = \operatorname{proj}_{\mathsf{C}_i}\overline{\mathsf{x}}_i = \mathsf{Q}_i\overline{\mathsf{x}}_i = \mathsf{q}_i. \end{cases} \tag{3.50}$$

On the other hand, (3.45) yields

$$d_{\mathsf{C}_i}(\mathsf{x}_i) > \omega_i \quad\Rightarrow\quad \mathsf{Q}_i\mathsf{x}_i = \mathsf{x}_i, \tag{3.51}$$

while (3.49) and (3.47) yield

$$d_{\mathsf{C}_i}(\mathsf{x}_i) > \omega_i \;\;\Rightarrow\;\; \mathsf{p}_i = \mathsf{F}_i\overline{\mathsf{x}}_i = \mathsf{F}_i\mathsf{x}_i = \operatorname{proj}_{\mathsf{C}_i}\mathsf{x}_i + \left(1 - \frac{\omega_i}{d_{\mathsf{C}_i}(\mathsf{x}_i)}\right)(\mathsf{x}_i - \operatorname{proj}_{\mathsf{C}_i}\mathsf{x}_i) \notin \mathsf{C}_i \tag{3.52}$$

$$\Rightarrow\;\; \mathsf{F}_i\overline{\mathsf{x}}_i = \operatorname{proj}_{\mathsf{C}_i}\overline{\mathsf{x}}_i + \left(1 - \frac{\omega_i}{d_{\mathsf{C}_i}(\overline{\mathsf{x}}_i)}\right)(\overline{\mathsf{x}}_i - \operatorname{proj}_{\mathsf{C}_i}\overline{\mathsf{x}}_i) \;\text{ and }\; d_{\mathsf{C}_i}(\overline{\mathsf{x}}_i) > \omega_i \tag{3.53}$$

$$\Rightarrow\;\; \mathsf{q}_i = \mathsf{Q}_i\overline{\mathsf{x}}_i = \overline{\mathsf{x}}_i. \tag{3.54}$$

Therefore, in view of (3.51), it remains to show that $\overline{\mathsf{x}}_i = \mathsf{x}_i$. Set $\mathsf{r}_i = \operatorname{proj}_{\mathsf{C}_i}\mathsf{p}_i$. We deduce from (3.52), (3.53), and [5, Proposition 3.21] that $\mathsf{r}_i = \operatorname{proj}_{\mathsf{C}_i}\overline{\mathsf{x}}_i = \operatorname{proj}_{\mathsf{C}_i}\mathsf{x}_i$. Thus, (3.52) and (3.53) yield

$$\mathsf{p}_i - \mathsf{r}_i = \left(1 - \frac{\omega_i}{\|\mathsf{x}_i - \mathsf{r}_i\|}\right)(\mathsf{x}_i - \mathsf{r}_i) = \left(1 - \frac{\omega_i}{\|\overline{\mathsf{x}}_i - \mathsf{r}_i\|}\right)(\overline{\mathsf{x}}_i - \mathsf{r}_i). \tag{3.55}$$

Taking the norm of both sides yields $\|\mathsf{x}_i - \mathsf{r}_i\| = \|\overline{\mathsf{x}}_i - \mathsf{r}_i\|$ and hence $\overline{\mathsf{x}}_i = \mathsf{x}_i$. □

### 3.2.3 A block-iterative extrapolated algorithm for best approximation

We propose a flexible algorithm to solve the following abstract best approximation problem. This new algorithm, which is of interest in its own right, will be specialized in Section 3.2.4 to the setting of Problem 3.2.1.

**Problem 3.2.16** Let $\mathcal{H}$ be a real Hilbert space, let $(C_i)_{i\in I}$ be an at most countable family of closed convex subsets of $\mathcal{H}$ with nonempty intersection $C$, and let $x_0 \in \mathcal{H}$. The goal is to find $\operatorname{proj}_C x_0$, i.e., to

$$\text{minimize } \|x - x_0\| \quad \text{subject to} \quad x \in \bigcap_{i\in I} C_i. \tag{3.56}$$

In 1968, Yves Haugazeau proposed in his unpublished thesis [34] an iterative method to solve Problem 3.2.16 when $I$ is finite. His algorithm proceeds by periodic projections onto the individual sets.

**Proposition 3.2.17** [34, Théorème 3-2] *In Problem 3.2.16, suppose that $I$ is finite, say $I =$*

$\{0, \dots, m-1\}$, *where* $2 \leqslant m \in \mathbb{N}$. *Given* $(s,t) \in \mathcal{H}^2$ *such that*

$$D = \big\{ x \in \mathcal{H} \;\big|\; \langle x - s \mid x_0 - s \rangle \leqslant 0 \;\text{ and }\; \langle x - t \mid s - t \rangle \leqslant 0 \big\} \neq \varnothing, \tag{3.57}$$

*set* $\chi = \langle x_0 - s \mid s - t \rangle$, $\mu = \|x_0 - s\|^2$, $\nu = \|s - t\|^2$, *and* $\rho = \mu\nu - \chi^2$, *and define*

$$Q(x_0, s, t) = \operatorname{proj}_D x_0 = \begin{cases} t, & \text{if } \rho = 0 \text{ and } \chi \geqslant 0; \\ x_0 + \left(1 + \dfrac{\chi}{\nu}\right)(t - s), & \text{if } \rho > 0 \text{ and } \chi\nu \geqslant \rho; \\ s + \dfrac{\nu}{\rho}\big(\chi(x_0 - s) + \mu(t - s)\big), & \text{if } \rho > 0 \text{ and } \chi\nu < \rho. \end{cases} \tag{3.58}$$

*Construct a sequence* $(x_n)_{n\in\mathbb{N}}$ *by iterating*

$$\begin{array}{l} \textit{for } n = 0, 1, \dots \\ \quad \left\lfloor \begin{array}{l} t_n = \operatorname{proj}_{C_{n(\mathrm{mod}\, m)}} x_n \\ x_{n+1} = Q(x_0, x_n, t_n). \end{array} \right. \end{array} \tag{3.59}$$

*Then* $x_n \to \operatorname{proj}_C x_0$.

Haugazeau's algorithm uses only one set at each iteration. The following variant due to Guy Pierra uses all of them simultaneously.

**Proposition 3.2.18** [50, Théorème V.1] *In Problem 3.2.16, suppose that* $I$ *is finite, let* $Q$ *be as in Proposition 3.2.17, set* $\omega = 1/\operatorname{card} I$, *and fix* $\varepsilon \in \left]0, 1\right[$. *Construct a sequence* $(x_n)_{n\in\mathbb{N}}$ *by iterating*

$$\begin{array}{l} \textit{for } n = 0, 1, \dots \\ \quad \left\lfloor \begin{array}{l} \textit{for every } i \in I \\ \quad \left\lfloor \begin{array}{l} a_{i,n} = \operatorname{proj}_{C_i} x_n \\ \theta_{i,n} = \|a_{i,n} - x_n\|^2 \end{array} \right. \\ \theta_n = \omega \sum_{i\in I} \theta_{i,n} \\ \textit{if } \theta_n = 0 \\ \quad \left\lfloor \begin{array}{l} t_n = x_n \end{array} \right. \\ \textit{else} \\ \quad \left\lfloor \begin{array}{l} d_n = \omega \sum_{i\in I} a_{i,n} \\ y_n = d_n - x_n \\ \lambda_n = \theta_n / \|y_n\|^2 \\ t_n = x_n + \lambda_n y_n \end{array} \right. \\ x_{n+1} = Q(x_0, x_n, t_n). \end{array} \right. \end{array} \tag{3.60}$$

*Then* $x_n \to \operatorname{proj}_C x_0$.

**Remark 3.2.19** An attractive feature of Pierra's algorithm (3.60) is that, by convexity of $\|\cdot\|^2$,

the relaxation parameter $\lambda_n$ can extrapolate beyond 1, hence attaining large values that induce fast convergence [17, 50].

Propositions 3.2.17 and 3.2.18 were unified and extended in [15, Section 6.5] in the form of an algorithm for solving Problem 3.2.16 which is block-iterative in the sense that, at iteration $n \in \mathbb{N}$, only a subfamily of sets $(C_i)_{i\in I_n}$ needs to be activated, as opposed to all of them in (3.60). Block-iterative structures save time per iteration in two ways: firstly, they do not require that every constraint be activated; secondly, at every $n \in \mathbb{N}$, activation of each constraint indexed in $I_n$ can be performed in parallel and hence it is common to select card $I_n$ equal to the number of available processors. Furthermore, in [15, Section 6.5], the sets $(C_i)_{i\in I}$ were specified as lower level sets of certain functions and were activated by projections onto supersets instead of exact ones as in (3.59) and (3.60). Below, we propose an alternative block-iterative scheme (Algorithm 3.2.24) which is more sophisticated in that it leverages the affine structure of some sets $(C_i)_{i\in I'}$ to produce deeper relaxation steps, hence providing extra acceleration to the algorithm. Such affine-convex extrapolation techniques were first discussed in [6], where a weakly convergent method was designed to solve convex feasibility problems, i.e., to find an unspecified point in the intersection of closed convex sets. Additionally, as will be seen in Section 3.2.4, this new algorithm will be better suited to solve Problem 3.2.1 to the extent that it utilizes a fixed point model for the activation of the sets. The following notions and facts lay the groundwork for developing our best approximation algorithm.

**Definition 3.2.20** [5, Section 4.1] $\mathfrak{T}$ is the class of *firmly quasinonexpansive operators* from $\mathcal{H}$ to $\mathcal{H}$, i.e.,

$$\mathfrak{T} = \big\{T\colon \mathcal{H} \to \mathcal{H} \;\big|\; (\forall x \in \mathcal{H})(\forall y \in \operatorname{Fix} T)\ \langle y - Tx \mid x - Tx\rangle \leqslant 0\big\}. \tag{3.61}$$

**Example 3.2.21** [4, 5] Let $T\colon \mathcal{H} \to \mathcal{H}$ and set $C = \operatorname{Fix} T$. Then $T \in \mathfrak{T}$ in each of the following cases:

(i) $T$ is the projector onto a nonempty closed convex subset $C$ of $\mathcal{H}$.

(ii) $T$ is the proximity operator of a function $f \in \Gamma_0(\mathcal{H})$. Then $C = \operatorname{Argmin} f$.

(iii) $T$ is the resolvent of a maximally monotone operator $A\colon \mathcal{H} \to 2^{\mathcal{H}}$. Then $C = \big\{x \in \mathcal{H} \;\big|\; 0 \in Ax\big\}$ is the set of zeros of $A$.

(iv) $T$ is firmly nonexpansive.

(v) $R = 2T - \mathrm{Id}$ is quasinonexpansive: $(\forall x \in \mathcal{H})(\forall y \in \operatorname{Fix} R)\ \|Rx - y\| \leqslant \|x - y\|$. Then $C = \operatorname{Fix} R$.

(vi) $T$ is a subgradient projector onto the lower level set $C = \big\{x \in \mathcal{H} \;\big|\; f(x) \leqslant 0\big\} \neq \varnothing$ of a continuous convex function $f\colon \mathcal{H} \to \mathbb{R}$, that is, given a selection $s$ of the subdifferential of

$f$,

$$(\forall x \in \mathcal{H})\ \ Tx = \operatorname{sproj}_C x = \begin{cases} x - \dfrac{f(x)}{\|s(x)\|^2} s(x), & \text{if } \ f(x) > 0; \\ x, & \text{if } \ f(x) \leqslant 0. \end{cases} \tag{3.62}$$

**Lemma 3.2.22** [4, 5] *Let $T\colon \mathcal{H} \to \mathcal{H}$. If $T \in \mathfrak{T}$, then $\operatorname{Fix} T$ is closed and convex. Conversely, if $C$ is a nonempty closed convex subset of $\mathcal{H}$, then $C = \operatorname{Fix} T$, where $T = \operatorname{proj}_C \in \mathfrak{T}$.*

**Lemma 3.2.23** *Let $(T_n)_{n\in\mathbb{N}}$ be a sequence of operators in $\mathfrak{T}$ such that $\varnothing \neq C \subset \bigcap_{n\in\mathbb{N}} \operatorname{Fix} T_n$, let $x_0 \in \mathcal{H}$, let $Q$ be as in Proposition 3.2.17, and for every $n \in \mathbb{N}$, set $x_{n+1} = Q(x_0, x_n, T_n x_n)$. Then the following hold:*

(i) $(x_n)_{n\in\mathbb{N}}$ *is well defined.*

(ii) $\sum_{n\in\mathbb{N}} \|x_{n+1} - x_n\|^2 < +\infty$.

(iii) $\sum_{n\in\mathbb{N}} \|T_n x_n - x_n\|^2 < +\infty$.

(iv) $x_n \to \operatorname{proj}_C x_0$ *if and only if all the weak sequential cluster points of $(x_n)_{n\in\mathbb{N}}$ lie in $C$.*

*Proof*. In the case when $\varnothing \neq C = \bigcap_{n\in\mathbb{N}} \operatorname{Fix} T_n$, the results are shown in [4, Proposition 3.4(v) and Theorem 3.5]. However, an inspection of these proofs reveals that they remain true in our context. □

We are now in a position to introduce our best approximation algorithm for solving Problem 3.2.16. It incorporates ingredients of the best approximation method of [15, Section 6.5] and of the convex feasibility method of [6].

**Algorithm 3.2.24** Consider the setting of Problem 3.2.16 and denote by $(C_i)_{i\in I'}$ a subfamily of $(C_i)_{i\in I}$ of closed affine subspaces the projectors onto which are easy to implement; this subfamily is assumed to be nonempty as $\mathcal{H}$ can be included in it. Let $Q$ be as in Proposition 3.2.17, fix

$\varepsilon \in ]0,1[$, and iterate

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\quad \text{take } \mathrm{i}(n) \in I' \\
\quad z_n = \mathrm{proj}_{C_{\mathrm{i}(n)}} x_n \\
\quad \text{take a nonempty finite set } I_n \subset I \\
\quad \text{for every } i \in I_n \\
\quad\quad \text{take } T_{i,n} \in \mathfrak{T} \text{ such that } \operatorname{Fix} T_{i,n} = C_i \\
\quad\quad a_{i,n} = T_{i,n} z_n \\
\quad\quad \theta_{i,n} = \|a_{i,n} - z_n\|^2 \\
\quad \text{take } j_n \in I_n \text{ such that } \theta_{j_n,n} = \max_{i \in I_n} \theta_{i,n} \\
\quad \text{take } (\omega_{i,n})_{i \in I_n} \subset [0,1] \text{ such that } \sum_{i \in I_n} \omega_{i,n} = 1 \text{ and } \omega_{j_n,n} \geqslant \varepsilon \\
\quad I_n^+ = \big\{ i \in I_n \;\big|\; \omega_{i,n} > 0 \big\} \\
\quad \theta_n = \sum_{i \in I_n^+} \omega_{i,n} \theta_{i,n} \\
\quad \text{if } \theta_n = 0 \\
\quad\quad t_n = z_n \\
\quad \text{else} \\
\quad\quad d_n = \sum_{i \in I_n^+} \omega_{i,n} a_{i,n} \\
\quad\quad y_n = \mathrm{proj}_{C_{\mathrm{i}(n)}} d_n - z_n \\
\quad\quad \text{take } \lambda_n \in \big[\varepsilon \theta_n / \|d_n - z_n\|^2, \theta_n / \|y_n\|^2\big] \\
\quad\quad t_n = z_n + \lambda_n y_n \\
\quad x_{n+1} = Q(x_0, x_n, t_n).
\end{array}
\tag{3.63}
$$

**Remark 3.2.25** Let us highlight some special cases and features of Algorithm 3.2.24.

(i) If the only closed affine subspace is $\mathcal{H}$ then, for every $n \in \mathbb{N}$, $z_n = x_n$, and the resulting algorithm has a structure similar to that of [15, Section 6.5], except that the operators $(T_{i,n})_{i \in I_n}$ are chosen differently. In particular, this setting captures (3.59) and (3.60).

(ii) Suppose that the last step of the algorithm at iteration $n \in \mathbb{N}$ is replaced by $x_{n+1} = t_n$. Then we recover an instance of the (weakly convergent) convex feasibility algorithm of [6] to find an unspecified point in $C = \bigcap_{i \in I} C_i$.

(iii) At iteration $n \in \mathbb{N}$, a block of sets $(C_i)_{i \in I_n}$ is selected and each of its elements is activated via a firmly quasinonexpansive operator. Example 3.2.21 provides various options to choose these operators, depending on the nature of the sets.

(iv) If nontrivial affine sets are present then, at iteration $n \in \mathbb{N}$, we have $z_n \neq x_n$ in general. Thus, as discussed in [10] and its references in the context of feasibility algorithms (see (ii)), the resulting step $t_n$ is larger than when $z_n = x_n$, which typically yields faster convergence. This point will be illustrated numerically for our best approximation algorithm in Section 3.2.5.

We now establish the strong convergence of an arbitrary sequence $(x_n)_{n\in\mathbb{N}}$ generated by Algorithm 3.2.24 to the solution to Problem 3.2.16. The last component of the proof relies on Lemma 3.2.23(iv), i.e., showing that the weak sequential cluster points of $(x_n)_{n\in\mathbb{N}}$ lie in $C$. The same property is required in [6, Theorem 3.3] to show the weak convergence of the variant described in Remark 3.2.25(ii). This parallels the weak-to-strong convergence principle of [4], namely the transformation of weakly convergent feasibility methods into strongly convergent best approximation methods.

**Theorem 3.2.26** *In the setting of Problem 3.2.16, let $(x_n)_{n\in\mathbb{N}}$ be generated by Algorithm 3.2.24. Suppose that the following hold:*

[a] *There exist strictly positive integers $(M_i)_{i\in I}$ such that*

$$(\forall i \in I)(\forall n \in \mathbb{N}) \quad i \in \bigcup_{l=n}^{n+M_i-1} \{\mathrm{i}(l)\} \cup I_l. \tag{3.64}$$

[b] *For every $i \in I \smallsetminus I'$, every $x \in \mathcal{H}$, and every strictly increasing sequence $(r_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$,*

$$\begin{cases} i \in \bigcap_{n\in\mathbb{N}} I_{r_n} \\ \operatorname{proj}_{C_{\mathrm{i}(r_n)}} x_{r_n} \rightharpoonup x \\ T_{i,r_n}\big(\operatorname{proj}_{C_{\mathrm{i}(r_n)}} x_{r_n}\big) - \operatorname{proj}_{C_{\mathrm{i}(r_n)}} x_{r_n} \to 0 \end{cases} \quad \Rightarrow \quad x \in C_i. \tag{3.65}$$

*Then $x_n \to \operatorname{proj}_C x_0$.*

*Proof*. Let us fix $n \in \mathbb{N}$ temporarily. Define

$$L_n\colon \mathcal{H} \to \mathbb{R}\colon z \mapsto \begin{cases} \dfrac{\sum_{i\in I_n^+} \omega_{i,n}\|T_{i,n}z - z\|^2}{\big\|\sum_{i\in I_n^+} \omega_{i,n}T_{i,n}z - z\big\|^2}, & \text{if } z \notin \bigcap_{i\in I_n^+} C_i; \\ 1, & \text{if } z \in \bigcap_{i\in I_n^+} C_i \end{cases} \tag{3.66}$$

and

$$S_n\colon \mathcal{H} \to \mathcal{H}\colon z \mapsto z + L_n(z)\bigg(\sum_{i\in I_n^+} \omega_{i,n}T_{i,n}z - z\bigg). \tag{3.67}$$

We derive from [16, Proposition 2.4] that $S_n \in \mathfrak{T}$ and $\operatorname{Fix} S_n = \bigcap_{i\in I_n^+} \operatorname{Fix} T_{i,n} = \bigcap_{i\in I_n^+} C_i$. We also observe that

$$\theta_n = 0 \;\Leftrightarrow\; S_n z_n = z_n \;\Leftrightarrow\; z_n \in \bigcap_{i\in I_n^+} C_i = \operatorname{Fix} S_n. \tag{3.68}$$

Now define

$$K_n\colon \mathcal{H}\to\mathbb{R}\colon x\mapsto \begin{cases} \dfrac{\big\|S_n\big(\mathrm{proj}_{C_{\mathrm{i}(n)}}x\big)-\mathrm{proj}_{C_{\mathrm{i}(n)}}x\big\|^2}{\big\|\mathrm{proj}_{C_{\mathrm{i}(n)}}\big(S_n\big(\mathrm{proj}_{C_{\mathrm{i}(n)}}x\big)\big)-\mathrm{proj}_{C_{\mathrm{i}(n)}}x\big\|^2}, & \text{if } \mathrm{proj}_{C_{\mathrm{i}(n)}}x\notin\bigcap_{i\in I_n^+}C_i;\\ 1, & \text{if } \mathrm{proj}_{C_{\mathrm{i}(n)}}x\in\bigcap_{i\in I_n^+}C_i \end{cases} \tag{3.69}$$

and

$$T_n\colon \mathcal{H}\to\mathcal{H}\colon x\mapsto \mathrm{proj}_{C_{\mathrm{i}(n)}}x+\gamma_n(x)\Big(\mathrm{proj}_{C_{\mathrm{i}(n)}}\Big(S_n\big(\mathrm{proj}_{C_{\mathrm{i}(n)}}x\big)\Big)-\mathrm{proj}_{C_{\mathrm{i}(n)}}x\Big),$$
$$\text{where}\quad \gamma_n(x)\in[\varepsilon,K_n(x)]\,. \tag{3.70}$$

Then it follows from [6, Theorem 2.8] that $T_n\in\mathfrak{T}$ and

$$\varnothing\neq C\subset C_{\mathrm{i}(n)}\cap\bigcap_{i\in I_n^+}C_i=C_{\mathrm{i}(n)}\cap\operatorname{Fix}S_n=\operatorname{Fix}T_n. \tag{3.71}$$

If $\theta_n\neq 0$, using (3.63), (3.67), and the fact that $\mathrm{proj}_{C_{\mathrm{i}(n)}}$ is an affine operator [5, Corollary 3.22(ii)], we obtain

$$\begin{aligned}\mathrm{proj}_{C_{\mathrm{i}(n)}}\Big(S_n\big(\mathrm{proj}_{C_{\mathrm{i}(n)}}x_n\big)\Big)-\mathrm{proj}_{C_{\mathrm{i}(n)}}x_n&=\mathrm{proj}_{C_{\mathrm{i}(n)}}(S_nz_n)-z_n\\ &=\mathrm{proj}_{C_{\mathrm{i}(n)}}\Big(\big(1-L_n(z_n)\big)z_n+L_n(z_n)d_n\Big)-z_n\\ &=\big(1-L_n(z_n)\big)\mathrm{proj}_{C_{\mathrm{i}(n)}}z_n+L_n(z_n)\mathrm{proj}_{C_{\mathrm{i}(n)}}d_n-z_n\\ &=L_n(z_n)\big(\mathrm{proj}_{C_{\mathrm{i}(n)}}d_n-z_n\big)\\ &=L_n(z_n)y_n\end{aligned} \tag{3.72}$$

and, therefore,

$$\big\|\mathrm{proj}_{C_{\mathrm{i}(n)}}(S_nz_n)-z_n\big\|=L_n(z_n)\|y_n\|. \tag{3.73}$$

Hence, (3.69) and (3.67) yield

$$K_n(x_n)=\begin{cases}\dfrac{\|S_n(z_n)-z_n\|^2}{\big\|\mathrm{proj}_{C_{\mathrm{i}(n)}}\big(S_nz_n\big)-z_n\big\|^2}=\dfrac{\|L_n(z_n)(d_n-z_n)\|^2}{\|L_n(z_n)y_n\|^2}=\dfrac{\|d_n-z_n\|^2}{\|y_n\|^2}, & \text{if } \theta_n\neq 0;\\ 1, & \text{if } \theta_n=0.\end{cases} \tag{3.74}$$

At the same time, we derive from (3.66), (3.63), and (3.68) that

$$L_n(z_n)=\begin{cases}\dfrac{\theta_n}{\|d_n-z_n\|^2}, & \text{if } \theta_n\neq 0;\\ 1, & \text{if } \theta_n=0.\end{cases} \tag{3.75}$$

Altogether, it results from (3.70), (3.74), and (3.75) that, if $\theta_n \neq 0$,

$$\gamma_n(x_n)L_n(z_n) \in \big[\varepsilon L_n(z_n), K_n(x_n)L_n(z_n)\big] = \big[\varepsilon\theta_n/\|d_n - z_n\|^2, \theta_n/\|y_n\|^2\big] \tag{3.76}$$

and, in view of (3.63), we can therefore set $\lambda_n = \gamma_n(x_n)L_n(z_n)$. Thus, it follows from (3.63) and (3.72) that

$$\begin{aligned}
\theta_n \neq 0 \;\Rightarrow\; t_n &= z_n + \lambda_n y_n \\
&= z_n + \gamma_n(x_n)L_n(z_n)y_n \\
&= \operatorname{proj}_{C_{\mathrm{i}(n)}} x_n + \gamma_n(x_n)\Big(\operatorname{proj}_{C_{\mathrm{i}(n)}}\big(S_n\big(\operatorname{proj}_{C_{\mathrm{i}(n)}} x_n\big)\big) - \operatorname{proj}_{C_{\mathrm{i}(n)}} x_n\Big) \\
&= T_n x_n.
\end{aligned} \tag{3.77}$$

On the other hand, (3.63) and (3.68) yield

$$\theta_n = 0 \;\Rightarrow\; t_n = z_n = S_n z_n = T_n x_n. \tag{3.78}$$

Combining (3.77) and (3.78), we obtain

$$x_{n+1} = Q(x_0, x_n, T_n x_n). \tag{3.79}$$

Turning back to (3.70) and (3.63), we deduce from [5, Corollary 3.22(i)] that

$$\begin{aligned}
\|T_n x_n - x_n\|^2 &= \big\|z_n - x_n + \gamma_n(x_n)\big(\operatorname{proj}_{C_{\mathrm{i}(n)}}(S_n z_n) - z_n\big)\big\|^2 \\
&= \|z_n - x_n\|^2 + 2\gamma_n(x_n)\langle \operatorname{proj}_{C_{\mathrm{i}(n)}} x_n - x_n \mid \operatorname{proj}_{C_{\mathrm{i}(n)}}(S_n z_n) - \operatorname{proj}_{C_{\mathrm{i}(n)}} x_n\rangle \\
&\quad + |\gamma_n(x_n)|^2\big\|\operatorname{proj}_{C_{\mathrm{i}(n)}}(S_n z_n) - z_n\big\|^2 \\
&= \|z_n - x_n\|^2 + |\gamma_n(x_n)|^2\big\|\operatorname{proj}_{C_{\mathrm{i}(n)}}(S_n z_n) - z_n\big\|^2 \\
&\geqslant \|z_n - x_n\|^2 + \varepsilon^2\big\|\operatorname{proj}_{C_{\mathrm{i}(n)}}(S_n z_n) - z_n\big\|^2.
\end{aligned} \tag{3.80}$$

Since (3.71) implies that

$$\varnothing \neq C \subset \bigcap_{n\in\mathbb{N}} \operatorname{Fix} T_n, \tag{3.81}$$

we derive from (3.79) and Lemma 3.2.23(i) that $(x_n)_{n\in\mathbb{N}}$ is well defined. Furthermore, (3.80) and Lemma 3.2.23(iii) guarantee that

$$\sum_{n\in\mathbb{N}} \|z_n - x_n\|^2 < +\infty \tag{3.82}$$

and

$$\sum_{n\in\mathbb{N}} \|\operatorname{proj}_{C_{\mathrm{i}(n)}}(S_n z_n) - z_n\|^2 < +\infty. \tag{3.83}$$

Finally, in view of (3.81) and Lemma 3.2.23(iv), to conclude the proof, it is enough to show that all the weak sequential cluster points of $(x_n)_{n\in\mathbb{N}}$ lie in $C$. Since we have at our disposal [a], [b], (3.82), and (3.83), showing this inclusion can be done by following the same steps as in the proof of [6, Theorem 3.3(vi)]. □

**Remark 3.2.27** Condition [a] in Theorem 3.2.26 states that, for each $i \in I$, the set $C_i$ should be involved at least once every $M_i$ iterations. Condition [b] in Theorem 3.2.26 is discussed in [6, Section 3.4], where concrete scenarios that satisfy it are described.

### 3.2.4 Fixed point model and algorithm for Problem 3.2.1

To solve Problem 3.2.1, we are going to reformulate it as an instance of Problem 3.2.16. To this end, let us set

$$(\forall k \in K) \quad C_k = \big\{x \in \mathcal{H} \;\big|\; F_k x = p_k\big\} \quad \text{and} \quad T_k = p_k + \mathrm{Id} - F_k. \tag{3.84}$$

Then it follows from (3.9) that

$$(\forall k \in K) \quad T_k \text{ is firmly nonexpansive and } \operatorname{Fix} T_k = C_k. \tag{3.85}$$

We therefore deduce from Lemma 3.2.22 that $(C_k)_{k\in K}$ are closed convex subsets of $\mathcal{H}$. Thus, upon setting $I = J \cup K$, we recast Problem 3.2.1 is an instantiation of Problem 3.2.16. This leads us to the following solution method based on Algorithm 3.2.24.

**Proposition 3.2.28** *In the setting of Problem 3.2.1, let $Q$ be as in Proposition 3.2.17, fix $\varepsilon \in \left]0, 1\right[$, and denote by $(C_i)_{i\in I'}$ a subfamily of $(C_i)_{i\in J}$ of closed affine subspaces the projectors onto which*

*are easy to implement; this subfamily is assumed to be nonempty as $\mathcal{H}$ can be included in it. Iterate*

$$
\begin{array}{l}
\textit{for } n = 0, 1, \ldots \\
\quad \textit{take } \mathrm{i}(n) \in I' \\
\quad z_n = \mathrm{proj}_{C_{\mathrm{i}(n)}} x_n \\
\quad \textit{take a nonempty finite set } I_n \subset J \cup K \\
\quad \textit{for every } i \in I_n \\
\quad\quad \textit{if } i \in J \\
\quad\quad\quad \textit{take } T_{i,n} \in \mathfrak{T} \textit{ such that } \operatorname{Fix} T_{i,n} = C_i \\
\quad\quad\quad a_{i,n} = T_{i,n} z_n \\
\quad\quad \textit{if } i \in K \\
\quad\quad\quad a_{i,n} = p_i + z_n - F_i z_n \\
\quad\quad \theta_{i,n} = \|a_{i,n} - z_n\|^2 \\
\quad \textit{take } j_n \in I_n \textit{ such that } \theta_{j_n,n} = \max_{i \in I_n} \theta_{i,n} \\
\quad \textit{take } \{\omega_{i,n}\}_{i \in I_n} \subset [0,1] \textit{ such that } \sum_{i \in I_n} \omega_{i,n} = 1 \textit{ and } \omega_{j_n,n} \geqslant \varepsilon \\
\quad I_n^+ = \big\{ i \in I_n \;\big|\; \omega_{i,n} > 0 \big\} \\
\quad \theta_n = \sum_{i \in I_n^+} \omega_{i,n} \theta_{i,n} \\
\quad \textit{if } \theta_n = 0 \\
\quad\quad t_n = z_n \\
\quad \textit{else} \\
\quad\quad d_n = \sum_{i \in I_n^+} \omega_{i,n} a_{i,n} \\
\quad\quad y_n = \mathrm{proj}_{C_{\mathrm{i}(n)}} d_n - z_n \\
\quad\quad \textit{take } \lambda_n \in \big[ \varepsilon \theta_n / \|d_n - z_n\|^2, \theta_n / \|y_n\|^2 \big] \\
\quad\quad t_n = z_n + \lambda_n y_n \\
\quad x_{n+1} = Q(x_0, x_n, t_n).
\end{array}
\tag{3.86}
$$

*Suppose that condition* [a] *in Theorem 3.2.26 holds with $I = J \cup K$, as well as the following:*

[c] *For every $i \in J \smallsetminus I'$, every $x \in \mathcal{H}$, and every strictly increasing sequence $(r_n)_{n \in \mathbb{N}}$ in $\mathbb{N}$,* (3.65) *holds.*

*Then $(x_n)_{n \in \mathbb{N}}$ converges strongly to the solution to Problem 3.2.1.*

*Proof.* Let us bring into play (3.84) and (3.85). As discussed above, Problem 3.2.1 is an instance of Problem 3.2.16, where $I = J \cup K$. Now set

$$
(\forall k \in K)(\forall n \in \mathbb{N}) \quad T_{k,n} = T_k = p_k + \mathrm{Id} - F_k. \tag{3.87}
$$

Then (3.63) reduces to (3.86) and, in view of condition [c] above, to conclude via Theorem 3.2.26, it suffices to check that condition [b] in Theorem 3.2.26 holds for every $k \in K$. Towards this goal, let us fix $k \in K$ and a strictly increasing sequence $(r_n)_{n \in \mathbb{N}}$ in $\mathbb{N}$ such that $k \in \bigcap_{n \in \mathbb{N}} I_{r_n}$, and let us set $(\forall n \in \mathbb{N})\ u_n = \mathrm{proj}_{C_{\mathrm{i}(r_n)}} x_{r_n}$. Suppose that $u_n \rightharpoonup x \in \mathcal{H}$ and that

$T_{k,r_n}u_n - u_n \to 0$. Then (3.87) yields $T_k u_n - u_n \to 0$ and, since $T_k$ is nonexpansive by (3.85), it follows from Browder's demiclosedness principle [5, Corollary 4.28] that $x \in \operatorname{Fix} T_k = C_k$, which concludes the proof. □

As was mentioned in Remark 3.2.25(iv) and will be illustrated in Section 3.2.5, exploiting the presence of affine subspaces typically leads to faster convergence. Problem 3.2.1 can nonetheless be solved without taking the affine subspaces into account. Formally, this amounts to considering that $(C_i)_{i\in I'}$ consists solely of $\mathcal{H}$, in which case Proposition 3.2.28 leads to the following implementation.

**Corollary 3.2.29** *In the setting of Problem 3.2.1, let $Q$ be as in Proposition 3.2.17, and fix $\varepsilon \in ]0,1[$. Iterate*

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\quad \text{take a nonempty finite set } I_n \subset J \cup K \\
\quad \text{for every } i \in I_n \\
\quad\quad \text{if } i \in J \\
\quad\quad\quad \text{take } T_{i,n} \in \mathfrak{T} \text{ such that } \operatorname{Fix} T_{i,n} = C_i \\
\quad\quad\quad a_{i,n} = T_{i,n}x_n \\
\quad\quad \text{if } i \in K \\
\quad\quad\quad a_{i,n} = p_i + x_n - F_i x_n \\
\quad\quad \theta_{i,n} = \|a_{i,n} - x_n\|^2 \\
\quad \text{take } j_n \in I_n \text{ such that } \theta_{j_n,n} = \max_{i\in I_n} \theta_{i,n} \\
\quad \text{take } (\omega_{i,n})_{i\in I_n} \subset [0,1] \text{ such that } \sum_{i\in I_n} \omega_{i,n} = 1 \text{ and } \omega_{j_n,n} \geqslant \varepsilon \\
\quad I_n^+ = \big\{ i \in I_n \;\big|\; \omega_{i,n} > 0 \big\} \\
\quad \theta_n = \sum_{i\in I_n^+} \omega_{i,n}\theta_{i,n} \\
\quad \text{if } \theta_n = 0 \\
\quad\quad t_n = x_n \\
\quad \text{else} \\
\quad\quad y_n = \sum_{i\in I_n^+} \omega_{i,n} a_{i,n} - x_n \\
\quad\quad \text{take } \lambda_n \in \big[\varepsilon\theta_n/\|y_n\|^2, \theta_n/\|y_n\|^2\big] \\
\quad\quad t_n = x_n + \lambda_n y_n \\
\quad x_{n+1} = Q(x_0, x_n, t_n).
\end{array}
\tag{3.88}
$$

*Suppose that the following hold:*

[d] *There exist strictly positive integers $(M_i)_{i\in J\cup K}$ such that $(\forall i \in J \cup K)(\forall n \in \mathbb{N})\ i \in \bigcup_{l=n}^{n+M_i-1} I_l$.*

[e] *For every $i \in J$, every $x \in \mathcal{H}$, and every strictly increasing sequence $(r_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$,*

$$
\Big[ i \in \bigcap_{n\in\mathbb{N}} I_{r_n},\ x_{r_n} \rightharpoonup x,\ \text{and } T_{i,r_n}x_{r_n} - x_{r_n} \to 0 \Big] \quad \Rightarrow \quad x \in C_i. \tag{3.89}
$$

*Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to the solution to Problem 3.2.1.*

### 3.2.5 Numerical illustration

Let $\mathcal{H}$ be the standard Euclidean space $\mathbb{R}^N$, where $N = 1024$. The goal is to recover the original form of the signal $\overline{x} \in \mathcal{H}$ shown in Figure 3.1 from the following:

(i) $\overline{x}$ resides in the subspace $C_1$ of signals which are band-limited in the sense that their discrete Fourier transform vanishes outside of the $103$ lowest frequency components.

(ii) Let $\mathrm{tv} : \mathcal{H} \to \mathbb{R} \colon x = (\xi_i)_{1\leqslant i\leqslant N} \mapsto \sum_{1\leqslant i\leqslant N-1} |\xi_{i+1} - \xi_i|$ be the total variation function. An upper bound $\gamma \in \left]0, +\infty\right[$ on $\mathrm{tv}\,(\overline{x})$ is available. The associated constraint set is $C_2 = \big\{x \in \mathcal{H} \;\big|\; \mathrm{tv}\,(x) - \gamma \leqslant 0\big\}$. For this experiment, $\gamma = 1.5\mathrm{tv}\,(\overline{x})$.

(iii) $25$ observations $(q_k)_{k\in K}$ are available where, for every $k \in K = \{3, \ldots, 27\}$, $q_k$ is the isotonic regression of the coefficients of $\overline{x}$ in a dictionary $(e_{k,j})_{1\leqslant j\leqslant 10}$ of vectors in $\mathcal{H}$. More precisely (see Example 3.2.12(i)), set $\mathcal{G} = \mathbb{R}^{10}$ and $D = \big\{(\xi_j)_{1\leqslant j\leqslant 10} \in \mathcal{G} \;\big|\; \xi_1 \leqslant \cdots \leqslant \xi_{10}\big\}$. Then, for every $k \in K$, $q_k = \mathrm{proj}_D(L_k\overline{x})$, where $L_k \colon \mathcal{H} \to \mathcal{G} \colon x \mapsto (\langle x \mid e_{k,j}\rangle)_{1\leqslant j\leqslant 10}$.

We seek the minimal-energy signal consistent with the information above, i.e., we seek to

$$\text{minimize } \|x\| \quad \text{subject to} \quad x \in C_1 \cap C_2 \quad \text{and} \quad (\forall k \in K) \quad \mathrm{proj}_D(L_k x) = q_k. \tag{3.90}$$

Let us set $x_0 = 0$, $J = \{1, 2\}$, and, for every $k \in K$, $p_k = \|L_k\|^{-2} L_k^* q_k$, and $F_k = \|L_k\|^{-2} L_k^* \circ \mathrm{proj}_D \circ L_k$. For every $k \in K$, applying Proposition 3.2.9 with $\mathbb{I} = \{k\}$, $\mathcal{G}_k = \mathcal{G}$, $\beta_k = 1$, and $Q_k = \mathrm{proj}_D$ shows that $p_k$ is the proximal point of $\overline{x}$ relative to $F_k$ and, for every $x \in \mathcal{H}$, $F_k x = p_k \Leftrightarrow \mathrm{proj}_D(L_k x) = q_k$. We therefore arrive at an instance of Problem 3.2.1 which is equivalent to (3.90), namely

$$\text{minimize } \|x\| \quad \text{subject to} \quad x \in C_1 \cap C_2 \quad \text{and} \quad (\forall k \in K) \quad F_k x = p_k. \tag{3.91}$$

With an eye towards algorithm (3.86), since $C_1$ is an affine subspace with a straightforward projector [59], set $I' = \{1\}$. At iteration $n \in \mathbb{N}$, the constraint (ii) is activated by the subgradient projector $T_{2,n} = \mathrm{sproj}_{C_2}$ of (3.62) (see [19] for its computation) since the direct projector is hard to implement. The fact that condition [c] in Proposition 3.2.28 is satisfied follows from [5, Proposition 29.41(vi)(a)]. We solve (3.91) with algorithm (3.86) to obtain the solution $x_\infty$ shown in Figure 3.2 (see [33, Algorithm 8.1.1] for the computation of $\mathrm{proj}_D$).

To demonstrate the benefits of exploiting the presence of affine subspaces in algorithm (3.86), we show in Figure 3.3 the approximate solution it generates after $1000$ iterations. For the sake of comparison, we display in Figure 3.4 the approximate solution generated by algorithm (3.88) after $1000$ iterations. The following parameters are used:

- **Algorithm** (3.86)**:** For every $n \in \mathbb{N}$, $\mathrm{i}(n) = 1$, and whenever $\theta_n \neq 0$,

$$\lambda_n = \begin{cases} \dfrac{\theta_n}{2\|y_n\|^2}, & \text{if } n \equiv 0 \mod 3; \\ \dfrac{\theta_n}{\|y_n\|^2}, & \text{if } n \not\equiv 0 \mod 3. \end{cases} \tag{3.92}$$

  Additionally, $I_n$ is selected to activate $C_2$ at every iteration and periodically sweep through one entry of $K$ per iteration, hence satisfying condition [a] in Theorem 3.2.26 with $M_1 = M_2 = 1$, and, for every $k \in K$, $M_k = 25$. Moreover, for every $i \in I_n$, $\omega_{i,n} = 1/2$.

- **Algorithm** (3.88)**:** Iteration $n \in \mathbb{N}$ is executed with the same relaxation scheme (3.92) as in algorithm (3.86), and the same choice of the activation set $I_n$, with the exception that $I_n$ also activates $C_1$ at every iteration. In addition, for every $i \in I_n$, $\omega_{i,n} = 1/3$.

While both approaches are equivalent means of solving (3.91), Figures 3.3 and 3.4 demonstrate qualitatively that algorithm (3.86) yields faster convergence to the solution $x_\infty$ than algorithm (3.88). This is confirmed quantitatively by the error plots of Figure 3.5.

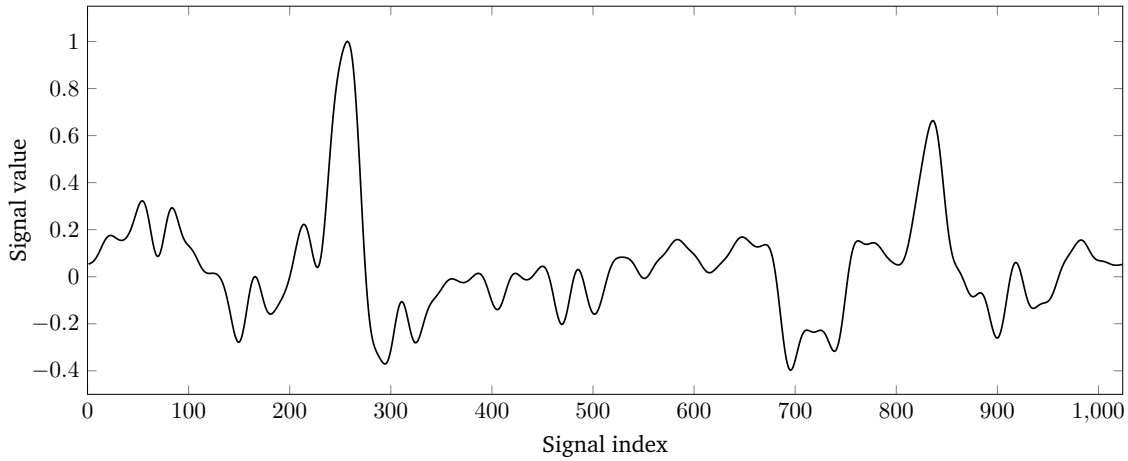



**Figure 3.1** Original signal $\overline{x}$.

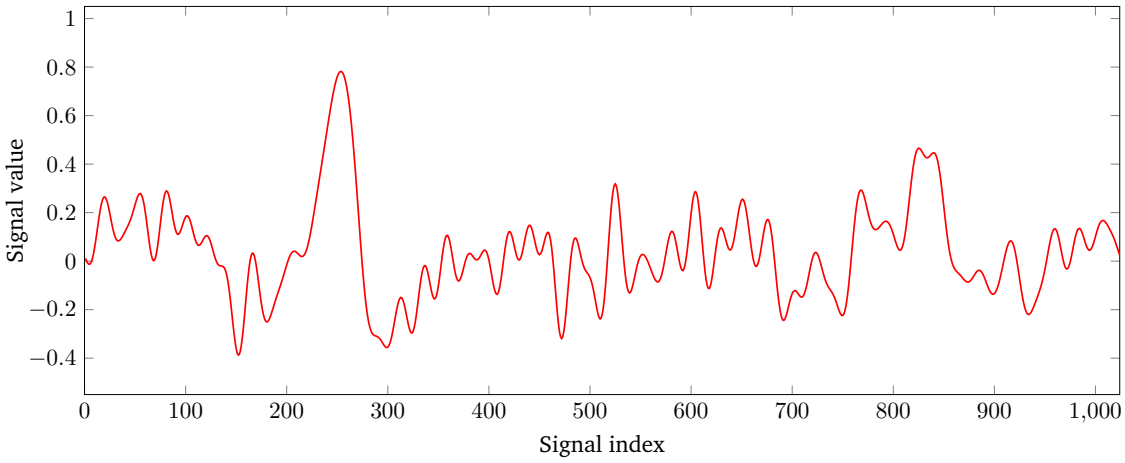



**Figure 3.2** Solution $x_\infty$ to (3.90).

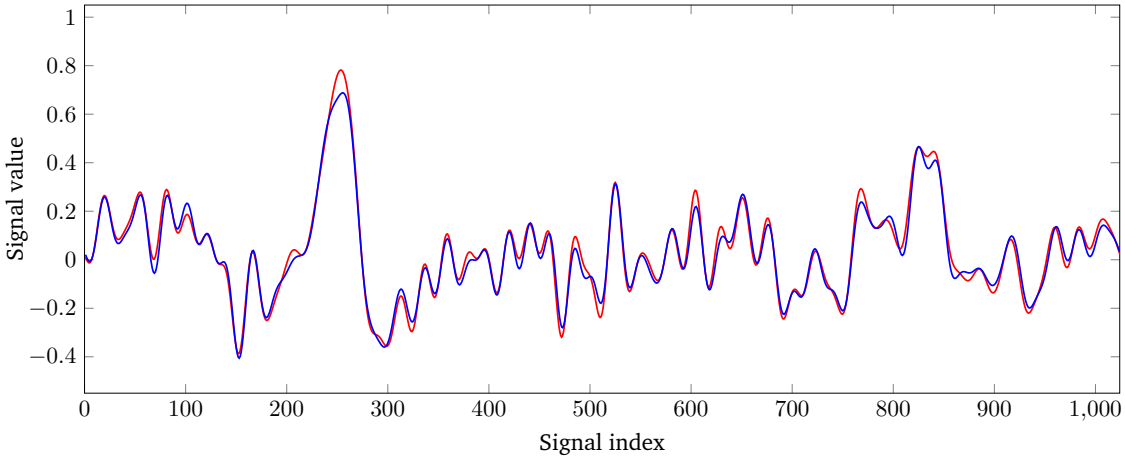



**Figure 3.3** Solution $x_\infty$ (red) and the approximate recovery obtained with 1000 iterations of algorithm (3.86), which exploits affine constraints (blue).

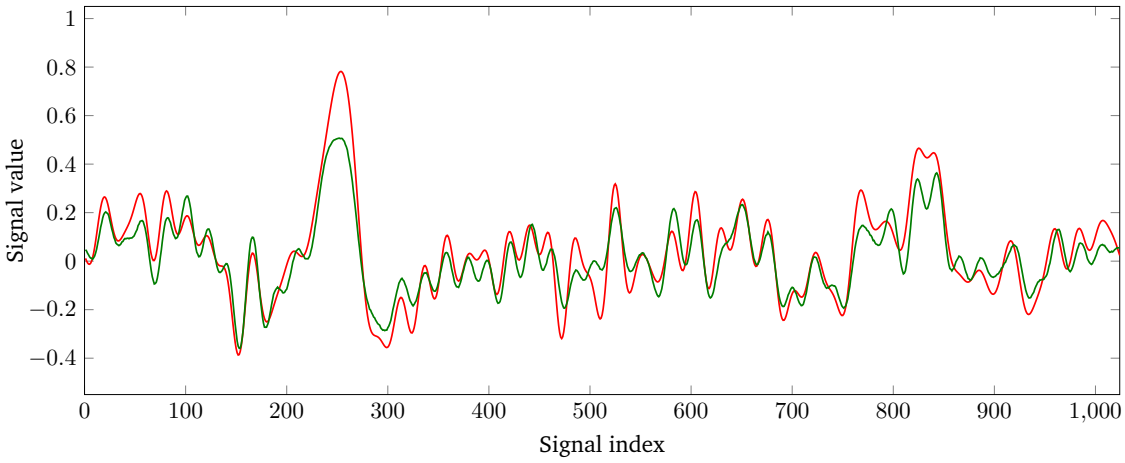



**Figure 3.4** Solution $x_\infty$ (red) and the approximate recovery obtained with 1000 iterations of algorithm (3.88), which does not exploit affine constraints (green).

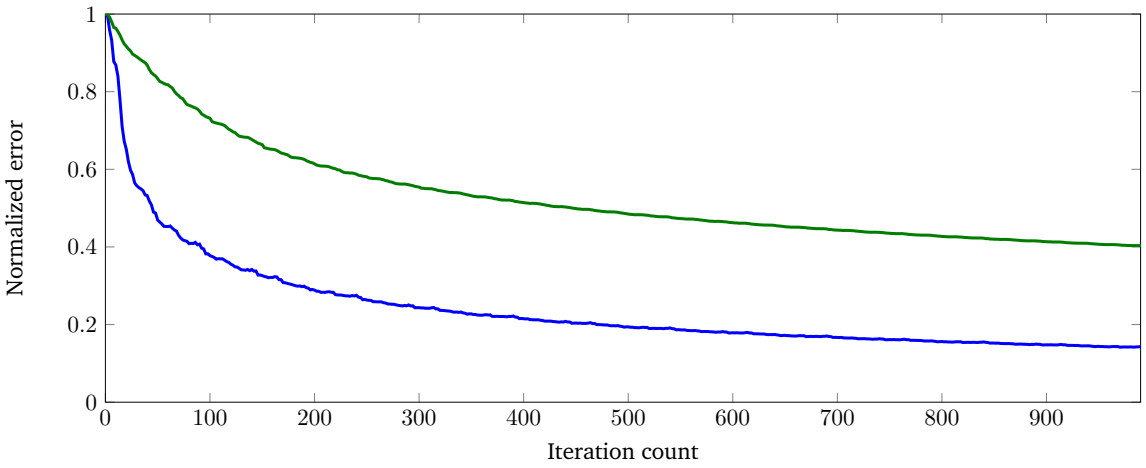



**Figure 3.5** Normalized error $\|x_n - x_\infty\|/\|x_0 - x_\infty\|$ versus iteration count $n \in \{0, \ldots, 1000\}$ for algorithm (3.86) (blue) and algorithm (3.88) (green).

## References


[1] F. Abramovich, T. Sapatinas, and B. W. Silverman, Wavelet thresholding via a Bayesian approach, *J. R. Stat. Soc. Ser. B Stat. Methodol.*, vol. 60, pp. 725–749, 1998.

[2] F. R. Ávila, M. P. Tcheou, and L. W. P. Biscainho, Audio soft declipping based on constrained weighted least squares, *IEEE Signal Process. Lett.*, vol. 24, pp. 1348–1352, 2017.

[3] A. L. Babierra and N. N. Reyes, A new characterization of the generalized inverse using projections on level sets, *J. Approx. Theory*, vol. 236, pp. 23–35, 2018.

[4] H. H. Bauschke and P. L. Combettes, A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces, *Math. Oper. Res.*, vol. 26, pp. 248–264, 2001.

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

[6] H. H. Bauschke, P. L. Combettes, and S. G. Kruk, Extrapolation algorithm for affine-convex feasibility problems, *Numer. Algorithms*, vol. 41, pp. 239–274, 2006.

[7] H. Boche, M. Guillemard, G. Kutyniok, and F. Philipp, Signal recovery from thresholded frame measurements, *Proc. 15th SPIE Wavelets Sparsity Conf.*, vol. 8858, pp. 80–86, 2013.

[8] L. M. Briceño-Arias and P. L. Combettes, Convex variational formulation with smooth coupling for multicomponent signal decomposition and recovery, *Numer. Math. Theory Methods Appl.*, vol. 2, pp. 485–508, 2009.

[9] B. Brogliato, *Nonsmooth Mechanics: Models, Dynamics and Control*, 3rd ed. Springer, New York, 2016.

[10] Y. Censor, W. Chen, P. L. Combettes, R. Davidi, and G. T. Herman, On the effectiveness of projection methods for convex feasibility problems with linear inequality constraints, *Comput. Optim. Appl.*, vol. 51, pp. 1065–1088, 2012.

[11] A. Chambolle, R. A. DeVore, N. Y. Lee, and B. J. Lucier, Nonlinear wavelet image processing: Variational problems, compression, and noise removal through wavelet shrinkage, *IEEE Trans. Image Process.*, vol. 7, pp. 319–335, 1998.

[12] G. Chierchia, N. Pustelnik, B. Pesquet-Popescu, and J.-C. Pesquet, A nonlocal structure tensor-based approach for multicomponent image recovery problems, *IEEE Trans. Image Process.*, vol. 23, pp. 5531–5544, 2014.

[13] C. K. Chui, F. Deutsch, and J. D. Ward, Constrained best approximation in Hilbert space II, *J. Approx. Theory*, vol. 71, pp. 213–238, 1992.

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

[15] P. L. Combettes, Strong convergence of block-iterative outer approximation methods for convex optimization, *SIAM J. Control Optim.*, vol. 38, pp. 538–565, 2000.

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

[17] P. L. Combettes, A block-iterative surrogate constraint splitting method for quadratic signal recovery, *IEEE Trans. Signal Process.*, vol. 51, pp. 1771–1782, 2003.

[18] P. L. Combettes, Monotone operator theory in convex optimization, *Math. Program.*, vol. B170, pp. 177–206, 2018.

[19] P. L. Combettes and J.-C. Pesquet, Image restoration subject to a total variation constraint, *IEEE Trans. Image Process.,* vol. 13, pp. 1213–1222, 2004.

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

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

[22] P. L. Combettes and N. N. Reyes, Functions with prescribed best linear approximations, *J. Approx. Theory*, vol. 162, pp. 1095–1116, 2010.

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

[24] R. Dai, H. Song, R. Foygel Barber, and G. Raskutti, The bias of isotonic regression, *Electron. J. Stat.*, vol. 14, pp. 801–834, 2020.

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

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

[27] F. Deutsch, W. Li, and J. D. Ward, A dual approach to constrained interpolation from a convex subset of Hilbert space, *J. Approx. Theory*, vol. 90, pp. 385–414, 1997.

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

[29] S. Enderby and Z. Baracska, Harmonic instability of digital soft clipping algorithms, *Proc. 15th Int. Conf. Digital Audio Effects*, pp. DAFX-1–DAFX-5. York, UK, Sept. 17–21, 2012.

[30] J. Favard, Sur l'interpolation, *J. Math. Pures Appl.*, vol. 19, pp. 281–306, 1940.

[31] S. Foucart and T. Needham, Sparse recovery from saturated measurements, *Inf. Inference*, vol. 6, pp. 196–212, 2017.

[32] L. Gosse, A Donoho-Stark criterion for stable signal recovery in discrete wavelet subspaces, *J. Comput. Appl. Math.*, vol. 235, pp 5024–5039, 2011.

[33] W. Härdle, *Applied Nonparametric Regression.* Cambridge University Press, Cambridge, 1990.

[34] Y. Haugazeau, *Sur les Inéquations Variationnelles et la Minimisation de Fonctionnelles Convexes*. Thèse, Université de Paris, 1968.

[35] P. Jaming, Nazarov's uncertainty principles in higher dimension, *J. Approx. Theory*, vol. 149, pp. 30–41, 2007.

[36] K. Konda, R. Memisevic, and D. Krueger, Zero-bias autoencoders and the benefits of co-adapting features, *Proc. Int. Conf. Learn. Represent.*, San Diego, CA, May 7–9, 2015.

[37] P. Krauss, C. Metzner, A. Schilling, C. Schütz, K. Tziridis, B. Fabry, and H. Schulze, Adaptive stochastic resonance for unknown and variable input signals, *Sci. Rep.*, vol. 7, art. 2450, 8 pp., 2017.

[38] Z. Li, C. A. Micchelli, and Y. Xu, Fixed-point proximity algorithm for minimal norm interpolation, *Appl. Comput. Harm. Anal.*, vol. 49, pp. 328–342, 2020.

[39] A. Marmin, A. Jezierska, M. Castella, and J.-C. Pesquet, Global optimization for recovery of clipped signals corrupted with Poisson-Gaussian noise, *IEEE Signal Process. Lett.*, vol. 27, pp. 970–974, 2020.

[40] J. M. Melenk and G. Zimmermann, Functions with time and frequency gaps, *J. Fourier Anal. Appl.*, vol. 2, pp. 611–614, 1996.

[41] C. A. Micchelli and F. I. Utreras, Smoothing and interpolation in a convex subset of a Hilbert space, *SIAM J. Sci. Statist. Comput.*, vol. 9, pp. 728–746, 1988.

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

[43] W. D. Montgomery, Optical applications of Von Neumann's alternating-projection theorem, *Optics Lett.*, vol. 7, pp. 1–3, 1982.

[44] J. J. Moreau, Fonctions convexes duales et points proximaux dans un espace hilbertien, *C. R. Acad. Sci. Paris*, vol. A255, pp. 2897–2899, 1962.

[45] J. J. Moreau, Les liaisons unilatérales et le principe de Gauss, *C. R. Acad. Sci. Paris*, vol. A256, pp. 871–874, 1963.

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

[47] J. J. Moreau, Quadratic programming in mechanics: Dynamics of one-sided constraints, *SIAM J. Control*, vol. 4, pp. 153–158, 1966.

[48] B. Mulansky and M. Neamtu, Interpolation and approximation from convex sets, *J. Approx. Theory*, vol. 92, pp. 82–100, 1998.

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

[50] G. Pierra, Éclatement de contraintes en parallèle pour la minimisation d'une forme quadratique, *Lecture Notes in Comput. Sci.*, vol. 41. Springer, New York, 1976, pp. 200–218.

[51] L. Rencker, F. Bach, W. Wang, and M. D. Plumbley, Sparse recovery and dictionary learning from nonlinear compressive measurements, *IEEE Trans. Signal Process.*, vol. 67, pp. 5659–5670, 2019.

[52] N. N. Reyes and L. J. D. Vallejo, Global growth of band-limited local approximations, *J. Math. Anal. Appl.*, vol. 400, pp. 418–424, 2013.

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

[54] C. Studer, P. Kuppinger, G. Pope, and H. Bölcskei, Recovery of sparsely corrupted signals, *IEEE Trans. Inform. Theory*, vol. 58, pp. 3115–3130, 2012.

[55] T. Tao and B. Vidakovic, Almost everywhere behavior of general wavelet shrinkage operators, *Appl. Comput. Harmon. Anal.*, vol. 9, pp. 72–82, 2000.

[56] E. Tarr, *Hack Audio.* Routledge, New York, 2019.

[57] V. N. Temlyakov, The best $m$-term approximation and greedy algorithms, *Adv. Comput. Math.*, vol. 8, pp. 249–265, 1998.

[58] T. Teshima, M. Xu, I. Sato, and M. Sugiyama, Clipped matrix completion: A remedy for ceiling effects, *Proc. AAAI Conf. Artif. Intell.*, pp. 5151–5158, 2019.

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

[60] M. Yuan and Y. Lin, Model selection and estimation in regression with grouped variables, *J. R. Stat. Soc. Ser. B Stat. Methodol.*, vol. 68, pp. 49–67, 2006.

Chapter 4

# A VARIATIONAL INEQUALITY MODEL FOR THE CONSTRUCTION OF SIGNALS FROM INCONSISTENT NONLINEAR EQUATIONS

## 4.1 Introduction and context

Due to noise or modeling errors, it may be the case that Problems 1.1.1, 1.1.3, or 1.1.4 have no solution. Furthermore, if no solution exists, then the algorithms for solving these problems in Chapters 2 and 3 are known to diverge. This chapter analyzes the relaxed formulation Problem 1.1.5 and presents an efficient block-iterative algorithm for its solution. It is worth noting that this relaxation captures the one proposed in Chapter 2 as a special case. In addition, further proxification results pertaining to matrix-valued operators are presented.

This chapter presents the following article.

> P. L. Combettes and Z. C. Woodstock, A variational inequality model for the construction of signals from inconsistent nonlinear equations, submitted.

## 4.2 Article: A variational inequality model for the construction of signals from inconsistent nonlinear equations

**Abstract**. Building up on classical linear formulations, we posit that a broad class of problems in signal synthesis and in signal recovery are reducible to the basic task of finding a point in a closed convex subset of a Hilbert space that satisfies a number of nonlinear equations involving firmly nonexpansive operators. We investigate this formalism in the case when, due to inaccurate

modeling or perturbations, the nonlinear equations are inconsistent. A relaxed formulation of the original problem is proposed in the form of a variational inequality. The properties of the relaxed problem are investigated and a provenly convergent block-iterative algorithm, whereby only blocks of the underlying firmly nonexpansive operators are activated at a given iteration, is devised to solve it. Numerical experiments illustrate robust recoveries in several signal and image processing applications.

### 4.2.1 Introduction

Signal construction encompasses forward problems such as image synthesis, holography, filter design, time-frequency distribution synthesis, and radiation therapy planning, as well as inverse problems such as density estimation, signal denoising, image interpolation, signal extrapolation, audio declipping, image reconstruction, or deconvolution; see, e.g., [4,16,19,29,31,32,45,47,48, 51,58]. Essential components in the mathematical modeling of signal construction problems are equations tying the ideal solution $\overline{x}$ in a space $\mathcal{H}$ to given prescriptions in a space $\mathcal{G}$, say $W\overline{x} = p$, where $W$ is an operator mapping $\mathcal{H}$ to $\mathcal{G}$. The prescription $p$ can be a design specification in forward problems, or an observation in inverse problems.

In 1978, Youla [60] elegantly brought to light the simple geometry that underlies many classical problems in signal construction by reducing them to the following formulation: given closed vector subspaces $C$ and $D$ in a real Hilbert space $\mathcal{H}$, and a point $p \in D$,

$$\text{find} \quad x \in C \quad \text{such that} \quad \operatorname{proj}_D x = p, \tag{4.1}$$

where $\operatorname{proj}_D$ denotes the projection operator onto $D$. In the context of signal recovery, the original signal of interest $\overline{x}$ is known to lie in $C$ and some observation $p$ of it is available in the form of its projection onto $D$. A natural nonlinear extension of this setting is obtained by considering nonempty closed convex sets $C$ in $\mathcal{H}$ and $D$ in a real Hilbert space $\mathcal{G}$, a bounded linear operator $L\colon \mathcal{H} \to \mathcal{G}$, a point $p \in D$, and setting as an objective to

$$\text{find} \quad x \in C \quad \text{such that} \quad \operatorname{proj}_D(Lx) = p. \tag{4.2}$$

An early instance of this model appears in [1], where $C$ is a set of bandlimited signals and $p$ is an observation of $N$ clipped samples of the original signal. Thus, $L\colon \mathcal{H} \to \mathbb{R}^N$ is the sampling operator and $D = \big\{y \in \mathbb{R}^N \;\big|\; \|y\|_\infty \leqslant \rho\big\}$ for some $\rho \in \left]0, +\infty\right[$. A key property of projectors onto closed convex sets is their firm nonexpansiveness. Recall that an operator $F\colon \mathcal{G} \to \mathcal{G}$ is firmly nonexpansive if [6]

$$(\forall x \in \mathcal{G})(\forall y \in \mathcal{G}) \quad \langle x - y \mid Fx - Fy\rangle \geqslant \|Fx - Fy\|^2. \tag{4.3}$$

In [26,27], it was shown that many nonlinear observation processes found in signal processing, machine learning, and inference problems can be represented through such operators. This

prompts us to consider the following formulation, whereby the prescriptions are modeled via Wiener systems (see Figure 4.1).

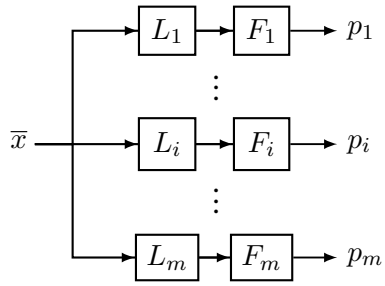



**Figure 4.1** Illustration of Problem 4.2.1 with $m$ prescriptions $(p_i)_{1\leqslant i\leqslant m}$. The $i$th prescription $p_i$ is the output produced when the ideal signal $\overline{x}$ is input to a Wiener system $W_i = F_i \circ L_i$, i.e., the concatenation of a linear system $L_i$ and a nonlinear system $F_i$ [49]. In the proposed model, $F_i$ is a firmly nonexpansive operator.

**Problem 4.2.1** Let $I$ be a nonempty finite set and let $C$ be a nonempty closed convex subset of a real Hilbert space $\mathcal{H}$. For every $i \in I$, let $\mathcal{G}_i$ be a real Hilbert space, let $p_i \in \mathcal{G}_i$, let $L_i \colon \mathcal{H} \to \mathcal{G}_i$ be a nonzero bounded linear operator, and let $F_i \colon \mathcal{G}_i \to \mathcal{G}_i$ be a firmly nonexpansive operator. The task is to

$$\text{find } \; x \in C \; \text{ such that } \; (\forall i \in I) \;\; F_i(L_i x) = p_i. \tag{4.4}$$

The work of [26, 27] assumes that the prescription equations in Problem 4.2.1 are exact and hence that a solution exists. In many instances, however, the prescription operators may be imperfectly known or the model may be corrupted by perturbations, so that Problem 4.2.1 may not have solutions, e.g., [17, 18, 31]. A dramatic consequence of this lack of feasibility is that the algorithms proposed [26, 27] are known to diverge in such situations. To deal robustly with possibly inconsistent equations, one must therefore come up with an appropriate relaxed formulation of Problem 4.2.1, i.e., one that seeks a point in $C$ that satisfies the nonlinear equations in an approximate sense, and coincides with the original problem (4.4) if it happens to be consistent. To guide our design of a relaxed problem, let us consider a classical instantiation of Problem 4.2.1.

**Example 4.2.2** Specialize Problem 4.2.1 by setting, for every $i \in I$,

$$p_i = 0 \;\; \text{and} \;\; F_i = \mathrm{Id} - \mathrm{proj}_{D_i}, \;\; \text{where } D_i \text{ is a nonempty closed convex subset of } \mathcal{G}_i, \tag{4.5}$$

and note that the operators $(F_i)_{i\in I}$ are firmly nonexpansive [6, Corollary 4.18]. In this context,

(4.4) reduces to the convex feasibility problem [15, 19, 62]

$$\text{find } x \in C \text{ such that } (\forall i \in I)\ L_i x \in D_i. \tag{4.6}$$

Let $(\omega_i)_{i\in I}$ be real numbers in $]0,1]$ such that $\sum_{i\in I}\omega_i = 1$ and, for every $i \in I$, let $d_{D_i}$ be the distance function to $D_i$. As seen in [23] (see also [16–18, 22, 32, 61] for special cases), a relaxation of (4.6) when it may be inconsistent is the least-squares problem

$$\underset{x\in C}{\text{minimize}}\ f(x), \quad \text{where} \quad f\colon x \mapsto \frac{1}{2}\sum_{i\in I}\omega_i d^2_{D_i}(L_i x) = \frac{1}{2}\sum_{i\in I}\omega_i \|L_i x - \text{proj}_{D_i}(L_i x)\|^2. \tag{4.7}$$

An important property of this formulation is that $f$ is a smooth convex function since [6, Corollary 12.31] asserts that

$$(\forall i \in I) \quad \nabla\frac{d^2_{D_i}\circ L_i}{2} = L_i^*\circ(\text{Id} - \text{proj}_{D_i})\circ L_i = L_i^*\circ F_i\circ L_i - L_i^* p_i. \tag{4.8}$$

It can therefore be solved by the projection-gradient algorithm [6, Corollary 28.10]. Let us also note that (4.7) is a valid relaxation of (4.6). Indeed, if the latter has solutions, then $f$ vanishes on $C$ at those points only, and (4.7) is therefore equivalent to (4.6). Historically, the first instance of the above relaxation process seems to be Legendre's least-squares methods [37]. There, $\mathcal{H} = \mathbb{R}^N = C$ and, for every $i \in I$, $\mathcal{G}_i = \mathbb{R}$, $D_i = \{\beta_i\}$, and $L_i = \langle\cdot \mid a_i\rangle$, where $\beta_i \in \mathbb{R}$ and $0 \neq a_i \in \mathbb{R}^N$. Set $b = (\beta_i)_{i\in I}$, let $A$ be the matrix with rows $(a_i)_{i\in I}$, and let $(\forall i \in I)$ $\omega_i = 1/\text{card}\, I$. Then (4.6) consists of solving the linear system $Ax = b$ and (4.7) of minimizing the function $x \mapsto \|Ax - b\|^2$.

In general, there is no suitable relaxation of Problem 4.2.1 in the form of a tractable convex minimization problem such as (4.7). For instance, in Example 4.2.2, we can rewrite (4.7) as

$$\underset{x\in C}{\text{minimize}}\ f(x), \quad \text{where} \quad f\colon x \mapsto \frac{1}{2}\sum_{i\in I}\omega_i \|F_i(L_i x) - p_i\|^2. \tag{4.9}$$

However, beyond the special case (4.5), $f$ is typically a nonconvex and nondifferentiable function [4, 43, 64], which makes it impossible to guarantee the construction of solutions. Another plausible formulation that captures (4.7) would be to introduce in Problem 4.2.1 the closed convex sets $(\forall i \in I)$ $D_i = \big\{y \in \mathcal{G}_i \;\big|\; F_i y_i = p_i\big\}$. However the resulting minimization problem (4.7) is intractable because we typically do not know how to evaluate the operators $(\text{proj}_{D_i})_{i\in I}$, and therefore cannot evaluate $f$ and its gradient.

Our strategy to relax (4.4) is to forego the optimization approach in favor of the broader framework of *variational inequalities*. To motivate this approach, let us go back to Example 4.2.2. Then it follows from Lemma 4.2.7 below and (4.8) that (4.7) equivalent to finding $x \in C$ such that $(\forall y \in C)$ $\sum_{i\in I}\omega_i\langle L_i(y-x) \mid F_i(L_i x) - p_i\rangle \geqslant 0$. We shall show that this variational inequality constitutes an appropriate relaxed formulation of Problem 4.2.1 in the presence of

general firmly nonexpansive operators $(F_i)_{i\in I}$, and that it can be solved iteratively through an efficient block-iterative fixed point algorithm. Here is a precise formulation of our relaxed problem.

**Problem 4.2.3** Let $I$ be a nonempty finite set, let $(\omega_i)_{i\in I}$ be real numbers in $]0,1]$ such that $\sum_{i\in I}\omega_i = 1$, and let $C$ be a nonempty closed convex subset of a real Hilbert space $\mathcal{H}$. For every $i \in I$, let $\mathcal{G}_i$ be a real Hilbert space, let $p_i \in \mathcal{G}_i$, let $L_i\colon \mathcal{H} \to \mathcal{G}_i$ be a nonzero bounded linear operator, and let $F_i\colon \mathcal{G}_i \to \mathcal{G}_i$ be a firmly nonexpansive operator. The task is to

$$\text{find } x \in C \text{ such that } (\forall y \in C) \sum_{i\in I}\omega_i\langle L_i(y-x) \mid F_i(L_ix) - p_i\rangle \geqslant 0. \tag{4.10}$$

The paper is organized as follows. Section 4.2.2 provides the notation and the necessary background, as well as preliminary results. It covers in particular the basics of monotone operator theory, which will play an essential role in the paper. In Section 4.3, we illustrate the flexibility and the breadth the proposed firmly nonexpansive Wiener model. In Section 4.3.1, we analyze various properties of Problem 4.2.3, in particular as a relaxation of Problem 4.2.1. We also provide in that section a block-iterative algorithm to solve Problem 4.2.3. Section 4.3.2 is devoted to numerical experiments in the area of signal and image processing.

## 4.2.2 Notation, background, and preliminary results

### 4.2.2.1 Notation

Our notation follows [6], to which one can refer for background on monotone operators and convex analysis. Let $\mathcal{H}$ be a real Hilbert space with scalar product $\langle\cdot \mid \cdot\rangle$, associated norm $\|\cdot\|$, and identity operator Id . The family of all subsets of $\mathcal{H}$ is denoted by $2^{\mathcal{H}}$. The Hilbert direct sum of a family of real Hilbert spaces $(\mathcal{H}_i)_{i\in I}$ is denoted by $\bigoplus_{i\in I}\mathcal{H}_i$.

Let $T\colon \mathcal{H} \to \mathcal{H}$. Then $T$ is *cocoercive* if there exists $\beta \in ]0,+\infty[$ such that

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \langle x-y \mid Tx-Ty\rangle \geqslant \beta\|Tx-Ty\|^2, \tag{4.11}$$

and *firmly nonexpansive* if $\beta = 1$ above. The set of *fixed points* of $T$ is $\operatorname{Fix} T = \big\{x \in \mathcal{H} \;\big|\; Tx = x\big\}$.

Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$. The *graph* of $A$ is $\operatorname{gra} A = \big\{(x,x^*) \in \mathcal{H}\times\mathcal{H} \;\big|\; x^* \in Ax\big\}$, the *domain* of $A$ is $\operatorname{dom} A = \big\{x \in \mathcal{H} \;\big|\; Ax \neq \varnothing\big\}$, the *range* of $A$ is $\operatorname{ran} A = \big\{x^* \in \mathcal{H} \;\big|\; (\exists\, x \in \mathcal{H})\; x^* \in Ax\big\}$, the set of zeros of $A$ is $\operatorname{zer} A = \big\{x \in \mathcal{H} \;\big|\; 0 \in Ax\big\}$, the *inverse* of $A$ is $A^{-1}\colon \mathcal{H} \to 2^{\mathcal{H}}\colon x^* \mapsto \big\{x \in \mathcal{H} \;\big|\; x^* \in Ax\big\}$, and the *resolvent* of $A$ is $J_A = (\operatorname{Id} + A)^{-1}$. Further, $A$ is *monotone* if

$$\big(\forall (x,x^*) \in \operatorname{gra} A\big)\big(\forall (y,y^*) \in \operatorname{gra} A\big) \quad \langle x-y \mid x^*-y^*\rangle \geqslant 0, \tag{4.12}$$

and *maximally monotone* if, for every $(x,x^*) \in \mathcal{H}\times\mathcal{H}$,

$$(x,x^*) \in \operatorname{gra} A \quad \Leftrightarrow \quad \big(\forall (y,y^*) \in \operatorname{gra} A\big)\;\; \langle x-y \mid x^*-y^*\rangle \geqslant 0. \tag{4.13}$$

If $A$ is maximally monotone, then $J_A$ is a single-valued firmly nonexpansive operator defined on $\mathcal{H}$. If $A$ is monotone and satisfies

$$(\forall (x, x^*) \in \operatorname{dom} A \times \operatorname{ran} A) \quad \sup \big\{ \langle x - y \mid y^* - x^* \rangle \;\big|\; (y, y^*) \in \operatorname{gra} A \big\} < +\infty, \tag{4.14}$$

then it is $3^*$ *monotone*.

$\Gamma_0(\mathcal{H})$ is the class of all lower semicontinuous convex functions from $\mathcal{H}$ to $]-\infty, +\infty]$ which are proper in the sense that they are not identically $+\infty$. Let $f \in \Gamma_0(\mathcal{H})$. The *domain* of $f$ is $\operatorname{dom} f = \big\{ x \in \mathcal{H} \;\big|\; f(x) < +\infty \big\}$, the *conjugate* of $f$ is the function

$$\Gamma_0(\mathcal{H}) \ni f^* \colon x^* \mapsto \sup_{x \in \mathcal{H}} \big( \langle x \mid x^* \rangle - f(x) \big), \tag{4.15}$$

and the *subdifferential* of $f$ is the maximally monotone operator

$$\partial f \colon \mathcal{H} \to 2^{\mathcal{H}} \colon x \mapsto \big\{ x^* \in \mathcal{H} \;\big|\; (\forall y \in \mathcal{H})\ \langle y - x \mid x^* \rangle + f(x) \leqslant f(y) \big\}. \tag{4.16}$$

The *Moreau envelope* of $f$ is

$$\widetilde{f} \colon \mathcal{H} \to \mathbb{R} \colon x \mapsto \inf_{y \in \mathcal{H}} \left( f(y) + \frac{\|x - y\|^2}{2} \right). \tag{4.17}$$

For every $x \in \mathcal{H}$, the infimum in (4.17) is achieved at a unique point, which is denoted by $\operatorname{prox}_f x$. This defines the *proximity operator* $\operatorname{prox}_f = J_{\partial f}$ of $f$.

Let $C$ be a nonempty closed and convex subset of $\mathcal{H}$. The *distance* from $x \in \mathcal{H}$ to $C$ is $d_C(x) = \inf_{y \in C} \|x - y\|$, the *indicator function* of $C$ is

$$\iota_C \colon \mathcal{H} \to ]-\infty, +\infty] : x \mapsto \begin{cases} 0, & \text{if } \ x \in C; \\ +\infty, & \text{if } \ x \notin C, \end{cases} \tag{4.18}$$

the *normal cone* to $C$ at $x \in \mathcal{H}$ is

$$N_C x = \partial \iota_C(x) = \begin{cases} \big\{ x^* \in \mathcal{H} \;\big|\; (\forall y \in C)\ \ \langle y - x \mid x^* \rangle \leqslant 0 \big\}, & \text{if } \ x \in C; \\ \varnothing, & \text{otherwise,} \end{cases} \tag{4.19}$$

and the *projection operator* onto $C$ is $\operatorname{proj}_C = \operatorname{prox}_{\iota_C} = J_{N_C}$.

The following facts will also come into play.

**Lemma 4.2.4** *Let $A \colon \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone, let $\mu \in ]0, +\infty[$, and let $\gamma \in ]0, 1/\mu[$. Set $B = A - \mu \operatorname{Id}$ and $\beta = 1 - \gamma\mu$. Then $J_{\gamma B} \colon \mathcal{H} \to \mathcal{H}$ is $\beta$-cocoercive. Furthermore, $J_{\gamma B} = J_{\beta^{-1}\gamma A} \circ (\beta^{-1} \operatorname{Id})$.*

*Proof*. Let $x$ and $q$ be in $\mathcal{H}$. Since $\beta^{-1}\gamma A$ is maximally monotone, its resolvent is single-valued

with domain $\mathcal{H}$. Therefore,

$$\begin{aligned} q \in J_{\gamma B}x &\Leftrightarrow x - q \in \gamma Bq \\ &\Leftrightarrow x - \beta q \in \gamma Aq \\ &\Leftrightarrow \beta^{-1}x - q \in \beta^{-1}\gamma Aq \\ &\Leftrightarrow q = J_{\beta^{-1}\gamma A}\big(\beta^{-1}x\big), \end{aligned} \tag{4.20}$$

which shows that $J_{\gamma B} = J_{\beta^{-1}\gamma A} \circ (\beta^{-1}\,\mathrm{Id}\,)$ is single-valued with domain $\mathcal{H}$. Finally, since $M = \beta\gamma A$ is maximally monotone, it follows from [6, Corollary 23.26] that $J_{\gamma B} = J_{\beta^{-2}M} \circ (\beta^{-1}\,\mathrm{Id}\,)$ is $\beta$-cocoercive. □

**Lemma 4.2.5 ([6, Proposition 24.68])** *Let $\mathcal{H}$ be the real Hilbert space of $N \times M$ matrices under the Frobenius norm, and set $s = \min\{N, M\}$. Denote the singular value decomposition of $x \in \mathcal{H}$ by $x = U_x \operatorname{diag}(\sigma_1(x), \ldots, \sigma_s(x))V_x^\top$. Let $\phi \in \Gamma_0(\mathbb{R})$ be even, and set*

$$F\colon \mathcal{H} \to \mathcal{H}\colon x \mapsto U_x \operatorname{diag}\Big(\operatorname{prox}_\phi\big(\sigma_1(x)\big), \ldots, \operatorname{prox}_\phi\big(\sigma_s(x)\big)\Big)V_x^\top. \tag{4.21}$$

*Then $F$ is firmly nonexpansive.*

#### 4.2.2.2 Variational inequalities

The following notion of a *variational inequality* was formulated in [12] (see Figure 4.2).

**Definition 4.2.6** Let $C$ be a nonempty closed convex set of $\mathcal{H}$ and let $B\colon \mathcal{H} \to \mathcal{H}$ be a monotone operator. The associated *variational inequality* problem is to

$$\text{find}\ \ x \in C\ \ \text{such that}\ \ (\forall y \in C)\ \ \langle y - x \mid Bx\rangle \geqslant 0. \tag{4.22}$$

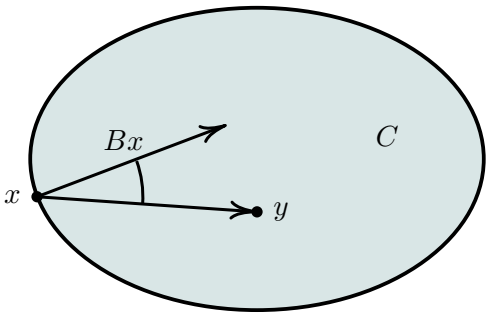



**Figure 4.2** Illustration of the variational inequality principle. The point $x$ solves (4.22) because it lies in $C$ and, for every $y \in C$, the angle between $y - x$ and $Bx$ is acute.

Variational inequalities are used in various areas of mathematics and its applications [8, 30, 35, 65]. They are also central in constrained minimization problems.

**Lemma 4.2.7** [6, Proposition 27.8] *Let $f\colon \mathcal{H} \to \mathbb{R}$ be a differentiable convex function, let $C$ be a nonempty closed convex subset of $\mathcal{H}$, and let $x \in \mathcal{H}$. Then $x$ minimizes $f$ over $C$ if and only if it satisfies the variational inequality*

$$x \in C \;\text{ and }\; (\forall y \in C)\;\; \langle y - x \mid \nabla f(x)\rangle \geqslant 0. \tag{4.23}$$

### 4.2.3 Composite sums of monotone operators

We shall require the following Brézis–Haraux-type theorem, which remains valid in general reflexive Banach spaces (see [10, Théorème 3] for the special case of the sum of two monotone operators).

**Lemma 4.2.8** *Let $\mathcal{H}$ be a real Hilbert space and let $(\mathcal{G}_i)_{i\in I}$ be a finite family of real Hilbert spaces. Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a $3^*$ monotone operator and, for every $i \in I$, let $B_i\colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ be a $3^*$ monotone operator and let $L_i\colon \mathcal{H} \to \mathcal{G}_i$ be a bounded linear operator. Suppose that $A + \sum_{i\in I} L_i^* \circ B_i \circ L_i$ is maximally monotone. Then*

$$\begin{cases} \operatorname{int}\big(\operatorname{ran} A + \sum_{i\in I} L_i^*(\operatorname{ran} B_i)\big) = \operatorname{int}\operatorname{ran}\big(A + \sum_{i\in I} L_i^* \circ B_i \circ L_i\big) \\ \overline{\operatorname{ran} A + \sum_{i\in I} L_i^*(\operatorname{ran} B_i)} = \overline{\operatorname{ran}}\,\big(A + \sum_{i\in I} L_i^* \circ B_i \circ L_i\big). \end{cases} \tag{4.24}$$

*Proof*. Clearly, $\operatorname{ran}(A + \sum_{i\in I} L_i^* \circ B_i \circ L_i) \subset (\operatorname{ran} A + \sum_{i\in I} L_i^*(\operatorname{ran} B_i))$. It is therefore enough to show that

$$\begin{cases} \operatorname{int}\big(\operatorname{ran} A + \sum_{i\in I} L_i^*(\operatorname{ran} B_i)\big) \subset \operatorname{ran}\big(A + \sum_{i\in I} L_i^* \circ B_i \circ L_i\big) \\ \operatorname{ran} A + \sum_{i\in I} L_i^*(\operatorname{ran} B_i) \subset \overline{\operatorname{ran}}\,\big(A + \sum_{i\in I} L_i^* \circ B_i \circ L_i\big). \end{cases} \tag{4.25}$$

Without loss of generality, set $I = \{1, \ldots, m\}$ and introduce the Hilbert direct sum $\boldsymbol{\mathcal{H}} = \mathcal{H} \oplus \mathcal{G}_1 \oplus \cdots \oplus \mathcal{G}_m$. Furthermore, introduce the bounded linear operator $\boldsymbol{L}\colon \mathcal{H} \to \boldsymbol{\mathcal{H}}\colon x \mapsto (x, L_1x, \ldots, L_mx)$ and the operator $\boldsymbol{M}\colon \boldsymbol{\mathcal{H}} \to 2^{\boldsymbol{\mathcal{H}}}\colon (x, y_1, \ldots, y_m) \mapsto Ax \times B_1y_1 \times \cdots \times B_my_m$, which is $3^*$ monotone since $A$ and $(B_i)_{i\in I}$ are. Note also that, since $\boldsymbol{L}^*\colon \boldsymbol{\mathcal{H}} \to \mathcal{H}\colon (x, y_1, \ldots, y_m) \mapsto x + \sum_{i\in I} L_i^* y_i$, the operator

$$\boldsymbol{L}^* \circ \boldsymbol{M} \circ \boldsymbol{L} = A + \sum_{i\in I} L_i^* \circ B_i \circ L_i \tag{4.26}$$

is maximally monotone. We can therefore apply [42, Theorem 5] to obtain

$$\begin{cases} \operatorname{int}\boldsymbol{L}^*(\operatorname{ran}\boldsymbol{M}) \subset \operatorname{ran}(\boldsymbol{L}^* \circ \boldsymbol{M} \circ \boldsymbol{L}) \\ \boldsymbol{L}^*(\operatorname{ran}\boldsymbol{M}) \subset \overline{\operatorname{ran}}\,(\boldsymbol{L}^* \circ \boldsymbol{M} \circ \boldsymbol{L}), \end{cases} \tag{4.27}$$

which is precisely (4.25). □

We consider below a monotone inclusion problem involving several operators.

**Problem 4.2.9** Let $(\omega_i)_{i\in I}$ be a finite family of real numbers in $]0,1]$ such that $\sum_{i\in I}\omega_i = 1$, let $A_0\colon \mathcal{H}\to 2^{\mathcal{H}}$ be maximally monotone and, for every $i\in I$, let $\beta_i\in\,]0,+\infty[$ and let $A_i\colon\mathcal{H}\to\mathcal{H}$ be $\beta_i$-cocoercive. The task is to find $x\in\mathcal{H}$ such that $0\in A_0x+\sum_{i\in I}\omega_iA_ix$.

**Proposition 4.2.10** [24, Proposition 4.9] *Consider the setting of Problem 4.2.9 under the assumption that it has a solution. Let $K$ be a strictly positive integer and let $(I_n)_{n\in\mathbb{N}}$ be a sequence of nonempty subsets of $I$ such that $(\forall n\in\mathbb{N})\ \bigcup_{k=0}^{K-1}I_{n+k}=I$. Let $\gamma\in\,]0,2\min_{1\leqslant i\leqslant m}\beta_i[$, let $x_0\in\mathcal{H}$, and let $(\forall i\in I)\ t_{i,-1}\in\mathcal{H}$. Iterate*

$$\begin{array}{l}\text{for } n=0,1,\ldots\\ \quad\text{for every } i\in I_n\\ \quad\quad t_{i,n}=x_n-\gamma A_ix_n\\ \quad\text{for every } i\in I\smallsetminus I_n\\ \quad\quad t_{i,n}=t_{i,n-1}\\ \quad x_{n+1}=J_{\gamma A_0}\Big(\sum_{i\in I}\omega_it_{i,n}\Big).\end{array}\tag{4.28}$$

*Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a solution to Problem 4.2.9.*

## 4.3 Firmly nonexpansive Wiener models

The proposed Wiener model (see Figure 4.1) involves a linear operator followed by a firmly nonexpansive operator acting on a real Hilbert space $\mathcal{G}$. Typical examples of linear transformations in the context of signal construction include the Fourier transform, the Radon transform, wavelet decompositions, frame decompositions, audio effects, or blurring operators. We show that firmly nonexpansive operators model many useful nonlinearities in this context. Key examples based on those of [27] are recalled and new ones are proposed. Following [27], we call $p\in\mathcal{G}$ a *proximal point* of $y\in\mathcal{G}$ relative to a firmly nonexpansive operator $F\colon\mathcal{G}\to\mathcal{G}$ if $Fy=p$.

#### 4.3.0.1 Projection operators

As seen in Section 4.2.2.1, the projection operator onto a nonempty closed convex set is firmly nonexpansive.

**Example 4.3.1** For every $j\in\{1,\ldots,m\}$, let $\mathsf{G}_j$ be a real Hilbert space and let $\mathsf{D}_j\subset\mathsf{G}_j$ be nonempty closed and convex. Suppose that $\mathcal{G}=\bigoplus_{1\leqslant j\leqslant m}\mathsf{G}_j$. The operator

$$F\colon(\mathsf{y}_j)_{1\leqslant j\leqslant m}\mapsto(\mathrm{proj}_{\mathsf{D}_j}\mathsf{y}_j)_{1\leqslant j\leqslant m},\tag{4.29}$$

which is also the projection onto the closed convex set $D = \bigtimes_{1 \leqslant j \leqslant m} \mathsf{D}_j$, is the hard clipper of [27, Example 2.11]. If we specialize to the case when, for every $j \in \{1, \ldots, m\}$, $\mathsf{G}_j = \mathbb{R}$, we obtain the standard hard clipping operators of [1, 31, 55].

**Example 4.3.2** Let $K \subset \mathcal{G}$ be a nonempty closed convex cone. The operator $F = \mathrm{proj}_K$ is used as a distortion model when $K$ is the positive orthant [53, Section 10.4.1]. Another instance of a conic projection operator arises in isotonic regression [5], where $K = \big\{(\xi_i)_{1 \leqslant i \leqslant N} \in \mathbb{R}^N \;\big|\; \xi_1 \leqslant \cdots \leqslant \xi_N\big\}$.

**Example 4.3.3** Compression schemes such as downsampling project a high-dimensional object of interest onto a closed convex subset of a low-dimensional subspace of $\mathcal{G}$ [41].

#### 4.3.0.2 Proximity operators

As seen in Section 4.2.2.1, the proximity operator of a function in $\Gamma_0(\mathcal{G})$ is firmly nonexpansive. The following construction will be particularly useful.

**Example 4.3.4** For every $j \in \{1, \ldots, m\}$, let $\mathsf{G}_j$ be a real Hilbert space and let $\mathsf{g}_j \in \Gamma_0(\mathsf{G}_j)$. Suppose that $\mathcal{G} = \bigoplus_{1 \leqslant j \leqslant m} \mathsf{G}_j$ and set $F \colon \mathcal{G} \to \mathcal{G} \colon (\mathsf{y}_j)_{1 \leqslant j \leqslant m} \mapsto (\mathrm{prox}_{\mathsf{g}_j} \mathsf{y}_j)_{1 \leqslant j \leqslant m}$. Then [6, Proposition 24.11] asserts that

$$F = \mathrm{prox}_g, \quad \text{where} \quad g \colon \mathcal{G} \to \left]-\infty, +\infty\right] : (\mathsf{y}_j)_{1 \leqslant j \leqslant m} \mapsto \sum_{j=1}^{m} \mathsf{g}_j(\mathsf{y}_j). \tag{4.30}$$

**Example 4.3.5** In Example 4.3.4 suppose that, for every $j \in \{1, \ldots, m\}$, $\mathsf{g}_j = \phi_j \circ \|\cdot\|$, where $\phi_j$ is an even function in $\Gamma_0(\mathbb{R})$ such that $\phi_j(0) = 0$ and $\phi_j \neq \iota_{\{0\}}$. Set $(\forall j \in \{1, \ldots, m\})$ $\rho_j = \max \partial \phi_j(0)$. Then we derive from [11, Proposition 2.1] that

$$F \colon \mathcal{G} \to \mathcal{G} \colon (\mathsf{y}_j)_{1 \leqslant j \leqslant m} \mapsto \Big( \big(\mathrm{prox}_{\phi_j} \|\mathsf{y}_j\|\big) \lfloor \mathsf{y}_j \rfloor_{\rho_j} \Big)_{1 \leqslant j \leqslant m},$$
$$\text{where} \quad \lfloor \mathsf{y}_j \rfloor_{\rho_j} = \begin{cases} \mathsf{y}_j / \|\mathsf{y}_j\|, & \text{if } \|\mathsf{y}_j\| > \rho_j; \\ 0, & \text{if } \|\mathsf{y}_j\| \leqslant \rho_j. \end{cases} \tag{4.31}$$

**Example 4.3.6** Consider the special case of Example 4.3.5 in which, for some $j \in \{1, \ldots, m\}$, $\phi_j$ is not differentiable at the origin, which implies that $\rho_j > 0$. Then $\mathrm{prox}_{\mathsf{g}_j}$ acts as a thresholder with respect to the $j$th variable in the sense that, if $\|\mathsf{y}_j\| \leqslant \rho_j$, then the $j$th coordinate of $Fy$ is zero. For instance, suppose that, for every $j \in \{1, \ldots, m\}$, $\phi_j = \rho_j |\cdot|$, hence $\partial \phi_j(0) = [-\rho_j, \rho_j]$ and $\mathsf{g}_j = \rho_j \|\cdot\|$. Then $Fy = p$ is acquired though the group-shrinkage operation [63]

$$p = \left( \left( 1 - \frac{\rho_j}{\max\{\|\mathsf{y}_j\|, \rho_j\}} \right) \mathsf{y}_j \right)_{1 \leqslant j \leqslant m}. \tag{4.32}$$

**Example 4.3.7** In contrast to the hard clipping operations of Example 4.3.1, soft clipping operators are not projection operators in general, but many turn out to be proximity operators [27] (see Figure 4.3). For instance, consider the setting of Example 4.3.5 with

$$(\forall j \in \{1, \dots, m\}) \quad \phi_j \colon \eta \mapsto \begin{cases} -|\eta| - \ln(1 - |\eta|) - \dfrac{\eta^2}{2}, & \text{if } |\eta| < 1; \\ +\infty, & \text{if } |\eta| \geqslant 1. \end{cases} \tag{4.33}$$

Then we obtain the soft clipping operator

$$F \colon (\mathsf{y}_j)_{1 \leqslant j \leqslant m} \mapsto \left( \frac{\mathsf{y}_j}{1 + \|\mathsf{y}_j\|} \right)_{1 \leqslant j \leqslant m} \tag{4.34}$$

used in [39]. Soft clipping operators model sensors in signal processing [4, 39, 53] and activation functions in neural networks [25].

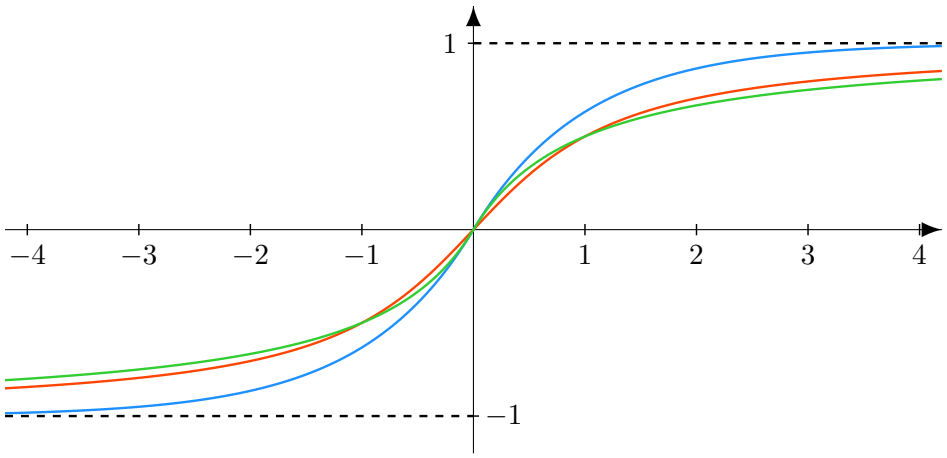



**Figure 4.3** Proximal soft clipping operators on $\mathbb{R}$ with saturation at $\pm 1$: $\eta \mapsto \operatorname{sign}(\eta)(1 - \exp(-|\eta|))$ [53, Section 10.6.3] (blue), $\eta \mapsto 2 \arctan(\eta)/\pi$ [25] (red), and $\eta \mapsto \eta/(1 + |\eta|)$ [39] (green).

#### 4.3.0.3 General firmly nonexpansive operators

Not all firmly nonexpansive operators are proximity operators [21].

**Example 4.3.8** Let $(R_j)_{1 \leqslant j \leqslant m}$ be nonexpansive operators on $\mathcal{G}$. Then the operator

$$F = \frac{\mathrm{Id} + R_1 \circ \cdots \circ R_m}{2} \tag{4.35}$$

is firmly nonexpansive [6, Proposition 4.4] but it is not a proximity operator [21, Example 3.5]. A concrete instance of (4.35) is found in audio signal processing. Consider a distortion $p \in \mathcal{G}$ of

a linearly degraded audio signal $L\overline{x} \in \mathcal{G}$ modeled by

$$F(L\overline{x}) = p, \tag{4.36}$$

where $L$ produces effects such as echo or reverberation [53, Chapter 11], and $F$ comprises several simpler operations $(R_j)_{1\leqslant j\leqslant m}$ which are actually firmly nonexpansive (see, e.g., Example 4.3.2, [27], and [53, Section 10.6.2]). These simpler distortion operators are then used in series and blended with a proportion of the input signal [53, Section 10.9], so that the overall process is described by (4.35) (see Figure 4.4). More generally $F$ remains firmly nonexpansive when $R_1 \circ \cdots \circ R_m$ is replaced by any nonexpansive operator.

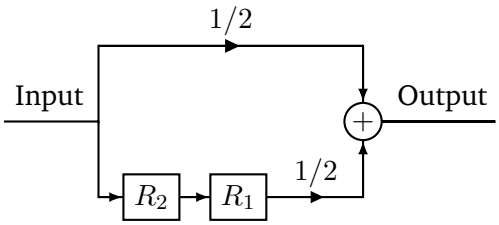



**Figure 4.4** The distortion operator $F$ in Example 4.3.8 for $m = 2$.

#### 4.3.0.4 Proximation

In some instances, a prescription $q \in \mathcal{G}$ may be given by an equation of the form $Qy = q$, where $Q\colon \mathcal{G} \to \mathcal{G}$ is not firmly nonexpansive. In this section, we provide constructive examples of *proximation*, by which we mean the replacement of the equality $Qy = q$ with an equivalent equality $Fy = p$, where $p \in \mathcal{G}$ and $F\colon \mathcal{G} \to \mathcal{G}$ is firmly nonexpansive.

**Definition 4.3.9** Let $Q\colon \mathcal{G} \to \mathcal{G}$ and let $q \in \operatorname{ran} Q$. Then $(Q, q)$ is *proximiable* if there exists a firmly nonexpansive operator $F\colon \mathcal{G} \to \mathcal{G}$ and $p \in \operatorname{ran} F$ such that $(\forall y \in \mathcal{G})\ Qy = q \Leftrightarrow Fy = p$. In this case $(F, p)$ is a *proximation* of $(Q, q)$.

We begin with a necessary condition describing when this technique is possible.

**Proposition 4.3.10** *Let $Q\colon \mathcal{G} \to \mathcal{G}$ and $q \in \operatorname{ran} Q$ be such that $(Q, q)$ is proximiable. Then*

$$Q^{-1}(\{q\}) = \big\{y \in \mathcal{G} \;\big|\; Qy = q\big\} \text{ is closed and convex.} \tag{4.37}$$

*Proof.* The proximation assumption means that there exists a firmly nonexpansive operator $F\colon \mathcal{G} \to \mathcal{G}$ and $p \in \operatorname{ran} F$ such that $Q^{-1}(\{q\}) = F^{-1}(\{p\})$. Now set $T = \operatorname{Id} - F + p$. Then it follows from [6, Proposition 4.4] that $T$ is firmly nonexpansive, and therefore from [6, Corollary 4.24] that $Q^{-1}(\{q\}) = F^{-1}(\{p\}) = \operatorname{Fix} T$ is closed and convex. □

Interestingly, condition (4.37) is also assumed in various nonlinear recovery problems [45, 46, 56]. However, the solution techniques of these papers require the ability to project onto $Q^{-1}(\{q\})$ – a capability which rarely occurs when $\dim \mathcal{G} > 1$. The numerical approach proposed in Section 4.3.1 will circumvent this requirement and lead to provenly-convergent algorithms which instead rely on evaluating the associated firmly nonexpansive operator $F\colon \mathcal{G} \to \mathcal{G}$.

**Example 4.3.11 ([27, Proposition 2.14])** For every $j \in \{1, \dots, m\}$, let $\mathsf{G}_j$ be a real Hilbert space, let $\mathsf{D}_j$ be a nonempty closed convex subset of $\mathsf{G}_j$, let $\gamma_j \in ]0, +\infty[$, and set

$$\mathsf{Q}_j\colon \mathsf{G}_j \to \mathsf{G}_j\colon \mathsf{y}_j \mapsto \begin{cases} \mathsf{y}_j, & \text{if } \ d_{\mathsf{D}_j}(\mathsf{y}_j) > \gamma_j; \\ \operatorname{proj}_{\mathsf{D}_j} \mathsf{y}_j, & \text{if } \ d_{\mathsf{D}_j}(\mathsf{y}_j) \leqslant \gamma_j \end{cases} \tag{4.38}$$

and

$$\mathsf{S}_j\colon \mathsf{G}_j \to \mathsf{G}_j\colon \mathsf{y}_j \mapsto \begin{cases} \mathsf{y}_j + \dfrac{\gamma_j}{d_{\mathsf{D}_j}(\mathsf{y}_j)}(\operatorname{proj}_{\mathsf{D}_j} \mathsf{y}_j - \mathsf{y}_j), & \text{if } \ \mathsf{y}_j \notin \mathsf{D}_j; \\ \mathsf{y}_j, & \text{if } \ \mathsf{y}_j \in \mathsf{D}_j. \end{cases} \tag{4.39}$$

Suppose that $\mathcal{G} = \bigoplus_{1 \leqslant j \leqslant m} \mathsf{G}_j$, set $Q\colon \mathcal{G} \to \mathcal{G}\colon (\mathsf{y}_j)_{1 \leqslant j \leqslant m} \mapsto (\mathsf{Q}_j \mathsf{y}_j)_{1 \leqslant j \leqslant m}$, and let $q \in \operatorname{ran} Q$. Even though $Q$ is discontinuous, $(Q, q)$ is proximiable. Indeed, set $S\colon \mathcal{G} \to \mathcal{G}\colon (\mathsf{y}_j)_{1 \leqslant j \leqslant m} \mapsto (\mathsf{S}_j \mathsf{y}_j)_{1 \leqslant j \leqslant m}$, $F\colon \mathcal{G} \to \mathcal{G}\colon (\mathsf{y}_j)_{1 \leqslant j \leqslant m} \mapsto (\mathsf{S}_j(\mathsf{Q}_j \mathsf{y}_j))_{1 \leqslant j \leqslant m}$, and $p = Sq$. Then $(F, p)$ is a proximation of $(Q, q)$. In particular if, for every $j \in \{1, \dots, m\}$, $\mathsf{D}_j = \{0\}$, then $Q$ is the block thresholding estimation operator of [34, Section 2.3].

**Example 4.3.12** Consider Example 4.3.11 with, for every $j \in \{1, \dots, m\}$, $\mathsf{G}_j = \mathbb{R}$, $\mathsf{D}_j = \{0\}$, and $\gamma_j = \gamma \in ]0, +\infty[$. Then each operator $\mathsf{Q}_j$ in (4.38) reduces to the hard thresholder

$$\operatorname{hard}_\gamma : \eta \mapsto \begin{cases} \eta, & \text{if } \ |\eta| > \gamma; \\ 0, & \text{if } \ |\eta| \leqslant \gamma, \end{cases} \tag{4.40}$$

$\mathsf{S}_j\colon \eta \mapsto \eta - \gamma \operatorname{sign}(\eta)$, and

$$\mathsf{S}_j \circ \operatorname{hard}_\gamma \ = \ \operatorname{soft}_\gamma : \eta \mapsto \operatorname{sign}(\eta) \max\{|\eta| - \gamma, 0\} \tag{4.41}$$

is the soft thresholder on $[-\gamma, \gamma]$. Furthermore, it follows from Example 4.3.11 that $(F, p)$ is a proximation of $(Q, q)$. The resulting transformation $Q$ is used for signal compression in [28, 54], and as a sensing model in [9].

Next, we combine Example 4.3.12 with Lemma 4.2.5 to address low rank matrix approximation. Note the properties of $\phi$ in Lemma 4.2.5 imply that $\operatorname{prox}_\phi 0 = 0$. Therefore, operators of the form (4.21) cannot increase the rank of a matrix.

**Example 4.3.13** Let $\mathcal{G}$ be the real Hilbert space of $N \times M$ matrices under the Frobenius norm, set $s = \min\{N, M\}$, and let us denote the singular value decomposition of $y \in \mathcal{G}$ by

$y = U_y \operatorname{diag}(\sigma_1(y), \dots, \sigma_s(y)) V_y^\top$. Let $\rho \in ]0, +\infty[$, let $\operatorname{hard}_\rho$ be given by (4.40), set $\mathsf{S}\colon \mathbb{R} \to \mathbb{R}\colon \eta \mapsto \eta - \rho \operatorname{sign}(\eta)$, and set

$$\begin{cases} Q\colon \mathcal{G} \to \mathcal{G}\colon y \mapsto U_y \operatorname{diag}\Big(\operatorname{hard}_\rho(\sigma_1(y)), \dots, \operatorname{hard}_\rho(\sigma_s(y))\Big) V_y^\top \\ S\colon \mathcal{G} \to \mathcal{G}\colon y \mapsto U_y \operatorname{diag}\Big(\mathsf{S}\big(\sigma_1(y)\big), \dots, \mathsf{S}\big(\sigma_s(y)\big)\Big) V_y^\top. \end{cases} \tag{4.42}$$

Let $q \in \operatorname{ran} Q$, and set $F = S \circ Q$ and $p = Sq$. Since $\operatorname{soft}_\rho = \operatorname{prox}_{\rho|\cdot|}$ and $\rho|\cdot|$ is even, it follows from Example 4.3.12 and Lemma 4.2.5 that $(F, p)$ is a proxification of $(Q, q)$. The operator $Q$ is used in image compression to produce low rank approximations [3, 36, 44, 59], and the associated firmly nonexpansive operator $F$ soft-thresholds singular values at level $\rho$.

**Remark 4.3.14** In the setting of Example 4.3.13, consider the compression technique performed by the nonconvex projection operator $R\colon \mathcal{G} \to \mathcal{G}$ [13] which truncates singular values at a given rank $r \in \{1, \dots, s-1\}$, i.e., $R\colon y \mapsto U_y \operatorname{diag}\big(\sigma_1(y), \dots, \sigma_r(y), 0, \dots, 0\big) V_y^\top$. Let $y \in \mathcal{G}$ and set $q = Ry$. Then, for every $\rho \in ]\sigma_{r+1}(y), \sigma_r(y)[$, $Qy = q$. Therefore, knowledge of the low rank approximation $q$ to $y$ can be exploited in our framework by proxifying $(Q, q)$ using Example 4.3.13. Note that $\rho$ can be estimated from $q$ since one has access to $\sigma_r(q) = \sigma_r(y)$.

Our last example illustrates how proxification can be used to handle a prescription arising from an extension of the notion of a proximity operator for nonconvex functions.

**Example 4.3.15** Let $\mu \in ]0, +\infty[$, let $\gamma \in ]0, 1/\mu[$, set $\beta = 1 - \gamma\mu$, and let $g\colon \mathcal{G} \to ]-\infty, +\infty]$ be proper, lower semicontinuous, and $\mu$-*weakly convex* in the sense that $g + \mu\|\cdot\|^2/2$ is convex. For every $y \in \mathcal{G}$, $g + \|y - \cdot\|^2/(2\gamma)$ is a strongly convex function in $\Gamma_0(\mathcal{G})$ and, by [6, Corollary 11.17], it therefore admits a unique minimizer $Q_{\gamma g} y$, which defines the operator $Q_{\gamma g}\colon \mathcal{G} \to \mathcal{G}$. Now let $q \in \operatorname{ran} Q_{\gamma g}$ and set $A = \partial(g + \mu\|\cdot\|^2/2)$, $B = A - \mu \operatorname{Id}$, $F = \beta Q_{\gamma g}$, and $p = \beta q$. Then $A$ is maximally monotone but in general, since $g$ is not convex, $Q_{\gamma g}$ is not firmly nonexpansive. However,

$$\begin{aligned} (\forall (y, p) \in \mathcal{G} \times \mathcal{G}) \quad Q_{\gamma g} y = p &\Leftrightarrow p \in \operatorname{zer}\Big(\partial\Big(\gamma g + \frac{\gamma\mu}{2}\|\cdot\|^2 - \frac{\gamma\mu}{2}\|\cdot\|^2 + \frac{1}{2}\|y - \cdot\|^2\Big)\Big) \\ &\Leftrightarrow p \in \operatorname{zer}(\gamma A + \beta \operatorname{Id} - y) = \operatorname{zer}(\operatorname{Id} + \gamma B - y) \\ &\Leftrightarrow J_{\gamma B} y = p, \end{aligned} \tag{4.43}$$

so Lemma 4.2.4 implies that $Q_{\gamma g} = J_{\gamma B}$ is $\beta$-cocoercive. Thus, $(F, p)$ is a proxification of $(Q_{\gamma g}, q)$. Operators of the form $Q_{\gamma g}$ are used for shrinkage in [7, 38, 50] in the same spirit as in Example 4.3.6. For instance, for $\mathcal{G} = \mathbb{R}$ and $\rho \in ]0, +\infty[$, the penalty $g = \ln(\rho + |\cdot|)$ of [38, 50]

is $\rho^{-2}$-weakly convex and yields

$$Q_{\gamma g}\colon y \mapsto \begin{cases} \frac{1}{2}\big(y-\rho+\sqrt{|y+\rho|^2-4\gamma}\,\big), & \text{if } \ y > \frac{\gamma}{\rho}; \\ 0, & \text{if } \ |y| \leqslant \frac{\gamma}{\rho}; \\ \frac{1}{2}\big(y+\rho-\sqrt{|y-\rho|^2-4\gamma}\,\big), & \text{if } \ y < -\frac{\gamma}{\rho}. \end{cases} \tag{4.44}$$

#### 4.3.0.5 Operators arising from monotone equilibria

The property that the object of interest is a zero of the sum of two monotone operators can be modeled in our framework as follows.

**Example 4.3.16** Let $A\colon \mathcal{G} \to 2^{\mathcal{G}}$ be maximally monotone, let $\beta \in \left]0,+\infty\right[$, and let $B\colon \mathcal{G} \to \mathcal{G}$ be $\beta$-cocoercive. Let $\gamma \in \left]0, 2\beta\right[$ and set

$$F = \left(1-\frac{\gamma}{4\beta}\right)\big(\operatorname{Id} - J_{\gamma A} \circ (\operatorname{Id} - \gamma B)\big) \quad \text{and} \quad p = 0. \tag{4.45}$$

Then $F$ is firmly nonexpansive and, for every $y \in \mathcal{G}$, $Fy = p \Leftrightarrow y \in \operatorname{zer}(A+B)$. Indeed, set $R = J_{\gamma A} \circ (\operatorname{Id} - \gamma B)$. By [6, Proposition 26.1(iv)], $R$ is $(2-\gamma/2\beta)^{-1}$-averaged and $\operatorname{zer} F = \operatorname{Fix} R = \operatorname{zer}(A+B)$. It follows from [6, Proposition 4.39] that $\operatorname{Id} - R$ is $(1-\gamma/(4\beta))$-cocoercive, which makes $F$ firmly nonexpansive.

**Example 4.3.17** Let $f \in \Gamma_0(\mathcal{G})$, let $\beta \in \left]0,+\infty\right[$, and let $g\colon \mathcal{G} \to \mathbb{R}$ be a convex and differentiable function such that $\nabla g$ is $\beta^{-1}$-Lipschitzian. Consider the task of enforcing the property

$$y \in \operatorname{Argmin}(f+g). \tag{4.46}$$

Set $A = \partial f$ and $B = \nabla g$. Then $B$ is $\beta$-cocoercive [6, Corollary 18.17], and (4.46) holds if and only if $y \in \operatorname{zer}(A+B)$. Therefore, Example 4.3.16 yields a proximal point representation $(F,p)$ of (4.46).

### 4.3.1 Analysis and numerical solution of Problem 4.2.3

We first show that Problem 4.2.3 is an appropriate relaxation of Problem 4.2.1.

**Proposition 4.3.18** *Suppose that the set of solutions to Problem 4.2.1 is nonempty. Then it coincides with that of solutions to Problem 4.2.3.*

*Proof.* Let $\overline{x}$ be a solution to Problem 4.2.1. Then it is clear that $\overline{x}$ solves Problem 4.2.3. Now let $x$ be a solution to Problem 4.2.3. Then $x \in C$ and

$$(\forall y \in C) \quad \sum_{i\in I} \omega_i \langle L_i(x-y) \mid F_i(L_i x) - p_i\rangle \leqslant 0. \tag{4.47}$$

Therefore, since $\overline{x} \in C$ and, for every $i \in I$, $F_i(L_i\overline{x}) = p_i$, we obtain

$$\sum_{i\in I} \omega_i \langle L_i x - L_i\overline{x} \mid F_i(L_i x) - F_i(L_i\overline{x})\rangle \leqslant 0 \tag{4.48}$$

and, by firm nonexpansiveness of the operators $(F_i)_{i\in I}$,

$$\sum_{i\in I} \omega_i \|F_i(L_i x) - F_i(L_i\overline{x})\|^2 \leqslant \sum_{i\in I} \omega_i \langle L_i x - L_i\overline{x} \mid F_i(L_i x) - F_i(L_i\overline{x})\rangle \leqslant 0. \tag{4.49}$$

We conclude that $(\forall i \in I)\ F_i(L_i x) = F_i(L_i\overline{x}) = p_i$. $\square$

**Remark 4.3.19** Consider the setting of Problem 4.2.3 and set $\boldsymbol{\mathcal{G}} = \bigoplus_{i\in I} \mathcal{G}_i$, $\boldsymbol{L}\colon \mathcal{H} \to \colon \boldsymbol{\mathcal{G}}\colon x \mapsto (L_i x)_{i\in I}$, $\boldsymbol{F}\colon \boldsymbol{\mathcal{G}} \to \boldsymbol{\mathcal{G}}\colon (y_i)_{i\in I} \mapsto (F_i y_i)_{i\in I}$, and $\boldsymbol{p} = (p_i)_{i\in I}$. Note that

$$\text{Problem 4.2.1 admits a solution if and only if } \boldsymbol{p} \in \boldsymbol{F}\big(\boldsymbol{L}(C)\big). \tag{4.50}$$

Thus, the quantity $d_{\boldsymbol{F}(\boldsymbol{L}(C))}(\boldsymbol{p})$ provides a measure of inconsistency of Problem 4.2.1. We can actually use a solution to Problem 4.2.3 to estimate it. Indeed, suppose that $\overline{x}_1$ and $\overline{x}_2$ are solutions to (4.10). Then (4.3) yields

$$\begin{aligned}
\sum_{i\in I} \omega_i \|F_i(L_i\overline{x}_1) - F_i(L_i\overline{x}_2)\|^2 &\leqslant \sum_{i\in I} \omega_i \langle L_i\overline{x}_1 - L_i\overline{x}_2 \mid F_i(L_i\overline{x}_1) - F_i(L_i\overline{x}_2)\rangle \\
&= \sum_{i\in I} \omega_i \langle L_i(\overline{x}_1 - \overline{x}_2) \mid F_i(L_i\overline{x}_1) - p_i\rangle \\
&\quad + \sum_{i\in I} \omega_i \langle L_i(\overline{x}_2 - \overline{x}_1) \mid F_i(L_i\overline{x}_2) - p_i\rangle \\
&\leqslant 0.
\end{aligned} \tag{4.51}$$

Hence, for every $i \in I$, there exists a unique $\overline{p}_i \in \mathcal{G}_i$ such that every solution $\overline{x}$ to Problem 4.2.3 satisfies

$$F_i(L_i\overline{x}) = \overline{p}_i. \tag{4.52}$$

In turn, if $\overline{x}$ is any solution to Problem 4.2.3, then

$$d_{\boldsymbol{F}(\boldsymbol{L}(C))}(\boldsymbol{p}) = \inf_{x\in C} \|\boldsymbol{p} - \boldsymbol{F}(\boldsymbol{L}x)\| \leqslant \|\boldsymbol{p} - \boldsymbol{F}(\boldsymbol{L}\overline{x})\| = \|\boldsymbol{p} - \overline{\boldsymbol{p}}\| = \sqrt{\sum_{i\in I} \|p_i - \overline{p}_i\|^2}. \tag{4.53}$$

Next, we turn to the existence of solutions.

**Proposition 4.3.20** *Problem 4.2.3 admits a solution in each of the following instances.*

(i) $\sum_{i\in I} \omega_i L_i^* p_i \in \operatorname{ran}\,(N_C + \sum_{i\in I} \omega_i L_i^* \circ F_i \circ L_i)$.

(ii) *$C$ is bounded.*

(iii) $\operatorname{ran} N_C + \sum_{i\in I} \omega_i L_i^*(\operatorname{ran} F_i) = \mathcal{H}$.

(iv) *For some $i \in I$, $L_i^*$ is surjective and one of the following holds:*

a) $L_i^*(\operatorname{ran} F_i) = \mathcal{H}$.

b) *$F_i$ is surjective.*

c) $\|F_i(y)\| \to +\infty$ *as* $\|y\| \to +\infty$.

d) $\operatorname{ran}(\operatorname{Id} - F_i)$ *is bounded.*

e) *There exists a continuous convex function $g_i\colon \mathcal{G}_i \to \mathbb{R}$ such that $F_i = \operatorname{prox}_{g_i}$.*

*Proof.* Set $A = N_C$ and $(\forall i \in I)\ B_i = \omega_i F_i$. Then the operators $(B_i)_{i\in I}$ are cocoercive. Now define

$$M = A + \sum_{i\in I} L_i^* \circ B_i \circ L_i. \tag{4.54}$$

It follows from [6, Proposition 4.12] that $B = \sum_{i\in I} L_i^* \circ B_i \circ L_i$ is cocoercive and hence maximally monotone by [6, Example 20.31], with $\operatorname{dom} B = \mathcal{H}$. On the other hand, [6, Example 20.26] asserts that $A$ is maximally monotone. We therefore derive from [6, Corollary 25.5(i)] that

$$M \text{ is maximally monotone.} \tag{4.55}$$

(i): Let $x \in \mathcal{H}$. In view of (4.19), $x$ solves Problem 4.2.3 if and only if

$$-\sum_{i\in I} \omega_i L_i^*\big(F_i(L_i x) - p_i\big) \in N_C x, \tag{4.56}$$

that is, $\sum_{i\in I} \omega_i L_i^* p_i \in Mx$.

(ii): Since $\operatorname{dom} M = \operatorname{dom} A = C$ is bounded, it follows from (4.55) and [6, Corollary 21.25] that $M$ is surjective, so (i) holds.

(iii): It follows from [6, Example 25.14] that $A$ is $3^*$ monotone and from [6, Example 25.20(i)] that the operators $(B_i)_{i\in I}$ are likewise. Hence, in view of (4.55) we invoke Lemma 4.2.8 to get

$$\operatorname{int}\operatorname{ran} M = \operatorname{int}\operatorname{ran}\Big(A + \sum_{i\in I} L_i^* \circ B_i \circ L_i\Big) = \operatorname{int}\Big(\operatorname{ran} A + \sum_{i\in I} L_i^*(\operatorname{ran} B_i)\Big) = \mathcal{H}. \tag{4.57}$$

So $M$ is surjective and (i) holds.

(iv)b)⇒(iv)a)⇒(iii): Clear.

(iv)c)⇒(iv)b): Since $F_i$ is maximally monotone by [6, Example 20.30], this follows from [6, Corollary 21.24].

(iv)d)⇒(iv)c): Set $\rho = \sup_{y\in\mathcal{G}_i} \|y - F_i y\|$. Then $\|F_i y\| \geqslant \|y\| - \|y - F_i y\| \geqslant \|y\| - \rho \to +\infty$ as $\|y\| \to +\infty$.

(iv)e)⇒(iv)b): We derive from [6, Proposition 16.27] that $\mathcal{G}_i = \operatorname{int}\operatorname{dom} g_i \subset \operatorname{dom}\partial g_i = \operatorname{dom}(\operatorname{Id} + \partial g_i) = \operatorname{ran}(\operatorname{Id} + \partial g_i)^{-1} = \operatorname{ran}\operatorname{prox}_{g_i}$. □

**Example 4.3.21** A simple instance when Problem 4.2.1 has no solution, while the relaxed Problem 4.2.3 does, is the following. Take disjoint nonempty closed convex subsets $C$ and $D$ of $\mathcal{H}$ such that $C$ is bounded, and let $I = 1$, $\mathcal{G}_1 = \mathcal{H}$, $L_1 = \mathrm{Id}$, $F_1 = \mathrm{Id} - \mathrm{proj}_D$, and $p_1 = 0$. Then the solution set of Problem 4.2.1 is $C \cap D = \varnothing$, while that of Problem 4.2.3 is $\mathrm{Fix}\,(\mathrm{proj}_C \circ \mathrm{proj}_D) \neq \varnothing$ [33].

We have described in Example 4.2.2 an instance of the relaxed Problem 4.2.3 which is in fact a minimization problem. The next proposition describes a general setting in which a minimization problem underlies Problem 4.2.3. It involves the Moreau envelope of (4.17).

**Proposition 4.3.22** *Consider the setting of Problem 4.2.3 and suppose that, for every $i \in I$, there exists $g_i \in \Gamma_0(\mathcal{G}_i)$ such that $F_i = \mathrm{prox}_{g_i}$. Then the objective of Problem 4.2.3 is to*

$$\underset{x \in C}{\text{minimize}} \;\; f(x), \quad \text{where} \quad f \colon x \mapsto \frac{1}{2} \sum_{i \in I} \omega_i \Big( \widetilde{g_i^*}(L_i x) - \langle L_i x \mid p_i \rangle \Big). \tag{4.58}$$

*Proof*. We derive from [6, Proposition 24.4] that $(\forall i \in I)\; \nabla \widetilde{g_i^*} = \mathrm{prox}_{g_i}$. In turn, $f$ is differentiable and

$$(\forall x \in \mathcal{H}) \quad \nabla f(x) = \sum_{i \in I} \omega_i L_i^* \big( \mathrm{prox}_{g_i}(L_i x) - p_i \big) = \sum_{i \in I} \omega_i L_i^* \big( F_i(L_i x) - p_i \big). \tag{4.59}$$

Consequently, (4.10) is equivalent to finding a solution to (4.23), i.e., by Lemma 4.2.7, to minimizing $f$ over $C$. □

Next, we present a block-iterative algorithm for solving Problem 4.2.3.

**Proposition 4.3.23** *Consider the setting of Problem 4.2.3 under the assumption that it has a solution. Let $K$ be a strictly positive integer and let $(I_n)_{n \in \mathbb{N}}$ be a sequence of nonempty subsets of $I$ such that*

$$(\forall n \in \mathbb{N}) \quad \bigcup_{k=0}^{K-1} I_{n+k} = I. \tag{4.60}$$

*Let $x_0 \in \mathcal{H}$, let $\gamma \in \left]0, 2\right[$, and, for every $i \in I$, let $t_{i,-1} \in \mathcal{H}$ and set $\gamma_i = \gamma / \|L_i\|^2$. Iterate*

$$\begin{array}{l} \text{for } n = 0, 1, \ldots \\ \quad \text{for every } i \in I_n \\ \quad\quad t_{i,n} = x_n - \gamma_i L_i^* \big( F_i(L_i x_n) - p_i \big) \\ \quad \text{for every } i \in I \smallsetminus I_n \\ \quad\quad t_{i,n} = t_{i,n-1} \\ \quad x_{n+1} = \mathrm{proj}_C \Big( \sum_{i=1}^{m} \omega_i t_{i,n} \Big). \end{array} \tag{4.61}$$

*Then $(x_n)_{n \in \mathbb{N}}$ converges weakly to a solution to Problem 4.2.3.*

*Proof.* Set $A_0 = N_C$ and $(\forall i \in I)\ A_i = \|L_i\|^{-2}(L_i^* \circ F_i \circ L_i - L_i^* p_i)$. For every $i \in I$, since $F_i$ is firmly nonexpansive, it follows from [6, Proposition 4.12] that $A_i$ is firmly nonexpansive, i.e., cocoercive with $\beta_i = 1$. Thus, (4.61) is a special case of (4.28), and the conclusion follows from Proposition 4.2.10. □

An attractive feature of (4.61) is its ability to activate only a subblock of operators $(F_i)_{i \in I_n}$ at iteration $n$, as opposed to all of them as in classical algorithms dealing with inconsistent common fixed point problems [16–18, 20]. This flexibility is of the utmost relevance for very large-scale applications. It will also be seen in Section 4.3.2 to lead to more efficient implementations. Condition (4.60) regulates the frequency of activation of the operators. Since $K$ can be chosen arbitrarily, it is actually quite mild.

### 4.3.2 Numerical experiments

In this section, we illustrate the ability of the proposed framework to efficiently model and solve various signal and image recovery problems with inconsistent nonlinear prescriptions. Each instance will use the block-iterative algorithm (4.61) which was shown in Proposition 4.3.23 to produce an exact solution of Problem 4.2.3 from any initial point in $\mathcal{H}$. Here, we implement it with $x_0 = 0$.

**Remark 4.3.24** In the modeling of signal construction problems as minimization problems, it is common practice to add a function $g$ to the objective in order to promote desirable properties in the solutions. Several functions are thus averaged and contribute collectively to defining solutions. A prominent example is the promotion of sparsity through the addition of a penalty such as the $\ell^1$ norm in $\mathbb{R}^N$ [14, 57]. In the more general variational inequality setting of Problem 4.2.3, this template can be mimicked by adding the prescription $Fy = 0$, where $F = \mathrm{Id} - \mathrm{prox}_g$, i.e., by Moreau's decomposition, $F = \mathrm{prox}_{g^*}$ [6, Remark 14.4]. Note that exact satisfaction of the equality $Fy = 0$ would just mean that one minimizes $g$ since $\mathrm{Fix}\,\mathrm{prox}_g = \mathrm{Argmin} g$. In general, when incorporated to Problem 4.2.3, the pair $(F, p) = (\mathrm{Id} - \mathrm{prox}_g, 0)$ is intended to promote the properties $g$ would in a standard minimization problem. We investigate in Sections 4.3.2.3 and 4.3.2.4 this technique to encourage sparsity in $\mathbb{R}^N$ through the incorporation of the operator $F = \mathrm{proj}_{B_\infty(0;\rho)} = \mathrm{Id} - \mathrm{prox}_{\rho\|\cdot\|_1}$, where $B_\infty(0;\rho)$ is the $\ell^\infty$ ball of $\mathbb{R}^N$ centered at the origin and with radius $\rho \in \left]0, +\infty\right[$.

#### 4.3.2.1 Image recovery

The goal is to recover the original image $\overline{x} \in \mathcal{H} = \mathbb{R}^N$ $(N = 256^2)$ shown in Figure 4.5(a) from the following.

- Bounds on pixel values: $\overline{x} \in C = [0, 255]^N$.
- The degraded image $p_1 \in \mathcal{G}_1 = \mathcal{H}$ shown in Figure 4.5(b), which is modeled as follows. The image $\overline{x}$ is blurred by $L_1 \colon \mathcal{H} \to \mathcal{G}_1$, which performs discrete convolution with a

$15 \times 15$ Gaussian kernel with standard deviation of $3.5$, then corrupted by an additive noise $w_1 \in \mathcal{G}_1$. The blurred image-to-noise ratio is $20 \log_{10}(\|L_1 \overline{x}\|/\|w_1\|) = 24.0$ dB. Pixel values beyond $60$ are then clipped. Altogether, $p_1 = \text{proj}_{D_1}(L_1 \overline{x} + w_1)$, where $D_1 = [0, 60]^N$. This process models a low-quality image acquired by a device which cannot detect photon counts beyond a certain threshold. We therefore use $F_1 = \text{proj}_{D_1}$ in (4.10).

- An approximation of the mean pixel value $\rho_2 = 138$ of $\overline{x}$. To enforce this information, following Example 4.2.2, we set $\mathcal{G}_2 = \mathcal{H}$, $L_2 = \text{Id}$, $p_2 = 0$, and

$$F_2 \colon (\eta_k)_{1 \leqslant k \leqslant N} \mapsto x - \left( \frac{\sum_{k=1}^N \eta_k}{N} - \rho_2 \right) \mathbf{1}. \tag{4.62}$$

- The phase $\theta \in [-\pi, \pi]^N$ of the 2-D discrete Fourier transform of a noise-corrupted version of $\overline{x}$, i.e., $\theta = \angle \text{DFT}\,(\overline{x} + w_3)$, where $w_3 \in \mathcal{H}$ yields an image-to-noise ratio $20 \log_{10}(\|\overline{x}\|/\|w_3\|) = 49.0$ dB. To model this information, we set $\mathcal{G}_3 = \mathcal{H}$, $L_3 = \text{Id}$, $p_3 = 0$, and

$$F_3 \colon y \mapsto y - \text{IDFT}\left( \left| \text{DFT}\, y \right| \max\left\{ \cos\left( \angle(\text{DFT}\, y) - \theta \right), 0 \right\} \exp(i\theta) \right). \tag{4.63}$$

Due to the noise present in $p_1$ and $\theta$, and the inexact estimation of $\rho_2$, this instance of Problem 4.2.1 ($I = \{1, 2, 3\}$) is inconsistent. We thus arrive at the relaxed Problem 4.2.3 by setting $\omega_1 = \omega_2 = \omega_3 = 1/3$. By Proposition 4.3.20(ii), since $C$ is bounded, Problem 4.2.3 is guaranteed to possess a solution. The solution shown in Figure 4.5(c) is computed using algorithm (4.61) with $\gamma = 1.9$ and $(\forall n \in \mathbb{N})\ I_n = I$. This experiment illustrates a nonlinear recovery scenario with inconsistent measurements which nonetheless produces realistic solutions obtained by exploiting all available information.

(a) (b) (c)

**Figure 4.5** Experiment of Section 4.3.2.1: (a) Original image $\overline{x}$. (b) Degraded image $p_1$. (c) Recovered image.

#### 4.3.2.2 Signal recovery

The goal is to recover the original signal $\overline{x} \in \mathcal{H} = C = \mathbb{R}^N$ ($N = 1024$) shown in Figure 4.6(a) from the following.

- A piecewise constant approximation $p_1$ of $\overline{x}$, given by $p_1 = \mathrm{proj}_{D_1}(\overline{x} + w_1)$, where $w_1 \in \mathcal{G}_1$ represents noise and $D_1$ is the subspace of signals in $\mathcal{G}_1 = \mathcal{H}$ which are constant by blocks along each of the 16 sets of 64 consecutive indices in $\{1, \dots, N\}$ (see Figure 4.6(b)). The signal-to-noise ratio is $20 \log_{10}(\|\overline{x}\|/\|w_1\|) = -2.3$ dB. We model this observation by setting $L_1 = \mathrm{Id}$ and $F_1 = \mathrm{proj}_{D_1}$.

- A bound $\rho_2 = 0.025$ on the magnitude of the finite differences of $\overline{x}$. To enforce this information, following Example 4.2.2, we set $\mathcal{G}_2 = \mathbb{R}^{N-1}$, $L_2 \colon \mathcal{H} \to \mathcal{G}_2 \colon (\xi_i)_{1 \leqslant i \leqslant N} \mapsto (\xi_{i+1} - \xi_i)_{1 \leqslant i \leqslant N-1}$, $p_2 = 0$, and $F_2 = \mathrm{Id} - \mathrm{proj}_{D_2}$, where $D_2 = \big\{ y \in \mathcal{G}_2 \;\big|\; \|y\|_\infty \leqslant \rho_2 \big\}$, that is, using (4.41),
$$F_2 \colon (\eta_k)_{1 \leqslant k \leqslant N-1} \mapsto \big( \mathrm{soft}_\gamma(\eta_k) \big)_{1 \leqslant k \leqslant N-1}. \tag{4.64}$$

- A collection of $m = 1200$ noisy thresholded scalar observations $r_3 = (\chi_j)_{j \in J} \in \mathbb{R}^m$ of $\overline{x}$, where $J = \{3, \dots, m+2\}$. The true data formation model is
$$(\forall j \in J) \quad \chi_j = R(\langle \overline{x} \mid e_j \rangle) + \nu_j, \tag{4.65}$$
where $(e_j)_{j \in J}$ is a dictionary of random vectors in $\mathbb{R}^N$ with zero-mean i.i.d. entries, the noise vector $w_3 = (\nu_j)_{j \in J}$ yields a signal-to-noise ratio of $20 \log_{10}(\|r_3\|/\|w_3\|) = 17.8$ dB, and $R$ is the thresholding operator of the type found in [2, 52] ($\rho = 0.05$), namely
$$R \colon \mathbb{R} \to \mathbb{R} \colon \eta \mapsto \begin{cases} \mathrm{sign}(\eta) \sqrt[4]{\eta^4 - \rho^4}, & \text{if } |\eta| > \rho; \\ 0, & \text{if } |\eta| \leqslant \rho. \end{cases} \tag{4.66}$$
We assume that $R$ is misspecified and that the presence of noise is unknown, so that the data acquisition process is incorrectly modeled as
$$(\forall j \in J) \quad \chi_j = Q(\langle \overline{x} \mid e_j \rangle), \tag{4.67}$$
where
$$Q \colon \mathbb{R} \to \mathbb{R} \colon \eta \mapsto \begin{cases} \mathrm{sign}(\eta) \sqrt{\eta^2 - \rho^2}, & \text{if } |\eta| > \rho; \\ 0, & \text{if } |\eta| \leqslant \rho. \end{cases} \tag{4.68}$$
Note that $Q$ is not Lipschitzian. Nonetheless, with
$$S \colon \mathbb{R} \to \mathbb{R} \colon \eta \mapsto \mathrm{sign}(\eta) \left( \sqrt{\eta^2 + \rho^2} - \rho \right), \tag{4.69}$$
it is straightforward to verify that $S \circ Q = \mathrm{soft}_\rho$ and that, for every $j \in J$, $(F_j, p_j) =$

$(\ \mathrm{soft}_\rho\,, S\chi_j)$ is a proxification of $(Q, \chi_j)$. Also, for every $j \in J$, set $\mathcal{G}_j = \mathbb{R}$ and $L_j = \langle \cdot \mid e_j \rangle$.

We thus obtain an instantiation of Problem 4.2.3 with $I = \{1, 2\} \cup J$ and, for every $i \in I$, $\omega_i = 1/(\text{card } I)$. Since $(e_j)_{j\in J}$ is overcomplete and, for every $j \in J$, $F_j$ is surjective, it follows that $\mathcal{H} = \left\{\sum_{j\in J} \omega_j \eta_j e_j \;\middle|\; \eta_j \in \text{ran } F_j\right\} = \sum_{j\in J} \omega_j L_j^*(\text{ran } F_j) \subset \sum_{i\in I} \omega_i L_i^*(\text{ran } F_i)$, so Problem 4.2.3 is guaranteed to possess a solution by Proposition 4.3.20(iii). Algorithm (4.61) produces the signal shown in Figure 4.6(c) with $\gamma = 1.9$ and the following activation strategy. At every iteration, $F_1$ and $F_2$ are activated, while we partition $J$ into four blocks of $300$ elements and cyclically activate one block per iteration, i.e.,

$$(\forall n \in \mathbb{N})(\forall j \in \{0, 1, 2, 3\}) \quad I_{4n+j} = \{1, 2, 3 + 300j, \ldots, 2 + 300(j + 1)\}, \tag{4.70}$$

which satisfies condition (4.60) with $K = 4$. This shows that, even when the data is noisy and poorly modeled, Problem 4.2.3 produces quite robust recoveries. The execution time savings resulting from the use of (4.70) compared to the full activation strategy (i.e., $I_n = I$ for every $n \in \mathbb{N}$) are displayed in Figure 4.7. Note that in very large-scale scenarios in which all data cannot be simultaneously loaded into memory, activation strategies such as (4.70) make algorithm (4.61) implementable.

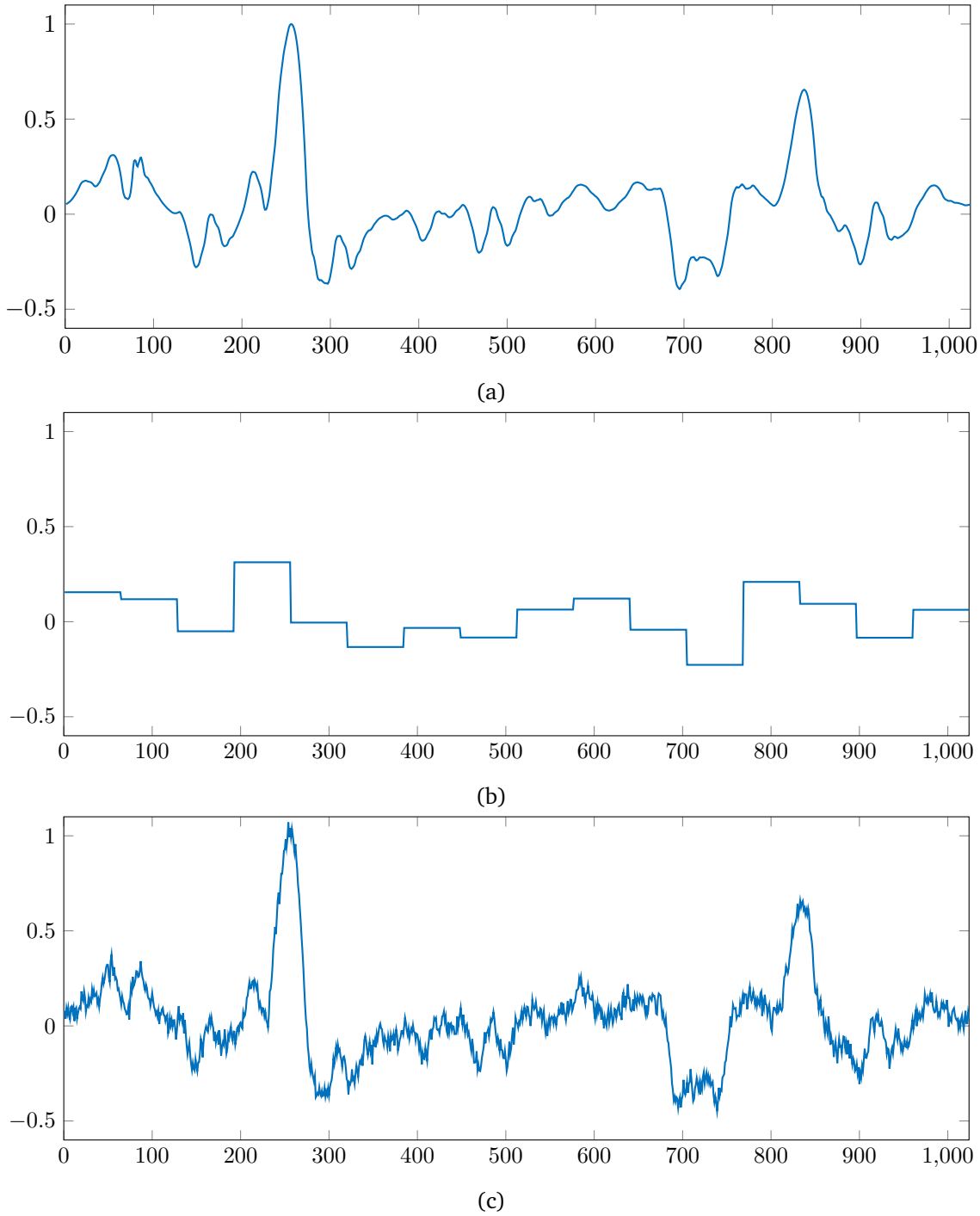



**Figure 4.6** Experiment of Section 4.3.2.2: (a): Original signal $\overline{x}$. (b): Piecewise constant approximation $p_1$. (c): Recovered signal.

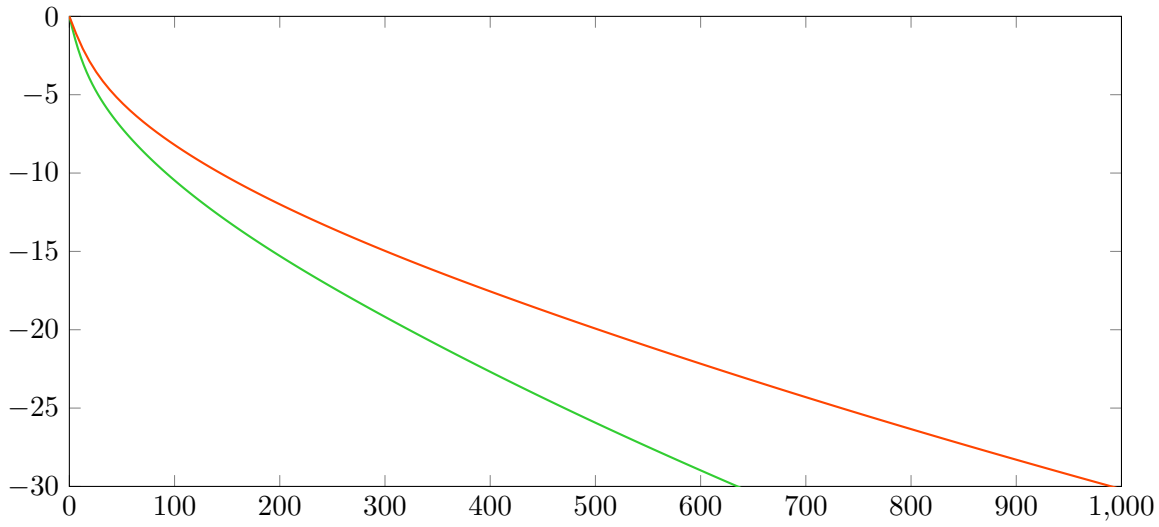



**Figure 4.7** Experiment of Section 4.3.2.2: Relative error $20\log_{10}(\|x_n - x_\infty\|/\|x_0 - x_\infty\|)$ (dB) versus execution time (seconds) for full activation (red) and cyclic activation (4.70) (green).

#### 4.3.2.3 Sparse image recovery

The goal is to recover the original image $\overline{x} \in \mathcal{H} = \mathbb{R}^N$ ($N = 256^2$) shown in Figure 4.8(a) from the following.

- Bounds on pixel values: $x \in C = [0, 255]^N$.
- The low rank approximation $q_1 \in \mathcal{G}_1 = \mathcal{H}$ displayed in Figure 4.8(b) of a blurred noisy version of $\overline{x}$ modeled as follows. The blurring operator $L_1 \colon \mathcal{H} \to \mathcal{G}_1$ applies a discrete convolution with a uniform $7 \times 7$ kernel, and the operators $Q$ and $S$ are as in Example 4.3.13, with threshold $\rho = 500$. Then $q_1 = Q(L_1\overline{x} + w_1)$ is a rank-85 compression, where $w_1 \in \mathcal{G}_1$ induces a blurred image-to-noise ratio of $20\log_{10}(\|L_1\overline{x}\|/\|w_1\|) = 17.6$ dB. By Example 4.3.13, we obtain a proxification of $(Q, q_1)$ with $(F_1, p_1) = (S \circ Q, Sq_1)$.
- $\overline{x}$ is sparse. To promote this property in the solutions to (4.10), following Remark 4.3.24, we set $\mathcal{G}_2 = \mathcal{H}$, $L_2 = \mathrm{Id}$, $p_2 = 0$, $\rho_2 = 1.5$, and $F_2 = \mathrm{proj}_{B_\infty(0;\rho_2)}$.

We therefore arrive at an instance of Problem 4.2.3 with $I = \{1, 2\}$ and $\omega_1 = \omega_2 = 1/2$. Since $C$ is bounded, Proposition 4.3.20(ii) guarantees that a solution exists. Algorithm (4.61) with $\gamma = 1$ yields the recovery in Figure 4.8(c). Even though computing $F_1$ requires only one singular value decomposition (not two, as (4.42) may suggest), it is the most numerically expensive operator in this problem. Therefore, we choose to activate $F_1$ only every 5 iterations, i.e.,

$$I_n = \begin{cases} I \smallsetminus \{1\}, & \text{if } n \not\equiv 0 \mod 5; \\ I, & \text{if } n \equiv 0 \mod 5. \end{cases} \tag{4.71}$$

Figure 4.10 displays the time savings resulting from the use of (4.71) compared to full activation (both activation strategies yield visually indistinguishable recoveries). Notice that, while the observation in Figure 4.8(b) is virtually illegible, many of the words in the recovery of Figure 4.8(c) can be discerned.


(a) (b) (c)

**Figure 4.8** Experiment of Section 4.3.2.3: (a) Original image $\overline{x}$. (b) Degraded image $q_1$. (c) Recovered image.

Finally, we examine the use of the non firmly nonexpansive sparsity-promoting operator of Example 4.3.15. Specifically, $Q_{\gamma g}$ is given by (4.44), which is induced by the logarithmic penalty with parameters $\rho = \rho_2$ and $\gamma = 0.05/\rho_2^2$. This implies that $0.95Q_{\gamma g}$ is firmly nonexpansive and hence that Id $-0.95Q_{\gamma g}$ is likewise. Figure 4.9 displays the result when $F_2$ is replaced by componentwise applications of Id $-0.95Q_{\gamma g}$. In this experiment, the $\ell^1$ penalty-based operator $F_2$ yields a sharper recovery in Figure 4.8(c) than the recovery in Figure 4.9, which is induced by the logarithmic penalty.

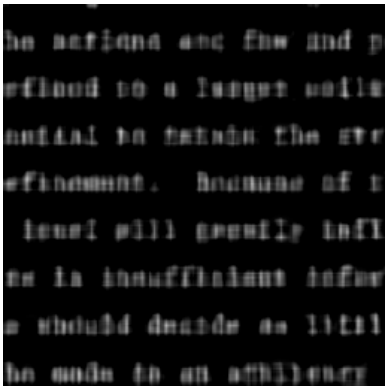


**Figure 4.9** Experiment of Section 4.3.2.3: Recovered image with logarithmic thresholding instead of soft thresholding.

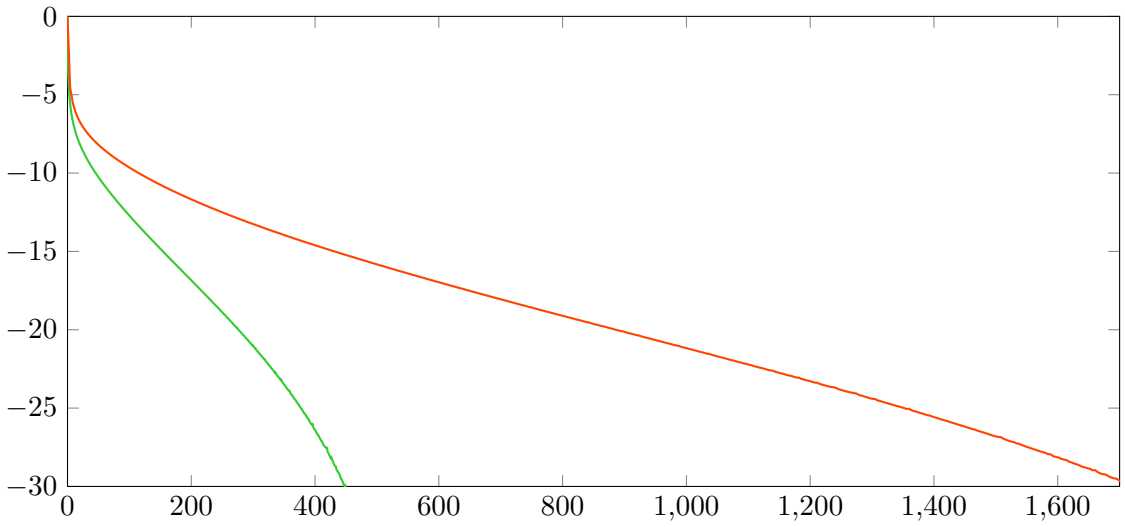



**Figure 4.10** Experiment of Section 4.3.2.3: Relative error $20\log_{10}(\|x_n - x_\infty\|/\|x_0 - x_\infty\|)$ (dB) versus execution time (seconds) for full-activation (red) and block activation (4.71) (green).

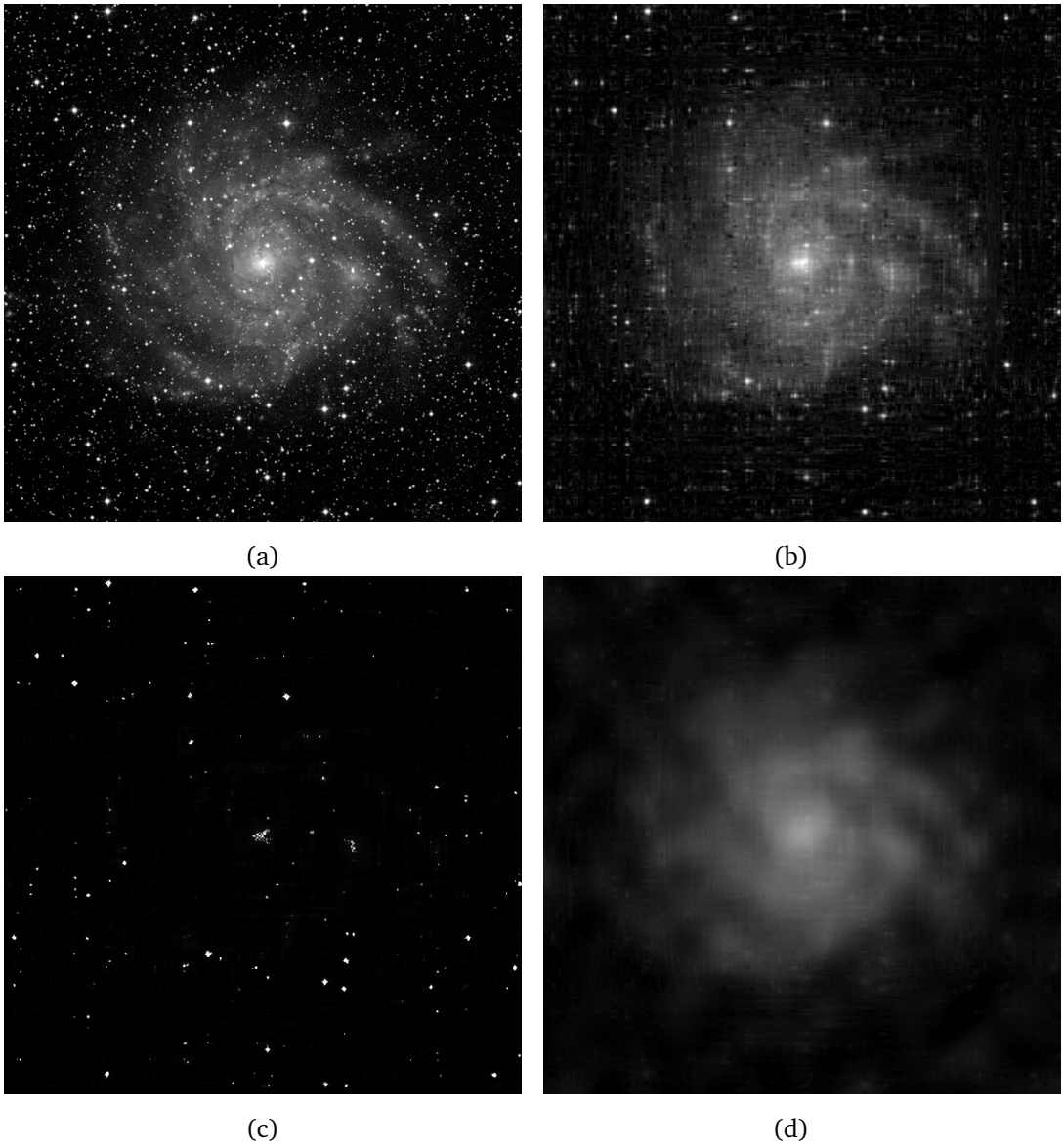
#### 4.3.2.4 Source separation

This experiment incorporates nonlinear compression to a problem in astronomy, which seeks to separate a background image $\overline{x}_1 \in \mathbb{R}^N$ ($N = 600^2$) of stars from a galaxy image $\overline{x}_2 \in \mathbb{R}^N$ [40]. The goal is to construct the image pair $(\overline{x}_1, \overline{x}_2) \in \mathcal{H} = \mathbb{R}^N \times \mathbb{R}^N$ given the following.

- Bounds on pixel values: $(\overline{x}_1, \overline{x}_2) \in C = [0, 255]^N \times [0, 255]^N$.
- The low rank approximation $q_1 \in \mathcal{G}_1 = \mathbb{R}^N$ shown in Figure 4.11(b) of the original superposition $\overline{x}_1 + \overline{x}_2$ shown in Figure 4.11(a), which is modeled as follows. Set $L_1 \colon \mathcal{H} \to \mathcal{G}_1 \colon (x_1, x_2) \mapsto x_1 + x_2$, and let $Q$ and $S$ be as in Example 4.3.13 with $\rho = 1500$. The resulting rank-22 approximation of $\overline{x}_1 + \overline{x}_2$ is given by $q_1 = Q(L_1(\overline{x}_1, \overline{x}_2))$. It follows from Example 4.3.13 that $(F_1, p_1) = (S \circ Q, Sq_1)$ is a proxification of $(Q, q_1)$.
- $\overline{x}_1$ is sparse, and $\overline{x}_2$ admits a sparse representation relative to the 2-D discrete cosine transform $L \colon \mathbb{R}^N \to \mathbb{R}^N$ [40]. To encourage these properties, as discussed in Remark 4.3.24, we set $\mathcal{G}_2 = \mathcal{H}$, $L_2 \colon (x_1, x_2) \mapsto (x_1, Lx_2)$, $p_2 = 0$, and $F_2 \colon (y_1, y_2) \mapsto (\mathrm{proj}_{B_\infty(0;10)} y_1, \mathrm{proj}_{B_\infty(0;45)} y_2)$. In view of Example 4.3.1, $F_2$ is firmly nonexpansive.

Thus, we arrive at an instance of Problem 4.2.3 with $I = \{1, 2\}$ and $\omega_1 = \omega_2 = 1/2$. By Proposition 4.3.20(ii) this problem is guaranteed to possess a solution, since $C$ is bounded. Algorithm (4.61) with $\gamma = 1$ provides the solution shown in Figure 4.11(c)–(d). To improve algorithmic performance, we adopt the activation strategy (4.71); see Figure 4.12 for time savings compared to the full activation strategy. As can be seen from Figure 4.11, this approach produces effective recoveries. Even though this problem involves a discontinuous observation

process, we can nonetheless solve it with algorithm (4.61), which exploits all of the information at hand.



(a) (b)

(c) (d)

**Figure 4.11** Experiment of Section 4.3.2.4: (a) Original image $\overline{x}_1 + \overline{x}_2$. (b) Low-rank compression of $\overline{x}_1 + \overline{x}_2$. (c) Recovered background (stars). (d) Recovered foreground (galaxy).

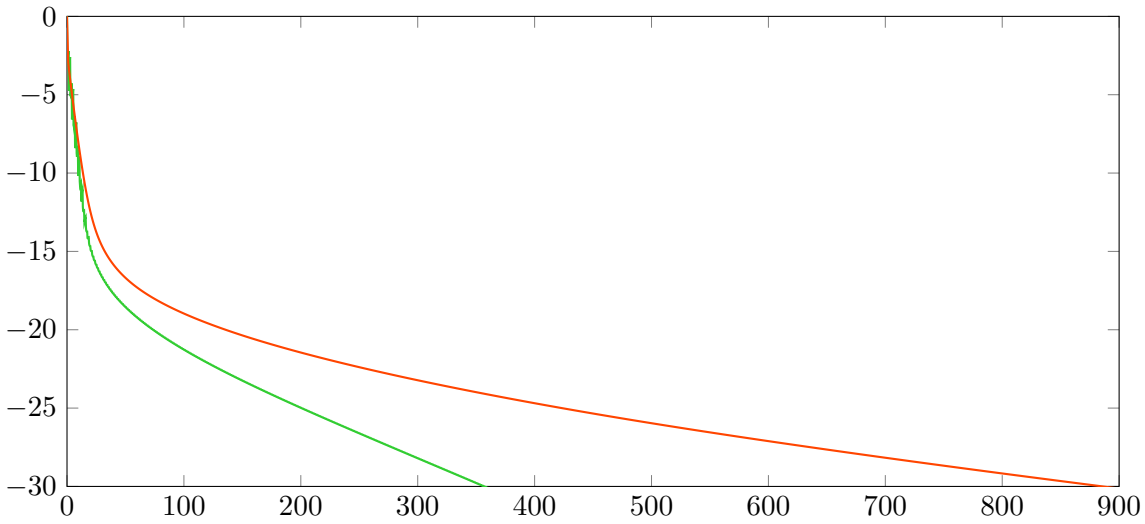



**Figure 4.12** Experiment of Section 4.3.2.4: Relative error $20\log_{10}(\|x_n - x_\infty\|/\|x_0 - x_\infty\|)$ (dB) versus execution time (seconds) for full-activation (red) and block activation (4.71) (green).

# References


[1] J. S. Abel and J. O. Smith, Restoring a clipped signal, *Proc. Int. Conf. Acoust. Speech Signal Process.*, vol. 3, pp. 1745–1748, 1991.

[2] F. Abramovich, T. Sapatinas, and B. W. Silverman, Wavelet thresholding via a Bayesian approach, *J. R. Stat. Soc. Ser. B Stat. Methodol.*, vol. 60, pp. 725–749, 1998.

[3] H. Andrews and C. Patterson, Singular value decomposition (SVD) image decoding, *IEEE Trans. Commun.*, vol. 24, pp. 425–432, 1976.

[4] F. R. Ávila, M. P. Tcheou, and L. W. P. Biscainho, Audio soft declipping based on constrained weighted least squares, *IEEE Signal Process. Lett.*, vol. 24, pp. 1348–1352, 2017.

[5] R. E. Barlow and H. D. Brunk, The isotonic regression problem and its dual, *J. Amer. Stat. Assoc.*, vol. 67, pp. 140–147, 1972,

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

[7] I. Bayram and S. Bulek, A penalty function promoting sparsity within and across groups, *IEEE Trans. Signal Process.*, vol. 65, pp. 4238–4251, 2017.

[8] A. Bensoussan and J.-L. Lions, *Applications des Inéquations Variationnelles en Contrôle Stochastique.* Bordas, Paris, 1978. English translation: *Applications of Variational Inequalities in Stochastic Control.* North Holland, New York, 1982.

[9] H. Boche, M. Guillemard, G. Kutyniok, and F. Philipp, Signal recovery from thresholded frame measurements, *Proc. 15th SPIE Wavelets Sparsity Conf.*, vol. 8858, pp. 80–86, 2013.

[10] H. Brézis and A. Haraux, Image d'une somme d'opérateurs monotones et applications, *Israel J. Math.*, vol. 23, pp. 165–186, 1976.

[11] L. M. Briceño-Arias and P. L. Combettes, Convex variational formulation with smooth coupling for multicomponent signal decomposition and recovery, *Numer. Math. Theory Methods Appl.*, vol. 2, pp. 485–508, 2009.

[12] F. E. Browder, Nonlinear monotone operators and convex sets in Banach spaces, *Bull. Amer. Math. Soc.*, vol. 71, pp. 780–785, 1965.

[13] J. A. Cadzow, Signal enhancement – A composite property mapping algorithm, *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, pp. 49–62, 1988.

[14] E. Candès and T. Tao, Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE Trans. Inform. Theory*, vol. 52, pp. 5406–5425, 2006.

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

[16] 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.

[17] 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.

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

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

[20] P. L. Combettes, Systems of structured monotone inclusions: Duality, algorithms, and applications, *SIAM J. Optim.*, vol. 23, pp. 2420–2447, 2013.

[21] P. L. Combettes, Monotone operator theory in convex optimization, *Math. Program.*, vol. B170, pp. 177–206, 2018.

[22] P. L. Combettes and P. Bondon, Hard-constrained inconsistent signal feasibility problems, *IEEE Trans. Signal Process.*, vol. 47, pp. 2460–2468, 1999.

[23] 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.

[24] P. L. Combettes and L. E. Glaudin, Solving composite fixed point problems with block updates, *Adv. Nonlinear Anal.*, vol. 10, pp. 1154–1177, 2021.

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

[26] P. L. Combettes and Z. C. Woodstock, A fixed point framework for recovering signals from nonlinear transformations, *Proc. Europ. Signal Process. Conf.*, pp. 2120–2124, 2020.

[27] P. L. Combettes and Z. C. Woodstock, Reconstruction of functions from prescribed proximal points, submitted, 2021.

[28] S. J. Dilworth, N. J. Kalton, D. Kutzarova, and V. N. Temlyakov, The thresholding greedy algorithm, greedy bases, and duality, *Constr. Approx.*, vol. 19, pp. 575–597, 2003.

[29] M. Ehrgott, Ç. Güler, H. M. Hamacher, and L. Shao, Mathematical optimization in intensity modulated radiation therapy, *Ann. Oper. Res.*, vol. 175, pp. 309–365, 2010.

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

[31] S. Foucart and J. Li, Sparse recovery from inaccurate saturated measurements, *Acta Appl. Math.*, vol. 158, pp. 49–66, 2018.

[32] M. Goldburg and R. J. Marks II, Signal synthesis in the presence of an inconsistent set of constraints, *IEEE Trans. Circuits Syst.*, vol. 32, pp. 647–663, 1985.

[33] L. G. Gubin, B. T. Polyak, and E. V. Raik, The method of projections for finding the common point of convex sets, *Comput. Math. Math. Phys.*, vol. 7, pp. 1–24, 1967.

[34] P. Hall, G. Kerkyacharian, and D. Picard, Block threshold rules for curve estimation using kernel and wavelet methods, *Ann. Statist.*, vol. 26, pp. 922–942, 1998.

[35] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and Their Applications.* Academic Press, New York, 1980.

[36] L. Knockaert, B. De Backer and D. De Zutter, SVD compression, unitary transforms, and computational complexity, *IEEE Trans. Signal Process.*, vol. 47, pp. 2724–2729, 1999.

[37] A. M. Legendre, *Nouvelles Méthodes pour la Détermination des Orbites des Comètes.* Firmin Didot, Paris, 1805.

[38] D. Malioutov and A. Aravkin, Iterative log thresholding, *Proc. Int. Conf. Acoust., Speech, Signal Process.*, pp. 7198–7202, 2014.

[39] A. Marmin, A. Jezierska, M. Castella, and J.-C. Pesquet, Global optimization for recovery of clipped signals corrupted with Poisson-Gaussian noise, *IEEE Signal Process. Lett.*, vol. 27, pp. 970–974, 2020.

[40] M. B. McCoy, V. Cevher, Q. T. Dinh, A. Asaei, and L. Baldassarre, Convexity in source separation: Models, geometry, and algorithms, *IEEE Signal Process. Mag.*, vol. 31, pp. 87–95, 2014.

[41] K. Nasrollahi and T. B. Moeslund, Super-resolution: A comprehensive survey, *Mach. Vis. Appl.*, vol. 25, pp. 1423–1468, 2014.

[42] T. Pennanen, On the range of monotone composite mappings, *J. Nonlinear Convex Anal.*, vol. 2, pp. 193–202, 2001.

[43] B. Peters, B. R. Smithyman, and F. J. Herrmann, Projection methods and applications for seismic nonlinear inverse problems with multiple constraints, *Geophys.*, vol. 84, pp. R251–R269, 2019.

[44] A. Ranade, S. S. Mahabalarao, and S. Kale, A variation on SVD based image compression, *Image Vis. Comput.*, vol. 25, pp. 771–777, 2007.

[45] L. Rencker, F. Bach, W. Wang, and M. D. Plumbley, Sparse recovery and dictionary learning from nonlinear compressive measurements, *IEEE Trans. Signal Process.*, vol. 67, pp. 5659–5670, 2019.

[46] D. Rzepka, M. Miśkowicz, D. Kościelnik, and N. T. Thao, Reconstruction of signals from level-crossing samples using implicit information, *IEEE Access*, vol. 6, pp. 35001–35011, 2018.

[47] B. E. A. Saleh, Image synthesis: Discovery instead of recovery, in: H. Stark (ed.) *Image Recovery: Theory and Application*, pp. 463–498. Academic Press, San Diego, CA, 1987.

[48] R. J. Samworth, Recent progress in log-concave density estimation, *Statist. Sci.*, vol. 33, pp. 493–509, 2018.

[49] M. Schetzen, Nonlinear system modeling based on the Wiener theory, *Proc. IEEE*, vol. 69, pp. 1557–1573, 1981.

[50] I. Selesnick and M. Farshchian, Sparse signal approximation via nonseparable regularization, *IEEE Trans. Signal Process.*, vol. 65, pp. 2561–2575, 2017.

[51] N. T. Shaked and J. Rosen, Multiple-viewpoint projection holograms synthesized by spatially incoherent correlation with broadband functions, *J. Opt. Soc. Amer. A*, vol. 25, pp. 2129–2138, 2008.

[52] T. Tao and B. Vidakovic, Almost everywhere behavior of general wavelet shrinkage operators, *Appl. Comput. Harmon. Anal.*, vol. 9, pp. 72–82, 2000.

[53] E. Tarr, *Hack Audio*. Routledge, New York, 2019.

[54] V. N. Temlyakov, The best $m$-term approximation and greedy algorithms, *Adv. Comput. Math.*, vol. 8, pp. 249–265, 1998.

[55] T. Teshima, M. Xu, I. Sato, and M. Sugiyama, Clipped matrix completion: A remedy for ceiling effects, *Proc. AAAI Conf. Artif. Intell.*, pp. 5151–5158, 2019.

[56] N. T. Thao and M. Vetterli, Deterministic analysis of oversampled A/D conversion and decoding improvement based on consistent estimates, *IEEE Trans. Signal Process.*, vol. 42, pp. 519–531, 1994.

[57] R. Tibshirani, Regression shrinkage and selection via the lasso: A retrospective, *J. R. Statist. Soc. B*, vol. 73, pp. 273–282, 2011.

[58] L. B. White, The wide-band ambiguity function and Altes' $Q$-distribution: Constrained synthesis and time-scale filtering, *IEEE Trans. Inform. Theory*, vol. 38, pp. 886–892, 1992.

[59] J.-F. Yang and C.-L. Lu, Combined techniques of singular value decomposition and vector quantization for image coding, *IEEE Trans. Image Process.*, vol. 4, pp. 1141–1146, 1995.

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

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

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

[63] M. Yuan and Y. Lin, Model selection and estimation in regression with grouped variables, *J. R. Stat. Soc. Ser. B Stat. Methodol.*, vol. 68, pp. 49–67, 2006.

[64] C. A. Zarzer, On Tikhonov regularization with non-convex sparsity constraints, *Inverse Problems*, vol. 15, art. 025006, 2009.

[65] E. Zeidler, *Nonlinear Functional Analysis and Its Applications II/B: Nonlinear Monotone Operators*. Springer, New York, 1990.

Chapter 5

# BLOCK-ACTIVATED ALGORITHMS FOR MULTICOMPONENT FULLY NONSMOOTH MINIMIZATION

## 5.1 Introduction and context

As seen in Chapters 2, 3, and 4, provenly-convergent and block-iterative algorithms are essential, especially for large-scale problems. This chapter considers algorithms with these properties which are designed for fully nonsmooth multicomponent convex minimization problems. We set out a list of requirements for large-scale minimization tasks in data science, and for the first time we compare the only algorithms which satisfy our requirements. Numerical experiments supplement our findings.

This chapter presents the following article.

> M. N. Bùi, P. L. Combettes, and Z. C. Woodstock, Block-activated algorithms for multicomponent fully nonsmooth minimization, submitted.

## 5.2 Article: Block-activated algorithms for multicomponent fully nonsmooth minimization

**Abstract**. We investigate block-activated proximal algorithms for multicomponent minimization problems involving a separable nonsmooth convex function penalizing the components individually, and nonsmooth convex coupling terms penalizing linear mixtures of the components. In the case of smooth coupling functions, several algorithms exist and they are well understood. By contrast, in the fully nonsmooth case, few block-activated methods are available and little effort has been devoted to assessing their merits and numerical performance. The goal of the paper is

to address this gap. The numerical experiments concern machine learning and signal recovery problems.

### 5.2.1 Introduction

The goal of many signal processing and machine learning tasks is to exploit the observed data and the prior knowledge to produce a solution that represents information of interest. In this process of extracting information from data, structured convex optimization has established itself as an effective modeling and algorithmic framework; see for instance [3, 5, 9, 15, 19]. In state-of-the-art applications, the sought solution is often a tuple of vectors which reside in different spaces [1, 2, 6, 7, 13, 14, 17]. The following multicomponent minimization problem will be shown to capture a wide range of concrete scenarios. It consists of a separable term penalizing the components individually, and of coupling terms penalizing linear mixtures of the components.

**Problem 5.2.1** Let $(\mathcal{H}_i)_{1\leqslant i\leqslant m}$ and $(\mathcal{G}_k)_{1\leqslant k\leqslant p}$ be Euclidean spaces. For every $i \in \{1,\ldots,m\}$ and every $k \in \{1,\ldots,p\}$, let $f_i\colon \mathcal{H}_i \to \left]-\infty,+\infty\right]$ and $g_k\colon \mathcal{G}_k \to \left]-\infty,+\infty\right]$ be proper lower semicontinuous convex functions, and let $L_{k,i}\colon \mathcal{H}_i \to \mathcal{G}_k$ be a linear operator. The objective is to

$$\underset{x_1\in\mathcal{H}_1,\ldots,x_m\in\mathcal{H}_m}{\operatorname{minimize}} \quad \underbrace{\sum_{i=1}^{m} f_i(x_i)}_{\text{separable term}} + \sum_{k=1}^{p} \underbrace{g_k\Bigg(\sum_{i=1}^{m} L_{k,i}x_i\Bigg)}_{k\text{th coupling term}}. \tag{5.1}$$

We denote the solution set by $\mathscr{P}$.

To solve Problem 5.2.1 reliably without adding restrictions on its constituents (for instance smoothness or strong convexity of some functions involved in the model), we focus on algorithms that have the following flexible features:

① **Nondifferentiability:** None of the functions $f_1,\ldots,f_m,g_1,\ldots,g_p$ is assumed to be differentiable.

② **Splitting:** The functions $f_1,\ldots,f_m,g_1,\ldots,g_p$ and the linear operators are activated separately.

③ **Block activation:** As $m$ and $p$ can be very large, only a block of the proximity operators of the functions $f_1,\ldots,f_m,g_1,\ldots,g_p$ is activated at each iteration.

④ **Operator norms:** Bounds on the norms of the linear operators involved in Problem 5.2.1 are not assumed.

⑤ **Convergence:** The algorithm produces a sequence which converges (possibly almost surely) to a solution to (5.1).

A consequence of features ① and ② is that the algorithms under consideration must activate the functions $f_1, \ldots, f_m, g_1, \ldots, g_p$ via their respective proximity operators (even if some functions happened to be smooth, proximal activation is often preferable [6, 11]). Feature ③ has a view towards current large-scale problems. In such scenarios, memory and computing power limitations make the execution of standard proximal splitting algorithms, which require activating all the proximity operators at each iteration, inefficient or simply impossible. As a result, we must turn our attention to algorithms which employ, at each iteration $n$, only blocks of functions $(f_i)_{i\in I_n}$ and $(g_k)_{k\in K_n}$. If the functions $(g_k)_{1\leqslant k\leqslant p}$ were all smooth, one could use block-activated versions of the forward-backward algorithm proposed in [16, 25] and the references therein; in particular, when $m = 1$, methods such as those of [12, 18, 23, 26] would be pertinent. Next, as noted in [16, Remark 5.10(iv)], another candidate of interest could be the randomly block-activated algorithm of [16, Section 5.2], which leads to block-activated versions of several primal-dual methods (see [24] for detailed developments and [8] for an inertial version when $m = 1$). However, this approach violates requirement ④ because it imposes bounds on the proximal scaling parameters which depend on the norms of the linear operators. Finally, requirement ⑤ rules out methods that guarantee merely minimizing sequences or ergodic convergence.

To the best of our knowledge, there seems to be two primary methods that fulfill ①–⑤:

- Algorithm 5.2.6: The stochastic primal-dual Douglas–Rachford algorithm of [16].
- Algorithm 5.2.8: The deterministic primal-dual projective splitting algorithm of [10].

In the case of smooth coupling functions $(g_k)_{1\leqslant k\leqslant p}$, extensive numerical experience has been accumulated to understand the behavior of block-activated methods, especially in the case of stochastic gradient methods. By contrast, to date, very few numerical experiments with the recent, fully nonsmooth Algorithms 5.2.6 and 5.2.8 have been conducted and no comparison of their merits and performance has been undertaken. Thus far, Algorithm 5.2.6 has been employed only in the context of machine learning (see also the variant for partially smooth problems proposed in [6]). On the other hand, Algorithm 5.2.8 has been used in image recovery in [11], but only in full activation mode, and in rare feature selection in [22], but with $m = 1$.

**Objectives:** This paper aims at filling the above gap by shedding light on the implementation, the features, and the behavior of the fully nonsmooth Algorithms 5.2.6 and 5.2.8, comparing their merits, and providing numerical experiments illustrating their performance.

**Contributions and outline:** In Section 5.2.2, we illustrate the pertinence and the versatility of the model proposed in Problem 5.2.1 through a panel of examples drawn from various fields. Algorithms 5.2.6 and 5.2.8 are presented in Section 5.2.3, where we analyze and compare their features, implementation, and asymptotic properties. This investigation is complemented in Section 5.2.4 by numerical experiments in the context of machine learning and image recovery.

### 5.2.2 Instantiations of Problem 5.2.1

We illustrate the pertinence and the versatility of the proposed model through a few examples.

**Example 5.2.2** Variational models in multispectral imaging naturally involve minimization over several components. Specific references are [4, 7].

**Example 5.2.3** In perspective maximum-likelihood type estimation, the goal is to estimate scale vectors $s = (\sigma_i)_{1\leqslant i\leqslant N}$ and $t = (\tau_i)_{1\leqslant i\leqslant P}$, as well as a regression vector $b \in \mathbb{R}^d$ [14]. The minimization problem assumes the form

$$\underset{s\in\mathbb{R}^N,\, t\in\mathbb{R}^P,\, b\in\mathbb{R}^d}{\operatorname{minimize}} \quad \varsigma(s) + \varpi(t) + \theta(b) + \sum_{i=1}^{N} \Phi_i\big(\sigma_i, X_i b\big) + \sum_{i=1}^{P} \Psi_i\big(\tau_i, L_i b\big), \tag{5.2}$$

where all the functions are convex, $(X_1, \ldots, X_N)$ are design matrices, and $(L_1, \ldots, L_P)$ are linear transformations.

**Example 5.2.4** We consider the latent group lasso formulation in machine learning [21]. Let $\{p_1, \ldots, p_m\} \subset [1, +\infty]$, let $\{G_1, \ldots, G_m\}$ be a covering of $\{1, \ldots, d\}$, and define

$$X = \big\{(x_1, \ldots, x_m) \;\big|\; x_i \in \mathbb{R}^d,\ \operatorname{support}(x_i) \subset G_i\big\}. \tag{5.3}$$

The solution is $\widetilde{y} = \sum_{i=1}^m \widetilde{x}_i$, where $(\widetilde{x}_1, \ldots, \widetilde{x}_m)$ solves

$$\underset{(x_1,\ldots,x_m)\in X}{\operatorname{minimize}} \quad \sum_{i=1}^{m} \tau_i \|x_i\|_{p_i} + \sum_{k=1}^{p} g_k\bigg(\sum_{i=1}^{m} \langle x_i \mid u_k\rangle\bigg), \tag{5.4}$$

with $\tau_i \in \left]0, +\infty\right[$, $u_k \in \mathbb{R}^d$, and $g_k \colon \mathbb{R} \to \left]-\infty, +\infty\right]$ convex.

**Example 5.2.5** Various signal recovery problems can be modeled as infimal convolution problems of the form

$$\underset{x\in\mathcal{H}}{\operatorname{minimize}} \quad f(x) + \sum_{k=1}^{p} (f_k \,\square\, g_k)(L_k x), \tag{5.5}$$

where all the functions are convex, $L_k \colon \mathcal{H} \to \mathcal{G}_k$ is linear, and $\square$ is the inf-convolution operation, e.g., [2, 11, 20]. Under mild conditions, (5.5) can be rephrased as

$$\underset{\substack{x\in\mathcal{H}\\ y_1\in\mathcal{G}_1,\ldots,y_p\in\mathcal{G}_p}}{\operatorname{minimize}} \quad f(x) + \sum_{k=1}^{p} f_k(y_k) + \sum_{k=1}^{p} g_k(L_k x - y_k). \tag{5.6}$$

### 5.2.3 Algorithms: presentation and discussion

The subdifferential, the conjugate, and the proximity operator of a proper lower semicontinuous convex function $f \colon \mathcal{H} \to \left]-\infty, +\infty\right]$ are denoted by $\partial f$, $f^*$, and $\operatorname{prox}_f$, respectively. Let us

consider the setting of Problem 5.2.1 and let us set $\boldsymbol{\mathcal{H}} = \mathcal{H}_1 \times \cdots \times \mathcal{H}_m$ and $\boldsymbol{\mathcal{G}} = \mathcal{G}_1 \times \cdots \times \mathcal{G}_p$. A generic element in $\boldsymbol{\mathcal{H}}$ is denoted by $\boldsymbol{x} = (x_i)_{1\leqslant i\leqslant m}$. We make the standing assumption that the Kuhn–Tucker set of Problem 5.2.1 is nonempty, that is, there exist $\widetilde{\boldsymbol{x}} \in \boldsymbol{\mathcal{H}}$ and $\widetilde{\boldsymbol{v}}^* \in \boldsymbol{\mathcal{G}}$ such that

$$\begin{cases} \big(\forall i \in \{1,\ldots,m\}\big) & -\sum_{k=1}^{p} L_{k,i}^* \widetilde{v}_k^* \in \partial f_i(\widetilde{x}_i) \\ \big(\forall k \in \{1,\ldots,p\}\big) & \sum_{i=1}^{m} L_{k,i}\widetilde{x}_i \in \partial g_k^*(\widetilde{v}_k^*). \end{cases} \tag{5.7}$$

This implies that the solution set $\mathscr{P}$ of Problem 5.2.1 is nonempty.

As discussed in Section 5.2.1, two primary algorithms seem to fulfill requirements ①–⑤. The first algorithm operates in the product space $\boldsymbol{\mathcal{H}} \times \boldsymbol{\mathcal{G}}$ and employs random activation of the blocks. To present it, let us introduce

$$\boldsymbol{L}\colon \boldsymbol{\mathcal{H}} \to \boldsymbol{\mathcal{G}}\colon \boldsymbol{x} \mapsto \left( \sum_{i=1}^{m} L_{1,i}x_i, \ldots, \sum_{i=1}^{m} L_{p,i}x_i \right) \tag{5.8}$$

and

$$\boldsymbol{V} = \big\{(\boldsymbol{z},\boldsymbol{y}) \in \boldsymbol{\mathcal{H}} \times \boldsymbol{\mathcal{G}} \;\big|\; \boldsymbol{y} = \boldsymbol{L}\boldsymbol{z}\big\}. \tag{5.9}$$

Let $\boldsymbol{z} \in \boldsymbol{\mathcal{H}}$ and $\boldsymbol{y} \in \boldsymbol{\mathcal{G}}$, and set $\boldsymbol{t} = (\mathbf{Id} + \boldsymbol{L}^*\boldsymbol{L})^{-1}(\boldsymbol{z} + \boldsymbol{L}^*\boldsymbol{y})$ and $\boldsymbol{s} = (\mathbf{Id} + \boldsymbol{L}\boldsymbol{L}^*)^{-1}(\boldsymbol{L}\boldsymbol{z} - \boldsymbol{y})$. Then the projection of $(\boldsymbol{z},\boldsymbol{y}) \in \boldsymbol{\mathcal{H}} \times \boldsymbol{\mathcal{G}}$ onto $\boldsymbol{V}$ is [16, Eq. (5.25)]

$$\operatorname{proj}_{\boldsymbol{V}}(\boldsymbol{z},\boldsymbol{y}) = (\boldsymbol{t},\boldsymbol{L}\boldsymbol{t}) = (\boldsymbol{z} - \boldsymbol{L}^*\boldsymbol{s}, \boldsymbol{y} + \boldsymbol{s}). \tag{5.10}$$

The coordinate operators of $\operatorname{proj}_{\boldsymbol{V}}$ are $(Q_j)_{1\leqslant j\leqslant m+p}$, i.e.,

$$\operatorname{proj}_{\boldsymbol{V}}(\boldsymbol{z},\boldsymbol{y}) = \big(Q_1(\boldsymbol{z},\boldsymbol{y}), \ldots, Q_{m+p}(\boldsymbol{z},\boldsymbol{y})\big). \tag{5.11}$$

**Algorithm 5.2.6 ( [16])** Let $\gamma \in \left]0, +\infty\right[$, let $\boldsymbol{x}_0$ and $\boldsymbol{z}_0$ be $\boldsymbol{\mathcal{H}}$-valued random variables (r.v.), let

$\boldsymbol{y}_0$ and $\boldsymbol{w}_0$ be $\boldsymbol{\mathcal{G}}$-valued r.v. Iterate

$$
\begin{array}{l}
\text{for } j = 1, \dots, m+p \\
\lfloor \text{compute } Q_j \text{ given by (5.8)–(5.11)} \\
\text{for } n = 0, 1, \dots \\
\quad \lambda_n \in \,]0, 2[ \\
\quad \text{select randomly } \varnothing \neq I_n \subset \{1, \dots, m\} \text{ and } \varnothing \neq K_n \subset \{1, \dots, p\} \\
\quad \text{for every } i \in I_n \\
\quad\quad x_{i,n+1} = Q_i(\boldsymbol{z}_n, \boldsymbol{y}_n) \\
\quad\quad z_{i,n+1} = z_{i,n} + \lambda_n \big(\operatorname{prox}_{\gamma f_i}(2x_{i,n+1} - z_{i,n}) - x_{i,n+1}\big) \\
\quad \text{for every } i \in \{1, \dots, m\} \smallsetminus I_n \\
\quad\quad (x_{i,n+1}, z_{i,n+1}) = (x_{i,n}, z_{i,n}) \\
\quad \text{for every } k \in K_n \\
\quad\quad w_{k,n+1} = Q_{m+k}(\boldsymbol{z}_n, \boldsymbol{y}_n) \\
\quad\quad y_{k,n+1} = y_{k,n} + \lambda_n \big(\operatorname{prox}_{\gamma g_k}(2w_{k,n+1} - y_{k,n}) - w_{k,n+1}\big) \\
\quad \text{for every } k \in \{1, \dots, p\} \smallsetminus K_n \\
\quad\quad (w_{k,n+1}, y_{k,n+1}) = (w_{k,n}, y_{k,n}).
\end{array}
\tag{5.12}
$$

**Theorem 5.2.7 ( [16])** *In the setting of Algorithm 5.2.6, define, for every $n \in \mathbb{N}$ and every $j \in \{1, \dots, m+p\}$,*

$$
\varepsilon_{j,n} = \begin{cases} 1, & \text{if } j \in I_n \text{ or } j - m \in K_n; \\ 0, & \text{otherwise}. \end{cases}
\tag{5.13}
$$

*Suppose that the following hold:*

[a] $\inf_{n\in\mathbb{N}} \lambda_n > 0$ *and* $\sup_{n\in\mathbb{N}} \lambda_n < 2$.

[b] *The r.v.* $(\boldsymbol{\varepsilon}_n)_{n\in\mathbb{N}}$ *are identically distributed.*

[c] *For every* $n \in \mathbb{N}$, *the r.v.* $\boldsymbol{\varepsilon}_n$ *and* $(\boldsymbol{z}_j, \boldsymbol{y}_j)_{0\leqslant j\leqslant n}$ *are mutually independent.*

[d] $(\forall j \in \{1, \dots, m+p\})$ $\mathsf{Prob}[\varepsilon_{j,0} = 1] > 0$.

*Then* $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *converges almost surely to a* $\mathscr{P}$*-valued r.v.*

The second algorithm operates by projecting onto hyperplanes which separate the current iterate from the Kuhn–Tucker set of Problem 5.2.1 and activating the blocks in a deterministic manner [10].

**Algorithm 5.2.8 ( [10])** Set $I_0 = \{1, \dots, m\}$ and $K_0 = \{1, \dots, p\}$. For every $i \in \{1, \dots, m\}$ and every $k \in \{1, \dots, p\}$, let $\{\gamma_i, \mu_k\} \subset \,]0, +\infty[$, $x_{i,0} \in \mathcal{H}_i$, and $v^*_{k,0} \in \mathcal{G}_k$. Iterate

for $n = 0, 1, \ldots$

- $\lambda_n \in \,]0, 2[$
- if $n > 0$
  - select $\varnothing \neq I_n \subset \{1, \ldots, m\}$ and $\varnothing \neq K_n \subset \{1, \ldots, p\}$
- for every $i \in I_n$
  - $x^*_{i,n} = x_{i,n} - \gamma_i \sum_{k=1}^{p} L^*_{k,i} v^*_{k,n}$
  - $a_{i,n} = \operatorname{prox}_{\gamma_i f_i} x^*_{i,n}$
  - $a^*_{i,n} = \gamma_i^{-1}(x^*_{i,n} - a_{i,n})$
- for every $i \in \{1, \ldots, m\} \smallsetminus I_n$
  - $(a_{i,n}, a^*_{i,n}) = (a_{i,n-1}, a^*_{i,n-1})$
- for every $k \in K_n$
  - $y^*_{k,n} = \mu_k v^*_{k,n} + \sum_{i=1}^{m} L_{k,i} x_{i,n}$
  - $b_{k,n} = \operatorname{prox}_{\mu_k g_k} y^*_{k,n}$
  - $b^*_{k,n} = \mu_k^{-1}(y^*_{k,n} - b_{k,n})$
  - $t_{k,n} = b_{k,n} - \sum_{i=1}^{m} L_{k,i} a_{i,n}$
- for every $k \in \{1, \ldots, p\} \smallsetminus K_n$
  - $(b_{k,n}, b^*_{k,n}) = (b_{k,n-1}, b^*_{k,n-1})$
  - $t_{k,n} = b_{k,n} - \sum_{i=1}^{m} L_{k,i} a_{i,n}$
- for every $i \in \{1, \ldots, m\}$
  - $t^*_{i,n} = a^*_{i,n} + \sum_{k=1}^{p} L^*_{k,i} b^*_{k,n}$
- $\tau_n = \sum_{i=1}^{m} \|t^*_{i,n}\|^2 + \sum_{k=1}^{p} \|t_{k,n}\|^2$
- if $\tau_n > 0$
  - $\pi_n = \sum_{i=1}^{m} \big(\langle x_{i,n} \mid t^*_{i,n}\rangle - \langle a_{i,n} \mid a^*_{i,n}\rangle\big) + \sum_{k=1}^{p} \big(\langle t_{k,n} \mid v^*_{k,n}\rangle - \langle b_{k,n} \mid b^*_{k,n}\rangle\big)$
- if $\tau_n > 0$ and $\pi_n > 0$
  - $\theta_n = \lambda_n \pi_n / \tau_n$
  - for every $i \in \{1, \ldots, m\}$
    - $x_{i,n+1} = x_{i,n} - \theta_n t^*_{i,n}$
  - for every $k \in \{1, \ldots, p\}$
    - $v^*_{k,n+1} = v^*_{k,n} - \theta_n t_{k,n}$
- else
  - for every $i \in \{1, \ldots, m\}$
    - $x_{i,n+1} = x_{i,n}$
  - for every $k \in \{1, \ldots, p\}$
    - $v^*_{k,n+1} = v^*_{k,n}$.

(5.14)

**Theorem 5.2.9 ( [10])** *In the setting of Algorithm 5.2.8, suppose that the following hold:*

[a] $\inf_{n\in\mathbb{N}} \lambda_n > 0$ *and* $\sup_{n\in\mathbb{N}} \lambda_n < 2$.

[b] *There exists* $T \in \mathbb{N}$ *such that, for every* $n \in \mathbb{N}$, $\bigcup_{j=n}^{n+T} I_j = \{1, \ldots, m\}$ *and* $\bigcup_{j=n}^{n+T} K_j =$

$\{1, \dots, p\}$.

*Then $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ converges to a point in $\mathscr{P}$.*

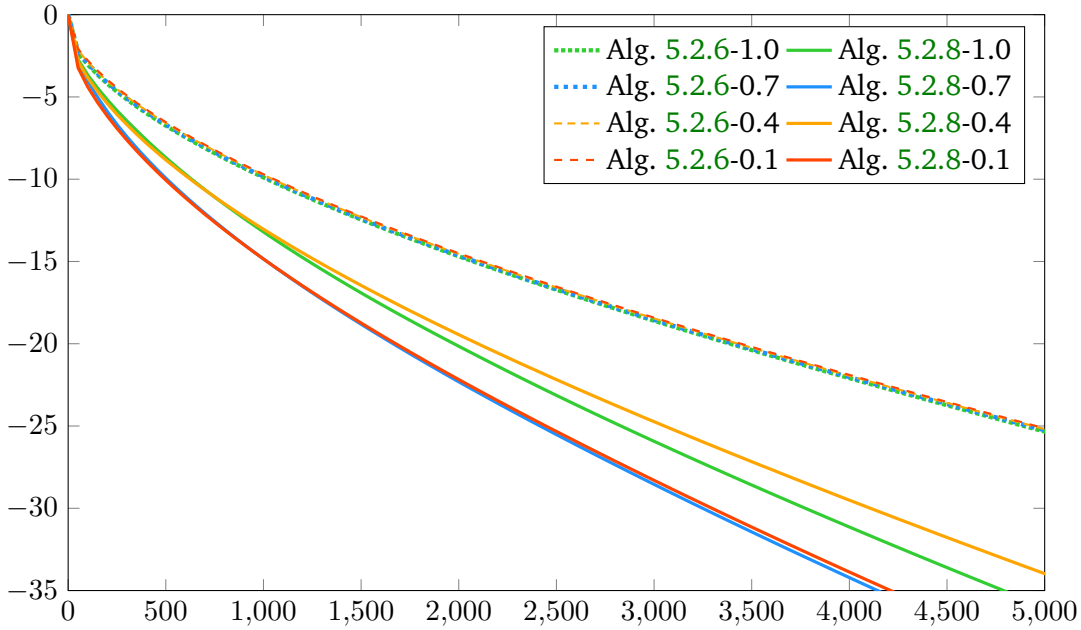



**Figure 5.1** Normalized error $20\log_{10}(\|\boldsymbol{x}_n - \boldsymbol{x}_\infty\|/\|\boldsymbol{x}_0 - \boldsymbol{x}_\infty\|)$ (dB), averaged over 20 runs, versus epoch count in Experiment 1. The variations around the averages were not significant. The computational load per epoch for both algorithms is comparable.

**Remark 5.2.10 (comparing Algorithms 5.2.6 and 5.2.8)**

(i) **Auxiliary tasks:** Algorithm 5.2.6 requires the construction and storage of the operators $(Q_j)_{1\leqslant j\leqslant m+p}$ of (5.10)–(5.11), which can be quite demanding as they involve inversion of a linear operator acting on the product space $\boldsymbol{\mathcal{H}}$ or $\boldsymbol{\mathcal{G}}$. By contrast, Algorithm 5.2.8 does not require such tasks.

(ii) **Proximity operators:** In both algorithms, only the proximity operators of the blocks of functions $(f_i)_{i\in I_n}$ and $(g_k)_{k\in K_n}$ need to be activated at iteration $n$.

(iii) **Linear operators:** In Algorithm 5.2.6, the operators $(Q_i)_{i\in I_n}$ and $(Q_{m+k})_{k\in K_n}$ selected at iteration $n$ are evaluated at $(z_{1,n}, \dots, z_{m,n}, y_{1,n}, \dots, y_{p,n}) \in \boldsymbol{\mathcal{H}} \times \boldsymbol{\mathcal{G}}$. On the other hand, Algorithm 5.2.8 activates the local operators $L_{k,i}\colon \mathcal{H}_i \to \mathcal{G}_k$ and $L^*_{k,i}\colon \mathcal{G}_k \to \mathcal{H}_i$ once or twice, depending on whether they are selected. For instance, if we set $N = \dim \boldsymbol{\mathcal{H}}$ and $M = \dim \boldsymbol{\mathcal{G}}$ and if all the linear operators are implemented in matrix form, then the corresponding load per iteration in full activation mode of Algorithm 5.2.6 is $\mathcal{O}((M+N)^2)$ versus $\mathcal{O}(MN)$ in Algorithm 5.2.8.

(iv) **Activation scheme:** As Algorithm 5.2.6 selects the blocks randomly, the user does not have complete control of the computational load of an iteration, whereas that of Algorithm 5.2.8 is more predictable because of its deterministic activation scheme.

(v) **Parameters:** A single scale parameter $\gamma$ is used in Algorithm 5.2.6, while Algorithm 5.2.8 allows the proximity operators to have their own scale parameters $(\gamma_1, \ldots, \gamma_m, \mu_1, \ldots, \mu_p)$. This gives Algorithm 5.2.8 more flexibility, but more effort may be needed to find efficient parameters. Furthermore, in both algorithms, there is no restriction on the parameter values.

(vi) **Convergence:** Algorithm 5.2.8 guarantees sure convergence under the mild sweeping condition [b] in Theorem 5.2.9, while 5.2.6 guarantees only almost sure convergence.

(vii) **Other features:** Although this point is omitted for brevity, unlike Algorithm 5.2.6, Algorithm 5.2.8 can be executed asynchronously with iteration-dependent scale parameters [10].

### 5.2.4 Numerical experiments

We present two experiments which are reflective of our numerical investigations in solving various problems using Algorithms 5.2.6 and 5.2.8.

#### 5.2.4.1 Experiment 1: group-sparse binary classification

We revisit the problem from [13], which is set as Example 5.2.4 with $g_k \colon \xi \mapsto \max\{0, 1 - \beta_k \xi\}$, where $\beta_k = \omega_k \operatorname{sign}(\langle \overline{y} \mid u_k \rangle)$ is the $k$th measurement of the true vector $\overline{y} \in \mathbb{R}^d$ ($d = 10000$) and $\omega_k \in \{-1, 1\}$ induces $25\%$ classification error. There are $p = 1000$ measurements and the goal is to reconstruct the group-sparse vector $\overline{y}$. There are $m = 1429$ groups. For every $i \in \{1, \ldots, m-1\}$, each $G_i$ has $10$ consecutive integers and an overlap with $G_{i+1}$ of length $3$. We obtain an instance of (5.1), where $\mathcal{H}_i = \mathbb{R}^{10}$, $f_i = 0.1\|\cdot\|_2$, and $L_{k,i} = \langle \cdot \mid u_k|_{G_i} \rangle$. The auxiliary tasks for Algorithm 5.2.6 (see Remark 5.2.10(i)) are negligible [13]. For each $\alpha \in \{0.1, 0.4, 0.7, 1.0\}$, at iteration $n \in \mathbb{N}$, $I_n$ has $\lceil \alpha m \rceil$ elements and the proximity operators of the scalar functions $(g_k)_{1 \leqslant k \leqslant p}$ are all used, i.e., $K_n = \{1, \ldots, p\}$. We display in Fig. 5.1 the normalized error versus the epoch, that is, the cumulative number of activated blocks in $\{1, \ldots, m\}$ divided by $m$.

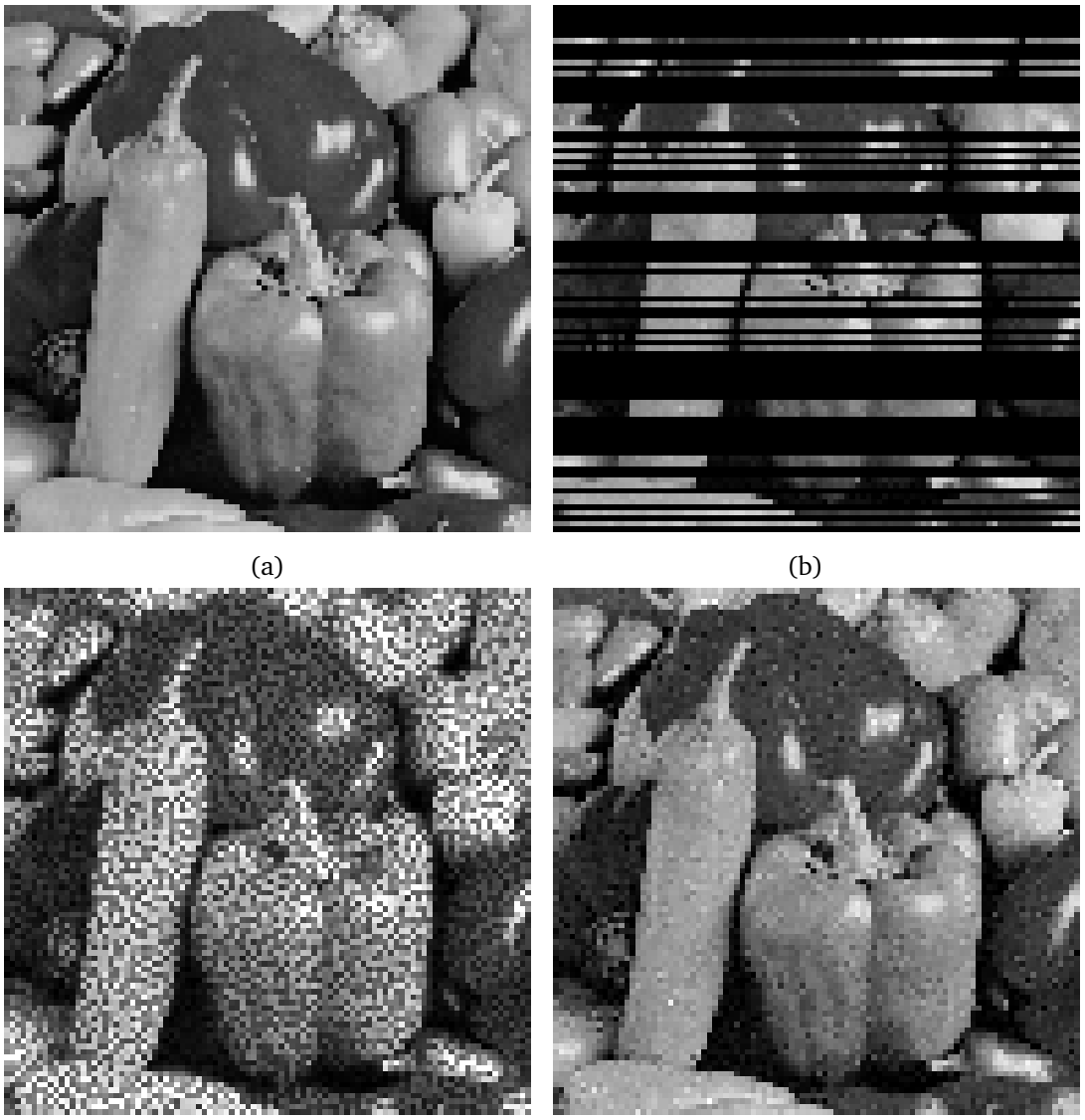
#### 5.2.4.2 Experiment 2: image recovery

We revisit the image interpolation problem from [11, Section 4.3]. The objective is to recover the image $\overline{x} \in C = [0, 255]^N$ ($N = 96^2$) of Fig. 5.2(a), given a noisy masked observation $b = M\overline{x} + w_1 \in \mathbb{R}^N$ and a noisy blurred observation $c = H\overline{x} + w_2 \in \mathbb{R}^N$. Here, $M$ masks all but $q = 39$ rows $(x^{(r_k)})_{1 \leqslant k \leqslant q}$ of an image $x$, and $H$ is a nonstationary blurring operator, while $w_1$ and $w_2$ yield signal-to-noise ratios of $28.5$ dB and $27.8$ dB, respectively. Since $H$ is sizable,

we split it into $s = 384$ subblocks: for every $k \in \{1, \dots, s\}$, $H_k \in \mathbb{R}^{24\times N}$ and the corresponding block of $c$ is denoted $c_k$. The goal is to

$$\underset{x\in C}{\text{minimize}} \ \|Dx\|_{1,2} + 10\sum_{k=1}^{q} \left\|x^{(r_k)} - b^{(r_k)}\right\|_2 + 5\sum_{k=1}^{s} \|H_k x - c_k\|_2^2, \tag{5.15}$$

where $D\colon \mathbb{R}^N \to \mathbb{R}^N \times \mathbb{R}^N$ models finite differences and $\|\cdot\|_{1,2}\colon (y_1, y_2) \mapsto \sum_{j=1}^{N} \|(\eta_{1,j}, \eta_{2,j})\|_2$. Thus, (5.15) is an instance of Problem 5.2.1, where $m = 1$; $p = q+s+1$; for every $k \in \{1, \dots, q\}$, $L_{k,1}\colon \mathbb{R}^N \to \mathbb{R}^{\sqrt{N}}\colon x \mapsto x^{(r_k)}$ and $g_k\colon y_k \mapsto 10\|y_k - b^{(r_k)}\|_2$; for every $k \in \{q+1, \dots, q+s\}$, $L_{k,1} = H_{k-q}$, $g_k = 5\|\cdot - c_k\|_2^2$, and $g_p = \|\cdot\|_{1,2}$; $L_{p,1} = D$; $f_1\colon x \mapsto 0$ if $x \in C$; $+\infty$ if $x \notin C$. At iteration $n$, $K_n$ has $\lceil \alpha p \rceil$ elements, where $\alpha \in \{0.1, 0.4, 0.7, 1.0\}$. The results are shown in Figs. 5.2–5.3, where the epoch is the cumulative number of activated blocks in $\{1, \dots, p\}$ divided by $p$.



(a) (b)

(c) (d)

**Figure 5.2** Experiment 2: (a) Original $\overline{x}$. (b) Observation $b$. (c) Observation $c$. (d) Recovery (all recoveries were visually indistinguishable).

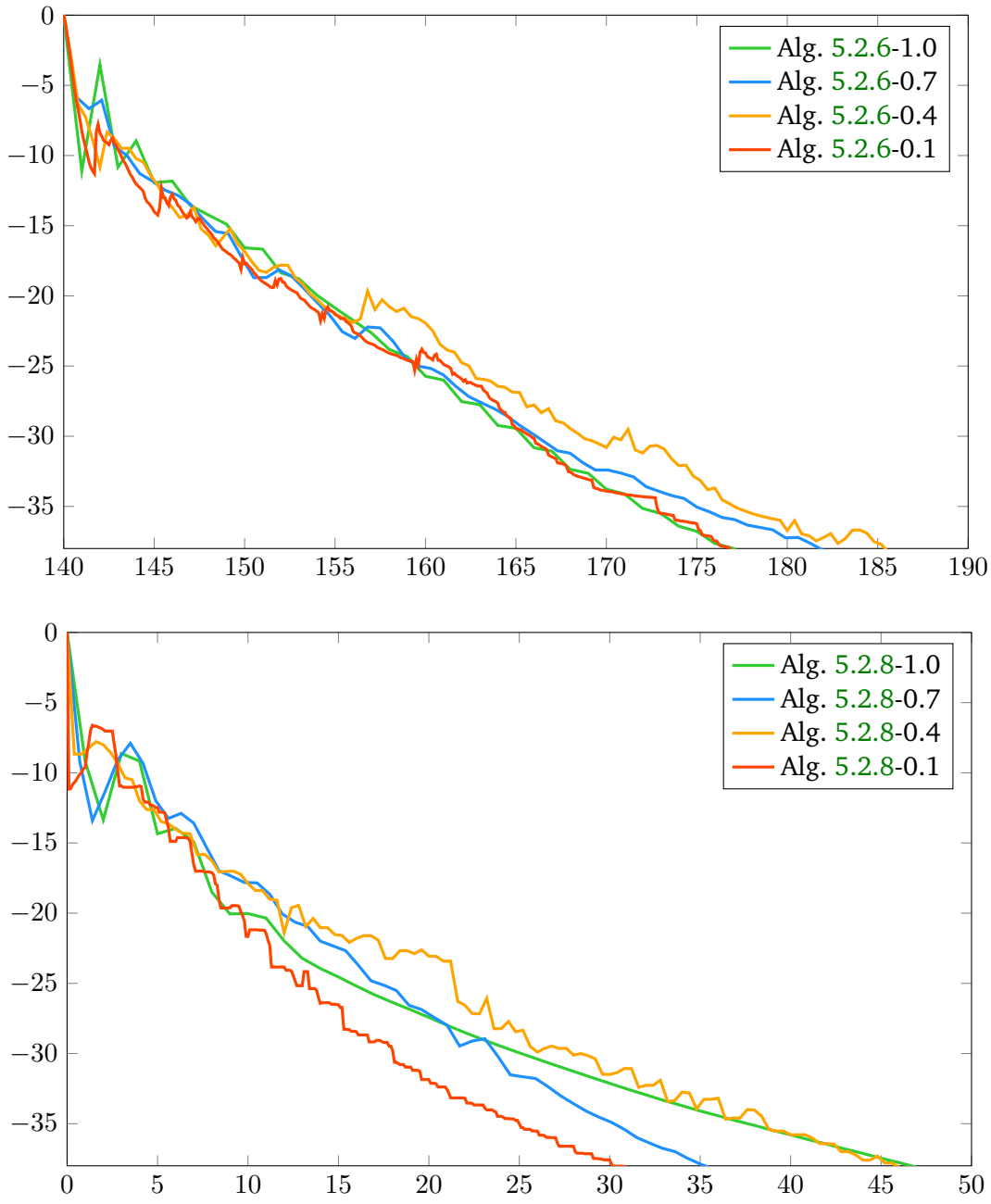



**Figure 5.3** Normalized error $20\log_{10}(\|\boldsymbol{x}_n - \boldsymbol{x}_\infty\|/\|\boldsymbol{x}_0 - \boldsymbol{x}_\infty\|)$ (dB) versus epoch count in Experiment 2. Top: Algorithm 5.2.6. The horizontal axis starts at 140 epochs to account for the auxiliary tasks (see Remark 5.2.10(i)). Bottom: Algorithm 5.2.8. The computational load per epoch for Algorithm 5.2.8 was about twice that of Algorithm 5.2.6.

#### 5.2.4.3 Discussion

Our first finding is that, for both Algorithms 5.2.6 and 5.2.8, even when full activation is possible, it may not be the best strategy (see Figs. 5.1 and 5.3). Second, Remark 5.2.10 and our experiments strongly suggest that Algorithm 5.2.8 may be preferable to 5.2.6. Let us add

that, in general, Algorithm 5.2.6 does not scale as well as 5.2.8. For instance, in Experiment 2, if the image size scales up, Algorithm 5.2.8 can still operate since it involves only individual applications of the local $L_{k,i}$ operators, while Algorithm 5.2.6 becomes unmanageable because of the size of the $Q_j$ operators (see Remark 5.2.10(i) and [6]).

Per Remark 5.2.10(vii), we are currently exploring the numerical benefits of implementing Algorithm 5.2.8 asynchronously.

# References


[1] A. Argyriou, R. Foygel, and N. Srebro, Sparse prediction with the $k$-support norm, *Proc. Adv. Neural Inform. Process. Syst. Conf.*, vol. 25, pp. 1457–1465, 2012.

[2] J.-F. Aujol and A. Chambolle, Dual norms and image decomposition models, *Int. J. Comput. Vision*, vol. 63, pp. 85–104, 2005.

[3] F. Bach, R. Jenatton, J. Mairal, and G. Obozinski, Optimization with sparsity-inducing penalties, *Found. Trends Machine Learn.*, vol. 4, pp. 1–106, 2012.

[4] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches, *IEEE J. Select. Topics Appl. Earth Observ. Remote Sensing*, vol. 5, pp. 354–379, 2012.

[5] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, *Found. Trends Machine Learn.*, vol. 3, pp. 1–122, 2010.

[6] L. M. Briceño-Arias, G. Chierchia, E. Chouzenoux, and J.-C. Pesquet, A random block-coordinate Douglas-Rachford splitting method with low computational complexity for binary logistic regression, *Comput. Optim. Appl.*, vol. 72, pp. 707–726, 2019.

[7] 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.

[8] A. Chambolle, M. J. Ehrhardt, P. Richtárik, and C.-B. Schönlieb, Stochastic primal-dual hybrid gradient algorithm with arbitrary sampling and imaging applications, *SIAM J. Optim.*, vol. 28, pp. 2783–2808, 2018.

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

[10] P. L. Combettes and J. Eckstein, Asynchronous block-iterative primal-dual decomposition methods for monotone inclusions, *Math. Program.,* vol. B168, pp. 645–672, 2018.

[11] 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.

[12] P. L. Combettes and L. E. Glaudin, Solving composite fixed point problems with block updates, *Adv. Nonlinear Anal.*, vol. 10, 2021.

[13] P. L. Combettes, A. M. McDonald, C. A. Micchelli, and M. Pontil, Learning with optimal interpolation norms, *Numer. Algorithms*, vol. 81, pp. 695–717, 2019.

[14] P. L. Combettes and C. L. Müller, Perspective maximum likelihood-type estimation via proximal decomposition, *Electron. J. Stat.*, vol. 14, pp. 207–238, 2020.

[15] P. L. Combettes and J.-C. Pesquet, Proximal splitting methods in signal processing, *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, pp. 185–212. Springer, 2011.

[16] P. L. Combettes and J.-C. Pesquet, Stochastic quasi-Fejér block-coordinate fixed point iterations with random sweeping, *SIAM J. Optim.*, vol. 25, pp. 1221–1248, 2015.

[17] J. Darbon and T. Meng, On decomposition models in imaging sciences and multi-time Hamilton–Jacobi partial differential equations, *SIAM J. Imaging Sci.*, vol. 13, pp. 971–1014, 2020.

[18] A. J. Defazio, T. S. Caetano, and J. Domke, Finito: A faster, permutable incremental gradient method for big data problems, *Proc. Intl. Conf. Machine Learn.*, pp. 1125–1133, 2014.

[19] R. Glowinski, S. J. Osher, and W. Yin (Eds.), *Splitting Methods in Communication, Imaging, Science, and Engineering*. Springer, 2016.

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

[21] L. Jacob, G. Obozinski, and J.-Ph. Vert, Group lasso with overlap and graph lasso, *Proc. Int. Conf. Machine Learn.*, pp. 433–440, 2009.

[22] P. R. Johnstone and J. Eckstein, Projective splitting with forward steps, *Math. Program.*, published online 2020-09-30.

[23] K. Mishchenko, F. Iutzeler, and J. Malick, A distributed flexible delay-tolerant proximal gradient algorithm, *SIAM J. Optim.*, vol. 30, pp. 933–959, 2020.

[24] J.-C. Pesquet and A. Repetti, A class of randomized primal-dual algorithms for distributed optimization, *J. Nonlinear Convex Anal.*, vol. 16, pp. 2453–2490, 2015.

[25] S. Salzo and S. Villa, Parallel random block-coordinate forward-backward algorithm: A unified convergence analysis, *Math. Program.*, to appear.

[26] M. Schmidt, N. Le Roux, and F. Bach, Minimizing finite sums with the stochastic average gradient, *Math. Program.*, vol. 162, pp. 83–112, 2017.

Chapter 6

# CONCLUSION

## 6.1 Summary

We have proposed a novel, flexible framework for enforcing firmly nonexpansive nonlinear equations in a variety of problems. This framework is shown to have utility in a broad range of pertinent applications. Even for nonlinear equations which are not in this class, many can be equivalently represented using firmly nonexpansive operators. These nonlinear equations are incorporated into feasibility problems, best approximation problems, and inconsistent feasibility problems. We have also developed a new best approximation algorithm, which covers a wider class of extrapolation strategies than previously available in the literature; furthermore, it is demonstrated that this strategy can yield significantly improved numerical performance. Also, for the first time, we identified, analyzed, and compared block-iterative algorithms designed for large-scale fully nonsmooth multicomponent convex minimization problems.

## 6.2 Future work

Future research directions suggested by this work are the following.

- The work in Chapter 3 produces the best approximation with respect to the Hilbertian norm on $\mathcal{H}$. A natural extension of this work would be to compute best approximations with respect to other norms, or Bregman distances.
- While Chapter 3 solves best approximation problems, developing a strategy for more general optimization tasks subject to nonlinear equations remains to be done.
- While the numerical experiments in Chapter 4 demonstrate that certain formulations of Problem 1.1.5 promote sparse recoveries, these formulations are purely based on heuristics. An extension of this work would be to investigate analytical guarantees concerning when a recovery will be sparse.

- Chapter 5 can be extended beyond the minimization setting to the more general setting of monotone inclusions. Furthermore, recent asynchronous algorithms in this field are yet to be thoroughly investigated.

- Consider a prescribed point $p \in \mathcal{H}$ of a firmly nonexpansive operator $F\colon \mathcal{H} \to \mathcal{H}$, and set $C = \big\{x \in \mathcal{H} \;\big|\; Fx = p\big\}$. A simple, yet fundamental observation in this work relies on the fact that, in order to enforce that $x \in C$ algorithmically, it suffices to have the ability to efficiently evaluate at least one operator from the class

$$\mathscr{C} = \{T\colon \mathcal{H} \to \mathcal{H} \;\mid T \;\text{ is firmly nonexpansive and }\; \operatorname{Fix} T = C\}\,. \tag{6.1}$$

  While traditional approaches insist on access to $\operatorname{proj}_C \in \mathscr{C}$, which is often costly to evaluate, we have identified that the operator $\operatorname{Id} - F + p$ resides in $\mathscr{C}$ and only relies on computation of the forward operator $F$. While this leads to a new avenue for the development of provenly-convergent, tractable algorithms for resolving such nonlinear equations, the relative behavior of operators in $\mathscr{C}$ has not been analyzed or evaluated numerically. For instance, given two operators in $\mathscr{C}$ is it possible to analytically characterize their relative performance in a given algorithm? In view of Chapter 4, is it possible to relate how selection of an operator in $\mathscr{C}$ may affect the solutions to the relaxed formulation Problem 1.1.5?

Raleigh, May 17th, 2021