# Reconstruction of Functions from Prescribed Proximal Points*

Patrick L. Combettes and Zev C. Woodstock

North Carolina State University, Department of Mathematics, Raleigh, NC 27695-8205, USA
plc@math.ncsu.edu, zwoodst@ncsu.edu

*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.

**Keywords**. Best approximation algorithm, constrained interpolation, firmly nonexpansive operator, nonlinear signal recovery, proximal point.

*Contact author: P. L. Combettes, plc@math.ncsu.edu, phone: +1 (919) 515 2671. The work of P. L. Combettes was supported by the National Science Foundation under grant CCF-1715671 and the work of Z. C. Woodstock was supported by the National Science Foundation under grant DGE-1746939.

## 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 $\text{proj}_U$ and $\text{proj}_V$, respectively, and let $p \in V$. The basic best approximation problem

$$\text{minimize} \ \|x - x_0\| \quad \text{subject to} \quad x \in U \quad \text{and} \quad \text{proj}_V x = p \tag{1.1}$$

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 (1.1) 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 (1.1) 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 (1.1) is feasible (see [22] for necessary and sufficient conditions), then the sequence $(x_n)_{n\in\mathbb{N}}$ constructed by iterating

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = p + \text{proj}_U x_n - \text{proj}_V(\text{proj}_U x_n) \tag{1.2}$$

converges strongly to its solution. The extension of (1.1) to finitely many vector subspaces $(U_j)_{j\in J}$ and $(V_k)_{k\in K}$ investigated in [22] is to

$$\text{minimize} \ \|x - x_0\| \quad \text{subject to} \ x \in \bigcap_{j\in J} U_j \ \text{and} \ (\forall k \in K) \ \text{proj}_{V_k} x = p_k, \ \text{where} \ p_k \in V_k, \tag{1.3}$$

and it can be solved using affine projection methods. In many applications, the constraint sets [12, 13, 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 (1.3) 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.,

$$\text{minimize} \ \|x - x_0\| \quad \text{subject to} \quad x \in \bigcap_{j\in J} C_j \quad \text{and} \quad (\forall k \in K) \quad F_k x = p_k. \tag{1.4}$$

In view of (1.3), 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

$$y \mapsto f_k(y) + \frac{1}{2}\|\overline{x} - y\|^2, \tag{1.5}$$

and wrote $p_k = \text{prox}_{f_k}\overline{x}$. This mechanism defines the proximity operator $\text{prox}_{f_k}\colon \mathcal{H} \to \mathcal{H}$ 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

$$(\forall x \in \mathcal{H}) \quad \iota_{D_k}(x) = \begin{cases} 0, & \text{if} \ x \in D_k; \\ +\infty, & \text{if} \ x \notin D_k \end{cases} \tag{1.6}$$

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, $\operatorname{prox}_{f_k} = (\operatorname{Id} + \partial f_k)^{-1}$, where

$$(\forall x \in \mathcal{H}) \quad \partial f_k(x) = \big\{ u \in \mathcal{H} \;\big|\; (\forall y \in \mathcal{H}) \;\; \langle y - x \mid u \rangle + f_k(x) \leqslant f_k(y) \big\}. \tag{1.7}$$

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

$$(\forall x \in \mathcal{H})(\forall u \in \mathcal{H}) \quad \big[\, u \in A_k x \quad \Leftrightarrow \quad (\forall y \in \mathcal{H})(\forall v \in A_k y) \quad \langle x - y \mid u - v \rangle \geqslant 0 \,\big]. \tag{1.8}$$

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 $J_{A_k} = (\operatorname{Id} + A_k)^{-1} \colon \mathcal{H} \to \mathcal{H}$ 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

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \|F_k x - F_k y\|^2 + \|(\operatorname{Id} - F_k)x - (\operatorname{Id} - F_k)y\|^2 \leqslant \|x - y\|^2. \tag{1.9}$$

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 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 (1.4). In light of these considerations, we propose to investigate the following nonlinear best approximation framework.

**Problem 1.1** 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

$$\text{minimize } \|x - x_0\| \quad \text{subject to} \quad x \in \bigcap_{j \in J} C_j \quad \text{and} \quad (\forall k \in K) \quad F_k x = p_k. \tag{1.10}$$

In Problem 1.1, 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 (1.3) 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 1.1, and to devise an efficient and flexible numerical method to solve it.

The remainder of the paper consists of four sections. In Section 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, 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 4, Problem 1.1 is rephrased in terms of a common fixed point problem and the algorithm of Section 3 is used to solve it. A numerical illustration of our framework is presented in Section 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 $\left]-\infty,+\infty\right]$ 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 $\operatorname{Fix}T=\big\{x\in\mathcal{H}\;\big|\;Tx=x\big\}$. 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].

## 2 Prescribed values as proximal points

We illustrate the fact that the proximal model adopted in Problem 1.1 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 2.1 and cocoercive ones in Section 2.2. In Section 2.3, we move to more general models in which the operators need not be Lipschitzian or even continuous.

### 2.1 Prescriptions derived from firmly nonexpansive operators

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

**Proposition 2.1** *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

$$(\forall x\in\mathcal{H})\quad \frac{1}{3}\sum_{i\in\mathbb{I}}\|\mathsf{F}_i\mathsf{x}_i\|^2 \leqslant \sum_{i\in\mathbb{I}}\|\mathsf{F}_i\mathsf{x}_i-\mathsf{F}_i\mathsf{z}_i\|^2+\sum_{i\in\mathbb{I}}\|\mathsf{F}_i\mathsf{z}_i-\mathsf{z}_i\|^2+\sum_{i\in\mathbb{I}}\|\mathsf{z}_i\|^2$$
$$\leqslant \sum_{i\in\mathbb{I}}\|\mathsf{x}_i-\mathsf{z}_i\|^2+\sum_{i\in\mathbb{I}}\|\mathsf{F}_i\mathsf{z}_i-\mathsf{z}_i\|^2+\|z\|^2$$
$$=\|x-z\|^2+\sum_{i\in\mathbb{I}}\|\mathsf{F}_i\mathsf{z}_i-\mathsf{z}_i\|^2+\|z\|^2$$
$$<+\infty. \tag{2.1}$$

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

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \|Fx - Fy\|^2 = \sum_{i\in\mathbb{I}} \|\mathsf{F}_i\mathsf{x}_i - \mathsf{F}_i\mathsf{y}_i\|^2$$
$$\leqslant \sum_{i\in\mathbb{I}} \|\mathsf{x}_i - \mathsf{y}_i\|^2 - \sum_{i\in\mathbb{I}} \|(\mathsf{Id} - \mathsf{F}_i)\mathsf{x}_i - (\mathsf{Id} - \mathsf{F}_i)\mathsf{y}_i\|^2$$
$$= \|x - y\|^2 - \|(\mathrm{Id} - F)x - (\mathrm{Id} - F)y\|^2. \tag{2.2}$$

Thus, $F$ is firmly nonexpansive. ☐

**Corollary 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. 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 0 is a minimizer of each of the functions $(\mathsf{f}_i)_{i\in\mathbb{I}}$, we derive from [5, Proposition 12.29] that $(\forall i \in \mathbb{I})\ \mathrm{prox}_{\mathsf{f}_i}0 = 0$. In turn, the condition $\sum_{i\in\mathbb{I}} \|\mathsf{F}_i\mathsf{z}_i - \mathsf{z}_i\|^2 < +\infty$ holds with $(\forall i \in \mathbb{I})\ \mathsf{z}_i = 0$. In view of Proposition 2.1, $p$ 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

$$f(\mathrm{prox}_f\overline{x}) + \frac{1}{2}\|\overline{x} - \mathrm{prox}_f\overline{x}\|^2 = \min_{y\in\mathcal{H}} \left( f(y) + \frac{1}{2}\|\overline{x} - y\|^2 \right)$$
$$= \min_{y\in\mathcal{H}} \sum_{i\in\mathbb{I}} \left( \mathsf{f}_i(\mathsf{y}_i) + \frac{1}{2}\|\overline{\mathsf{x}}_i - \mathsf{y}_i\|^2 \right)$$
$$= \sum_{i\in\mathbb{I}} \min_{\mathsf{y}_i\in\mathsf{H}_i} \left( \mathsf{f}_i(\mathsf{y}_i) + \frac{1}{2}\|\overline{\mathsf{x}}_i - \mathsf{y}_i\|^2 \right)$$
$$= \sum_{i\in\mathbb{I}} \left( \mathsf{f}_i(\mathrm{prox}_{\mathsf{f}_i}\overline{\mathsf{x}}_i) + \frac{1}{2}\|\overline{\mathsf{x}}_i - \mathrm{prox}_{\mathsf{f}_i}\overline{\mathsf{x}}_i\|^2 \right)$$
$$= f(p) + \frac{1}{2}\|\overline{x} - p\|^2, \tag{2.3}$$

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

**Corollary 2.3** *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 2.1 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 2.4** In the context of Corollary 2.3, for every $i \in \mathbb{I}$, let $\omega_i \in [0,1]$, let $\eta_i \in ]0,+\infty[$, let $\delta_i \in ]0,+\infty[$, and set $\varrho_i \colon \xi \mapsto (2\omega_i/\pi)\arctan(\eta_i\,\xi) + (1-\omega_i)\text{sign}(\xi)(1-\exp(-\delta_i|\xi|))$. 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

$$p = \sum_{i\in\mathbb{I}} \frac{\varrho_i(\langle \overline{x} \mid e_i\rangle)}{2\omega_i\eta_i/\pi + (1-\omega_i)\delta_i} e_i \tag{2.4}$$

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

**Example 2.5 (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

$$p = \sum_{i\in\mathbb{I}} \big(\text{prox}_{\phi_i}\langle \overline{x} \mid e_i\rangle\big) e_i \tag{2.5}$$

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 $\omega \in ]0,+\infty[$. This yields $p = \sum_{i\in\mathbb{I}}(\text{sign}(\langle \overline{x} \mid e_i\rangle)\max\{|\langle \overline{x} \mid e_i\rangle| - \omega, 0\})e_i$. In general, to see that $p$ in (2.5) is a proximal point of $\overline{x}$, it suffices to apply Corollary 2.3 with, for every $i \in \mathbb{I}$, $\beta_i = 1$ and $\varrho_i = \text{prox}_{\phi_i}$, 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 $f \colon \mathcal{H} \to ]-\infty,+\infty] : x \mapsto \sum_{i\in\mathbb{I}} \phi_i(\langle x \mid e_i\rangle)$.

**Example 2.6 (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 2.1 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 $\mathcal{F}_i = \big\{\Omega_i \cap S \bigm| S \in \mathcal{F}\big\}$. 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 2.2 and [8, Proposition 2.1] that the proximal point of $\overline{x}$ relative to $f \colon x \mapsto \sum_{i\in\mathbb{I}} \phi_i(\|\mathsf{x}_i\|)$ is

$$p = \Big(\big(\text{prox}_{\phi_i}\|\overline{\mathsf{x}}_i\|\big)\mathsf{u}_{\rho_i}(\overline{\mathsf{x}}_i)\Big)_{i\in\mathbb{I}}, \quad \text{where} \quad \mathsf{u}_{\rho_i} \colon \mathsf{H}_i \to \mathsf{H}_i \colon \mathsf{x}_i \mapsto \begin{cases} \mathsf{x}_i/\|\mathsf{x}_i\|, & \text{if } \|\mathsf{x}_i\| > \rho_i;\\ 0, & \text{if } \|\mathsf{x}_i\| \leqslant \rho_i. \end{cases} \tag{2.6}$$

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

**Example 2.7 (group shrinkage)** In Example 2.6, 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 $\rho_i \in ]0,+\infty[$. Then it follows from [5, Example 14.5] that the proximal point $p$ of (2.6) 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],

$$p = \left(\left(1 - \frac{\rho_1}{\max\{\|\overline{\mathsf{x}}_1\|, \rho_1\}}\right)\overline{\mathsf{x}}_1, \ldots, \left(1 - \frac{\rho_m}{\max\{\|\overline{\mathsf{x}}_m\|, \rho_m\}}\right)\overline{\mathsf{x}}_m\right). \tag{2.7}$$

### 2.2 Prescriptions derived from cocoercive operators

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

$$(\forall x \in \mathcal{G})(\forall y \in \mathcal{G}) \quad \langle x - y \mid Qx - Qy \rangle \geqslant \beta \|Qx - Qy\|^2, \tag{2.8}$$

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 2.8** *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 $\beta_i \in \left]0, +\infty\right[$, 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*

$$\beta = \frac{1}{\displaystyle\sum_{i\in\mathbb{I}} \frac{\|L_i\|^2}{\beta_i}}, \quad p = \beta \sum_{i\in\mathbb{I}} L_i^* q_i, \quad \text{and} \quad F = \beta \sum_{i\in\mathbb{I}} L_i^* \circ Q_i \circ L_i. \tag{2.9}$$

*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 (2.8) yields

$$\begin{aligned} 0 &= \frac{\langle Fx - F\overline{x} \mid x - \overline{x} \rangle}{\beta} \\ &= \sum_{i\in\mathbb{I}} \langle Q_i(L_i x) - Q_i(L_i\overline{x}) \mid L_i x - L_i \overline{x} \rangle \\ &\geqslant \sum_{i\in\mathbb{I}} \beta_i \|Q_i(L_i x) - Q_i(L_i\overline{x})\|^2 \\ &= \sum_{i\in\mathbb{I}} \beta_i \|Q_i(L_i x) - q_i\|^2, \end{aligned} \tag{2.10}$$

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 2.8 are obtained through proximity operators.

**Proposition 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 $g_i \in \Gamma_0(\mathcal{G}_i)$, 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) $(\forall x \in \mathcal{H})\ Fx = p \Leftrightarrow (\forall i \in \mathbb{I})\ \mathrm{prox}_{g_i}(L_i x) = q_i$.

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

$$F = \beta\, \mathrm{prox}_f, \quad \textit{where} \quad f = \left( \sum_{i \in \mathbb{I}} \left( g_i^* \,\square\, \frac{\|\cdot\|_{\mathcal{G}_i}^2}{2} \right) \circ L_i \right)^* - \frac{\|\cdot\|_{\mathcal{H}}^2}{2}. \tag{2.11}$$

*Proof*. (i)–(ii): Apply Proposition 2.8 with $(\forall i \in \mathbb{I})\ Q_i = \mathrm{prox}_{g_i}$ and $\beta_i = 1$.

(iii): This follows from [18, Proposition 3.9]. ☐

**Example 2.10 (scalar observations)** We specialize the setting of Proposition 2.9 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 $\chi_i = \mathrm{prox}_{g_i} \langle \overline{x} \mid a_i \rangle$ 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

$$\chi_i = \mathrm{proj}_D \langle \overline{x} \mid a_i \rangle = \begin{cases} \overline{\delta}, & \text{if } \langle \overline{x} \mid a_i \rangle > \overline{\delta}; \\ \langle \overline{x} \mid a_i \rangle, & \text{if } \langle \overline{x} \mid a_i \rangle \in D; \\ \underline{\delta}, & \text{if } \langle \overline{x} \mid a_i \rangle < \underline{\delta}, \end{cases} \tag{2.12}$$

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 $\mathrm{soft}_\Omega$ be the associated soft thresholder, i.e.,

$$\mathrm{soft}_\Omega \colon \mathbb{R} \to \mathbb{R} \colon \xi \mapsto \begin{cases} \xi - \overline{\omega}, & \text{if } \xi > \overline{\omega}; \\ 0, & \text{if } \xi \in \Omega; \\ \xi - \underline{\omega}, & \text{if } \xi < \underline{\omega}, \end{cases} \qquad \text{with} \qquad \begin{cases} \overline{\omega} = \sup \Omega \\ \underline{\omega} = \inf \Omega. \end{cases} \tag{2.13}$$

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

$$\chi_i = \mathrm{prox}_\psi\big(\mathrm{soft}_\Omega \langle \overline{x} \mid a_i \rangle\big) = \begin{cases} \mathrm{prox}_\psi(\langle \overline{x} \mid a_i \rangle - \overline{\omega}), & \text{if } \langle \overline{x} \mid a_i \rangle > \overline{\omega}; \\ 0, & \text{if } \langle \overline{x} \mid a_i \rangle \in \Omega; \\ \mathrm{prox}_\psi(\langle \overline{x} \mid a_i \rangle - \underline{\omega}), & \text{if } \langle \overline{x} \mid a_i \rangle < \underline{\omega}. \end{cases} \tag{2.14}$$

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

$$\chi_i = \mathrm{sign}(\langle \overline{x} \mid a_i \rangle) \max\{|\langle \overline{x} \mid a_i \rangle| - \omega, 0\} \tag{2.15}$$

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 (2.14) becomes

$$\chi_i = \begin{cases} \overline{\delta}, & \text{if } \langle \overline{x} \mid a_i\rangle \geqslant \overline{\delta} + \overline{\omega};\\ \langle \overline{x} \mid a_i\rangle - \overline{\omega}, & \text{if } \overline{\omega} < \langle \overline{x} \mid a_i\rangle < \overline{\delta} + \overline{\omega};\\ 0, & \text{if } \langle \overline{x} \mid a_i\rangle \in \Omega;\\ \langle \overline{x} \mid a_i\rangle - \underline{\omega}, & \text{if } \underline{\delta} + \underline{\omega} < \langle \overline{x} \mid a_i\rangle < \underline{\omega};\\ \underline{\delta}, & \text{if } \langle \overline{x} \mid a_i\rangle \leqslant \underline{\delta} + \underline{\omega}. \end{cases} \tag{2.16}$$

This operation combines hard clipping and soft thresholding.

(iv) Set

$$g_i\colon \xi \mapsto \begin{cases} \dfrac{(1+\xi)\ln(1+\xi) + (1-\xi)\ln(1-\xi) - \xi^2}{2}, & \text{if } |\xi| < 1;\\ \ln(2) - 1/2, & \text{if } |\xi| = 1;\\ +\infty, & \text{if } |\xi| > 1. \end{cases} \tag{2.17}$$

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

$$g_i\colon \xi \mapsto \begin{cases} -\dfrac{2}{\pi}\ln\left(\cos\left(\dfrac{\pi\xi}{2}\right)\right) - \dfrac{\xi^2}{2}, & \text{if } |\xi| < 1;\\ +\infty, & \text{if } |\xi| \geqslant 1. \end{cases} \tag{2.18}$$

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

$$g_i\colon \xi \mapsto \begin{cases} -|\xi| - \ln(1-|\xi|) - \xi^2/2, & \text{if } |\xi| < 1;\\ +\infty, & \text{if } |\xi| \geqslant 1. \end{cases} \tag{2.19}$$

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

$$g_i\colon \xi \mapsto \begin{cases} |\xi| + (1-|\xi|)\ln\big|1-|\xi|\big| - \xi^2/2, & \text{if } |\xi| < 1;\\ 1/2, & \text{if } |\xi| = 1;\\ +\infty, & \text{if } |\xi| > 1. \end{cases} \tag{2.20}$$

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,

$$\big(\operatorname{Id} + g_i'\big)^{-1} = \operatorname{prox}_{g_i}\colon \xi \mapsto \operatorname{sign}(\xi)\big(1 - \exp(-|\xi|)\big) \tag{2.21}$$

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 $\eta_i \in \,]0,+\infty[$ and set

$$g_i \colon \xi \mapsto \eta_i \xi + \begin{cases} \xi \ln(\xi) + (1-\xi)\ln(1-\xi) - \xi^2/2, & \text{if } \xi \in \,]0,1[\,; \\ 0, & \text{if } \xi = 0; \\ -1/2, & \text{if } \xi = 1; \\ +\infty, & \text{if } \xi \in \mathbb{R} \smallsetminus [0,1]. \end{cases} \tag{2.22}$$

Proceeding as in (vii), we obtain

$$\chi_i = \frac{1}{1+\exp(\eta_i - \langle \overline{x} \mid a_i \rangle)}, \tag{2.23}$$

which is an encoding scheme used in [36].

**Example 2.11** In Proposition 2.9 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

$$q_i = \text{proj}_{D_i}(L_i \overline{x}) \tag{2.24}$$

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 $\rho_i \in \,]0,+\infty[$, then (2.24) reduces to the hard saturation process

$$q_i = \begin{cases} \dfrac{\rho_i}{\|L_i\overline{x}\|} L_i\overline{x}, & \text{if } \|L_i\overline{x}\| > \rho_i; \\ L_i\overline{x}, & \text{if } \|L_i\overline{x}\| \leqslant \rho_i, \end{cases} \tag{2.25}$$

which can be viewed as an infinite dimensional version of Example 2.10(i).

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

$$q_i = \begin{cases} L_i\overline{x} + \dfrac{\text{prox}_{\phi_i^*} d_{D_i}(L_i\overline{x})}{d_{D_i}(L_i\overline{x})} \big(\text{proj}_{D_i}(L_i\overline{x}) - L_i\overline{x}\big), & \text{if } d_{D_i}(L_i\overline{x}) > \rho_i; \\ \text{proj}_{D_i}(L_i\overline{x}), & \text{if } d_{D_i}(L_i\overline{x}) \leqslant \rho_i. \end{cases} \tag{2.26}$$

In particular, assume that $D_i = \{0\}$. Then (2.26) reduces to the abstract soft thresholding process

$$q_i = \begin{cases} L_i\overline{x} - \dfrac{\text{prox}_{\phi_i^*} \|L_i\overline{x}\|}{\|L_i\overline{x}\|} L_i\overline{x}, & \text{if } \|L_i\overline{x}\| > \rho_i; \\ 0, & \text{if } \|L_i\overline{x}\| \leqslant \rho_i, \end{cases} \tag{2.27}$$

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 $\rho_i \in \,]0,+\infty[$. Then $\phi_i^* = \iota_{[-\rho_i,\rho_i]}$, $\partial\phi_i(0) = [-\rho_i,\rho_i]$, and (2.27) becomes

$$q_i = \begin{cases} \left(1 - \dfrac{\rho_i}{\|L_i\overline{x}\|}\right) L_i\overline{x}, & \text{if } \|L_i\overline{x}\| > \rho_i; \\ 0, & \text{if } \|L_i\overline{x}\| \leqslant \rho_i, \end{cases} \tag{2.28}$$

which can be viewed as an infinite dimensional version of (2.15).

### 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 1.1. 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 2.12** In the spirit of the shrinkage ideas of Corollary 2.3 and Example 2.5, 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 2.3, 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

$$\sum_{i\in\mathbb{I}} \big|\varrho_i(\langle \overline{x} \mid e_i\rangle)\big|^2 \leqslant \delta^2 \sum_{i\in\mathbb{I}} |\langle \overline{x} \mid e_i\rangle|^2 = \delta^2\|\overline{x}\|^2 < +\infty, \tag{2.29}$$

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

$$(\forall \xi \in \mathbb{R}) \quad \big[\, \chi_i = \varrho_i(\xi) \quad \Leftrightarrow \quad \sigma_i(\chi_i) = \sigma_i\big(\varrho_i(\xi)\big) = \varphi_i(\xi) \,\big]. \tag{2.30}$$

Using Corollary 2.3 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]

$$\varrho_i \colon \xi \mapsto \begin{cases} \operatorname{sign}(\xi)\sqrt{\xi^2 - \omega_i^2}, & \text{if } |\xi| > \omega_i; \\ 0, & \text{if } |\xi| \leqslant \omega_i. \end{cases} \tag{2.31}$$

It is straightforward to verify that (2.30) holds with

$$\sigma_i \colon \xi \mapsto \operatorname{sign}(\xi)\left(\sqrt{\xi^2 + \omega_i^2} - \omega_i\right), \tag{2.32}$$

in which case $\varphi_i = \sigma_i \circ \varrho_i$ is the soft thresholder on $[-\omega_i, \omega_i]$ of (2.13).

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

$$\varrho_i = \operatorname{hard}_{[-\omega_i,\omega_i]} \colon \xi \mapsto \begin{cases} \xi, & \text{if } |\xi| > \omega_i; \\ 0, & \text{if } |\xi| \leqslant \omega_i, \end{cases} \tag{2.33}$$

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 (2.30) is satisfied with

$$\sigma_i \colon \xi \mapsto \xi - \omega_i \operatorname{sign}(\xi), \tag{2.34}$$

in which case $\varphi_i = \sigma_i \circ \varrho_i$ turns out to be the soft thresholder on $[-\omega_i, \omega_i]$ of (2.13).

Next, we revisit Proposition 2.1 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 2.13** *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 2.1. □

The following result illustrates the process described in Proposition 2.13, through a generalization of the discontinuous hard thresholding operator of Example 2.12(ii), which corresponds to the case when $\mathsf{H}_i=\mathbb{R}$ and $\mathsf{C}_i=\{0\}$ in (2.35) below.

**Proposition 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. For every* $i\in\mathbb{I}$*, let* $\omega_i\in\,]0,+\infty[$*, let* $\mathsf{C}_i$ *be a nonempty closed convex subset of* $\mathsf{H}_i$*, set*

$$\mathsf{Q}_i\colon\mathsf{H}_i\to\mathsf{H}_i\colon\mathsf{x}_i\mapsto\begin{cases}\mathsf{x}_i, & \text{if } d_{\mathsf{C}_i}(\mathsf{x}_i)>\omega_i;\\ \mathrm{proj}_{\mathsf{C}_i}\mathsf{x}_i, & \text{if } d_{\mathsf{C}_i}(\mathsf{x}_i)\leqslant\omega_i,\end{cases}\tag{2.35}$$

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

$$\mathsf{S}_i\colon\mathsf{H}_i\to\mathsf{H}_i\colon\mathsf{x}_i\mapsto\begin{cases}\mathsf{x}_i+\dfrac{\omega_i}{d_{\mathsf{C}_i}(\mathsf{x}_i)}(\mathrm{proj}_{\mathsf{C}_i}\mathsf{x}_i-\mathsf{x}_i), & \text{if } \mathsf{x}_i\notin\mathsf{C}_i;\\ \mathsf{x}_i, & \text{if } \mathsf{x}_i\in\mathsf{C}_i\end{cases}\quad\text{and}\quad\begin{cases}\mathsf{F}_i=\mathsf{S}_i\circ\mathsf{Q}_i\\ \mathsf{p}_i=\mathsf{S}_i\mathsf{q}_i.\end{cases}\tag{2.36}$$

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

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

(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* $\big[\,(\forall i\in\mathbb{I})\ \mathsf{Q}_i\mathsf{x}_i=\mathsf{q}_i\,\big]\Leftrightarrow\mathrm{prox}_f x=p$.

*Proof*. We derive from (2.35), (2.36), and [5, Proposition 3.21] that

$$(\forall i\in\mathbb{I})(\forall\mathsf{x}_i\in\mathsf{H}_i)\quad\mathsf{F}_i\mathsf{x}_i=\begin{cases}\mathrm{proj}_{\mathsf{C}_i}\mathsf{x}_i+\left(1-\dfrac{\omega_i}{d_{\mathsf{C}_i}(\mathsf{x}_i)}\right)\big(\mathsf{x}_i-\mathrm{proj}_{\mathsf{C}_i}\mathsf{x}_i\big)\notin\mathsf{C}_i, & \text{if } d_{\mathsf{C}_i}(\mathsf{x}_i)>\omega_i;\\ \mathrm{proj}_{\mathsf{C}_i}\mathsf{x}_i\in\mathsf{C}_i, & \text{if } d_{\mathsf{C}_i}(\mathsf{x}_i)\leqslant\omega_i.\end{cases}\tag{2.37}$$

(i): This is a consequence of (2.37) and [5, Example 24.28].

(ii): If $\mathbb{I}$ is infinite, $(\forall i \in \mathbb{I})\ 0 \in \mathsf{C}_i \Rightarrow d_{\mathsf{C}_i}(0) = 0 \Rightarrow \mathsf{F}_i(0) = 0$ by (2.37). In turn, the claim follows from Corollary 2.2 and (i).

(iii): We first note that Corollary 2.2 and (i) imply that

$$(\mathsf{F}_i\mathsf{x}_i)_{i\in\mathbb{I}} = (\mathrm{prox}_{\omega_i d_{\mathsf{C}_i}}\mathsf{x}_i)_{i\in\mathbb{I}} = \mathrm{prox}_f x. \tag{2.38}$$

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, (2.38) yields $\mathrm{prox}_f x = (\mathsf{F}_i\mathsf{x}_i)_{i\in\mathbb{I}} = p$. Conversely, suppose that $\mathrm{prox}_f x = p$ and fix $i \in \mathbb{I}$. We derive from (2.38) and (2.36) that

$$\mathsf{F}_i\mathsf{x}_i = \mathsf{p}_i = \mathsf{S}_i\mathsf{q}_i = \mathsf{S}_i(\mathsf{Q}_i\overline{\mathsf{x}}_i) = \mathsf{F}_i\overline{\mathsf{x}}_i. \tag{2.39}$$

We must show that $\mathsf{Q}_i\mathsf{x}_i = \mathsf{q}_i$. It follows from (2.35), (2.37), and (2.39) that

$$d_{\mathsf{C}_i}(\mathsf{x}_i) \leqslant \omega_i \ \Leftrightarrow\ \mathsf{Q}_i\mathsf{x}_i = \mathrm{proj}_{\mathsf{C}_i}\mathsf{x}_i = \mathsf{F}_i\mathsf{x}_i = \mathsf{F}_i\overline{\mathsf{x}}_i \in \mathsf{C}_i \ \Rightarrow\ \begin{cases} d_{\mathsf{C}_i}(\overline{\mathsf{x}}_i) \leqslant \omega_i \\ \mathsf{Q}_i\mathsf{x}_i = \mathrm{proj}_{\mathsf{C}_i}\overline{\mathsf{x}}_i = \mathsf{Q}_i\overline{\mathsf{x}}_i = \mathsf{q}_i. \end{cases} \tag{2.40}$$

On the other hand, (2.35) yields

$$d_{\mathsf{C}_i}(\mathsf{x}_i) > \omega_i \quad \Rightarrow \quad \mathsf{Q}_i\mathsf{x}_i = \mathsf{x}_i, \tag{2.41}$$

while (2.39) and (2.37) yield

$$\begin{aligned} d_{\mathsf{C}_i}(\mathsf{x}_i) > \omega_i \quad &\Rightarrow \quad \mathsf{p}_i = \mathsf{F}_i\overline{\mathsf{x}}_i = \mathsf{F}_i\mathsf{x}_i = \mathrm{proj}_{\mathsf{C}_i}\mathsf{x}_i + \left(1 - \frac{\omega_i}{d_{\mathsf{C}_i}(\mathsf{x}_i)}\right)(\mathsf{x}_i - \mathrm{proj}_{\mathsf{C}_i}\mathsf{x}_i) \notin \mathsf{C}_i &(2.42)\\ &\Rightarrow \quad \mathsf{F}_i\overline{\mathsf{x}}_i = \mathrm{proj}_{\mathsf{C}_i}\overline{\mathsf{x}}_i + \left(1 - \frac{\omega_i}{d_{\mathsf{C}_i}(\overline{\mathsf{x}}_i)}\right)(\overline{\mathsf{x}}_i - \mathrm{proj}_{\mathsf{C}_i}\overline{\mathsf{x}}_i) \ \text{ and } \ d_{\mathsf{C}_i}(\overline{\mathsf{x}}_i) > \omega_i &(2.43)\\ &\Rightarrow \quad \mathsf{q}_i = \mathsf{Q}_i\overline{\mathsf{x}}_i = \overline{\mathsf{x}}_i. &(2.44) \end{aligned}$$

Therefore, in view of (2.41), it remains to show that $\overline{\mathsf{x}}_i = \mathsf{x}_i$. Set $\mathsf{r}_i = \mathrm{proj}_{\mathsf{C}_i}\mathsf{p}_i$. We deduce from (2.42), (2.43), and [5, Proposition 3.21] that $\mathsf{r}_i = \mathrm{proj}_{\mathsf{C}_i}\overline{\mathsf{x}}_i = \mathrm{proj}_{\mathsf{C}_i}\mathsf{x}_i$. Thus, (2.42) and (2.43) yield

$$\mathsf{p}_i - \mathsf{r}_i = \left(1 - \frac{\omega_i}{\|\mathsf{x}_i - \mathsf{r}_i\|}\right)(\mathsf{x}_i - \mathsf{r}_i) = \left(1 - \frac{\omega_i}{\|\overline{\mathsf{x}}_i - \mathsf{r}_i\|}\right)(\overline{\mathsf{x}}_i - \mathsf{r}_i). \tag{2.45}$$

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 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 4 to the setting of Problem 1.1.

**Problem 3.1** 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 $\mathrm{proj}_C x_0$, i.e., to

$$\text{minimize} \ \ \|x - x_0\| \quad \text{subject to} \quad x \in \bigcap_{i\in I} C_i. \tag{3.1}$$

In 1968, Yves Haugazeau proposed in his unpublished thesis [34] an iterative method to solve Problem 3.1 when $I$ is finite. His algorithm proceeds by periodic projections onto the individual sets.

**Proposition 3.2** [34, Théorème 3-2] *In Problem 3.1, 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*

$$D = \big\{ x \in \mathcal{H} \;\big|\; \langle x - s \mid x_0 - s\rangle \leqslant 0 \;\; \textit{and} \;\; \langle x - t \mid s - t\rangle \leqslant 0 \big\} \neq \varnothing, \tag{3.2}$$

*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*

$$Q(x_0, s, t) = \text{proj}_D x_0 = \begin{cases} t, & \textit{if} \;\; \rho = 0 \;\; \textit{and} \;\; \chi \geqslant 0;\\ x_0 + \left(1 + \dfrac{\chi}{\nu}\right)(t-s), & \textit{if} \;\; \rho > 0 \;\; \textit{and} \;\; \chi\nu \geqslant \rho;\\ s + \dfrac{\nu}{\rho}\big(\chi(x_0 - s) + \mu(t-s)\big), & \textit{if} \;\; \rho > 0 \;\; \textit{and} \;\; \chi\nu < \rho. \end{cases} \tag{3.3}$$

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

$$\begin{array}{l} \textit{for} \;\; n = 0, 1, \dots \\ \left\lfloor \begin{array}{l} t_n = \text{proj}_{C_{n (\text{mod } m)}} x_n \\ x_{n+1} = Q(x_0, x_n, t_n). \end{array} \right. \end{array} \tag{3.4}$$

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

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

**Proposition 3.3** [50, Théorème V.1] *In Problem 3.1, suppose that $I$ is finite, let $Q$ be as in Proposition 3.2, set $\omega = 1/\text{card}\, I$, and fix $\varepsilon \in \left]0,1\right[$. Construct a sequence $(x_n)_{n\in\mathbb{N}}$ by iterating*

$$\begin{array}{l} \textit{for} \;\; n = 0, 1, \dots \\ \left\lfloor \begin{array}{l} \textit{for every} \;\; i \in I \\ \left\lfloor \begin{array}{l} a_{i,n} = \text{proj}_{C_i} x_n \\ \theta_{i,n} = \|a_{i,n} - x_n\|^2 \end{array} \right. \\ \theta_n = \omega \sum_{i \in I} \theta_{i,n} \\ \textit{if} \; \theta_n = 0 \\ \left\lfloor \; t_n = x_n \right. \\ \textit{else} \\ \left\lfloor \begin{array}{l} d_n = \omega \sum_{i\in I} a_{i,n} \\ y_n = d_n - x_n \\ \lambda_n = \theta_n / \|y_n\|^2 \\ t_n = x_n + \lambda_n y_n \end{array} \right. \\ x_{n+1} = Q(x_0, x_n, t_n). \end{array} \right. \end{array} \tag{3.5}$$

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

**Remark 3.4** An attractive feature of Pierra's algorithm (3.5) 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 and 3.3 were unified and extended in [15, Section 6.5] in the form of an algorithm for solving Problem 3.1 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 (3.5). 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 $\operatorname{card} I_n$ 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 (3.4) and (3.5). Below, we propose an alternative block-iterative scheme (Algorithm 3.9) 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 4, this new algorithm will be better suited to solve Problem 1.1 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 3.5** [5, Section 4.1] $\mathfrak{T}$ is the class of *firmly quasinonexpansive operators* from $\mathcal{H}$ to $\mathcal{H}$, i.e.,

$$\mathfrak{T} = \big\{T\colon \mathcal{H} \to \mathcal{H} \;\big|\; (\forall x \in \mathcal{H})(\forall y \in \operatorname{Fix} T)\;\; \langle y - Tx \mid x - Tx\rangle \leqslant 0\big\}. \tag{3.6}$$

**Example 3.6** [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 $f$,

$$(\forall x \in \mathcal{H})\;\; Tx = \operatorname{sproj}_C x = \begin{cases} x - \dfrac{f(x)}{\|s(x)\|^2}s(x), & \text{if } f(x) > 0;\\ x, & \text{if } f(x) \leqslant 0.\end{cases} \tag{3.7}$$

**Lemma 3.7** [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 3.8** *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, 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 3.1. It incorporates ingredients of the best approximation method of [15, Section 6.5] and of the convex feasibility method of [6].

**Algorithm 3.9** Consider the setting of Problem 3.1 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, fix $\varepsilon \in \left]0,1\right[$, and iterate

$$
\begin{array}{l}
\text{for } n = 0,1,\ldots\\
\left\lfloor
\begin{array}{l}
\text{take } \mathrm{i}(n) \in I'\\
z_n = \operatorname{proj}_{C_{\mathrm{i}(n)}} x_n\\
\text{take a nonempty finite set } I_n \subset I\\
\text{for every } i \in I_n\\
\left\lfloor
\begin{array}{l}
\text{take } T_{i,n} \in \mathfrak{T} \text{ such that } \operatorname{Fix} T_{i,n} = C_i\\
a_{i,n} = T_{i,n} z_n\\
\theta_{i,n} = \|a_{i,n} - z_n\|^2
\end{array}
\right.\\
\text{take } j_n \in I_n \text{ such that } \theta_{j_n,n} = \max_{i\in I_n} \theta_{i,n}\\
\text{take } \{\omega_{i,n}\}_{i\in I_n} \subset [0,1] \text{ such that } \sum_{i\in I_n} \omega_{i,n} = 1 \text{ and } \omega_{j_n,n} \geqslant \varepsilon\\
I_n^+ = \big\{ i \in I_n \;\big|\; \omega_{i,n} > 0 \big\}\\
\theta_n = \sum_{i\in I_n^+} \omega_{i,n}\theta_{i,n}\\
\text{if } \theta_n = 0\\
\left\lfloor\; t_n = z_n \right.\\
\text{else}\\
\left\lfloor
\begin{array}{l}
d_n = \sum_{i\in I_n^+} \omega_{i,n} a_{i,n}\\
y_n = \operatorname{proj}_{C_{\mathrm{i}(n)}} d_n - z_n\\
\text{take } \lambda_n \in \big[\varepsilon\theta_n/\|d_n - z_n\|^2, \theta_n/\|y_n\|^2\big]\\
t_n = z_n + \lambda_n y_n
\end{array}
\right.\\
x_{n+1} = Q(x_0, x_n, t_n).
\end{array}
\right.
\end{array}
\tag{3.8}
$$

**Remark 3.10** Let us highlight some special cases and features of Algorithm 3.9.

(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 (3.4) and (3.5).

(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 3.6 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 5.

We now establish the strong convergence of an arbitrary sequence $(x_n)_{n\in\mathbb{N}}$ generated by Algorithm 3.9 to the solution to Problem 3.1. The last component of the proof relies on Lemma 3.8(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 3.10(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 3.11** *In the setting of Problem 3.1, let $(x_n)_{n\in\mathbb{N}}$ be generated by Algorithm 3.9. Suppose that the following hold:*

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

$$(\forall i \in I)(\forall n \in \mathbb{N}) \quad i \in \bigcup_{l=n}^{n+M_i-1} \{\mathrm{i}(l)\} \cup I_l. \tag{3.9}$$

[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 $\mathbb{N}$,*

$$\begin{cases} i \in \bigcap_{n\in\mathbb{N}} I_{r_n} \\ \mathrm{proj}_{C_{\mathrm{i}(r_n)}} x_{r_n} \rightharpoonup x \\ T_{i,r_n}\big(\mathrm{proj}_{C_{\mathrm{i}(r_n)}} x_{r_n}\big) - \mathrm{proj}_{C_{\mathrm{i}(r_n)}} x_{r_n} \to 0 \end{cases} \quad \Rightarrow \quad x \in C_i. \tag{3.10}$$

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

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

$$L_n \colon \mathcal{H} \to \mathbb{R} \colon z \mapsto \begin{cases} \dfrac{\sum_{i\in I_n^+} \omega_{i,n}\|T_{i,n}z - z\|^2}{\big\|\sum_{i\in I_n^+} \omega_{i,n}T_{i,n}z - z\big\|^2}, & \text{if } z \notin \bigcap_{i\in I_n^+} C_i; \\ 1, & \text{if } z \in \bigcap_{i\in I_n^+} C_i \end{cases} \tag{3.11}$$

and

$$S_n \colon \mathcal{H} \to \mathcal{H} \colon z \mapsto z + L_n(z)\Bigg(\sum_{i\in I_n^+} \omega_{i,n}T_{i,n}z - z\Bigg). \tag{3.12}$$

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

$$\theta_n = 0 \;\Leftrightarrow\; S_n z_n = z_n \;\Leftrightarrow\; z_n \in \bigcap_{i\in I_n^+} C_i = \mathrm{Fix}\, S_n. \tag{3.13}$$

Now define

$$K_n \colon \mathcal{H} \to \mathbb{R} \colon x \mapsto \begin{cases} \dfrac{\big\|S_n\big(\mathrm{proj}_{C_{\mathrm{i}(n)}} x\big) - \mathrm{proj}_{C_{\mathrm{i}(n)}} x\big\|^2}{\big\|\mathrm{proj}_{C_{\mathrm{i}(n)}}\big(S_n\big(\mathrm{proj}_{C_{\mathrm{i}(n)}} x\big)\big) - \mathrm{proj}_{C_{\mathrm{i}(n)}} x\big\|^2}, & \text{if } \mathrm{proj}_{C_{\mathrm{i}(n)}} x \notin \bigcap_{i\in I_n^+} C_i; \\ 1, & \text{if } \mathrm{proj}_{C_{\mathrm{i}(n)}} x \in \bigcap_{i\in I_n^+} C_i \end{cases} \tag{3.14}$$

and

$$T_n \colon \mathcal{H} \to \mathcal{H} \colon x \mapsto \operatorname{proj}_{C_{\mathrm{i}(n)}} x + \gamma_n(x)\Big(\operatorname{proj}_{C_{\mathrm{i}(n)}}\Big(S_n\big(\operatorname{proj}_{C_{\mathrm{i}(n)}} x\big)\Big) - \operatorname{proj}_{C_{\mathrm{i}(n)}} x\Big), \quad \text{where} \quad \gamma_n(x) \in [\varepsilon, K_n(x)]. \tag{3.15}$$

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

$$\varnothing \neq C \subset C_{\mathrm{i}(n)} \cap \bigcap_{i \in I_n^+} C_i = C_{\mathrm{i}(n)} \cap \operatorname{Fix} S_n = \operatorname{Fix} T_n. \tag{3.16}$$

If $\theta_n \neq 0$, using (3.8), (3.12), and the fact that $\operatorname{proj}_{C_{\mathrm{i}(n)}}$ is an affine operator [5, Corollary 3.22(ii)], we obtain

$$\begin{aligned}
\operatorname{proj}_{C_{\mathrm{i}(n)}}\Big(S_n\big(\operatorname{proj}_{C_{\mathrm{i}(n)}} x_n\big)\Big) - \operatorname{proj}_{C_{\mathrm{i}(n)}} x_n &= \operatorname{proj}_{C_{\mathrm{i}(n)}}(S_n z_n) - z_n \\
&= \operatorname{proj}_{C_{\mathrm{i}(n)}}\Big(\big(1 - L_n(z_n)\big) z_n + L_n(z_n) d_n\Big) - z_n \\
&= \big(1 - L_n(z_n)\big)\operatorname{proj}_{C_{\mathrm{i}(n)}} z_n + L_n(z_n)\operatorname{proj}_{C_{\mathrm{i}(n)}} d_n - z_n \\
&= L_n(z_n)\big(\operatorname{proj}_{C_{\mathrm{i}(n)}} d_n - z_n\big) \\
&= L_n(z_n) y_n
\end{aligned} \tag{3.17}$$

and, therefore,

$$\big\|\operatorname{proj}_{C_{\mathrm{i}(n)}}(S_n z_n) - z_n\big\| = L_n(z_n)\|y_n\|. \tag{3.18}$$

Hence, (3.14) and (3.12) yield

$$K_n(x_n) = \begin{cases} \dfrac{\|S_n(z_n) - z_n\|^2}{\big\|\operatorname{proj}_{C_{\mathrm{i}(n)}}(S_n z_n) - z_n\big\|^2} = \dfrac{\|L_n(z_n)(d_n - z_n)\|^2}{\|L_n(z_n) y_n\|^2} = \dfrac{\|d_n - z_n\|^2}{\|y_n\|^2}, & \text{if } \theta_n \neq 0; \\ 1, & \text{if } \theta_n = 0. \end{cases} \tag{3.19}$$

At the same time, we derive from (3.11), (3.8), and (3.13) that

$$L_n(z_n) = \begin{cases} \dfrac{\theta_n}{\|d_n - z_n\|^2}, & \text{if } \theta_n \neq 0; \\ 1, & \text{if } \theta_n = 0. \end{cases} \tag{3.20}$$

Altogether, it results from (3.15), (3.19), and (3.20) that, if $\theta_n \neq 0$,

$$\gamma_n(x_n) L_n(z_n) \in \big[\varepsilon L_n(z_n), K_n(x_n) L_n(z_n)\big] = \big[\varepsilon\theta_n/\|d_n - z_n\|^2, \theta_n/\|y_n\|^2\big] \tag{3.21}$$

and, in view of (3.8), we can therefore set $\lambda_n = \gamma_n(x_n) L_n(z_n)$. Thus, it follows from (3.8) and (3.17) that

$$\begin{aligned}
\theta_n \neq 0 \ \Rightarrow\ t_n &= z_n + \lambda_n y_n \\
&= z_n + \gamma_n(x_n) L_n(z_n) y_n \\
&= \operatorname{proj}_{C_{\mathrm{i}(n)}} x_n + \gamma_n(x_n)\Big(\operatorname{proj}_{C_{\mathrm{i}(n)}}\big(S_n\big(\operatorname{proj}_{C_{\mathrm{i}(n)}} x_n\big)\big) - \operatorname{proj}_{C_{\mathrm{i}(n)}} x_n\Big) \\
&= T_n x_n.
\end{aligned} \tag{3.22}$$

On the other hand, (3.8) and (3.13) yield

$$\theta_n = 0 \ \Rightarrow \ t_n = z_n = S_n z_n = T_n x_n. \tag{3.23}$$

Combining (3.22) and (3.23), we obtain

$$x_{n+1} = Q(x_0, x_n, T_n x_n). \tag{3.24}$$

Turning back to (3.15) and (3.8), we deduce from [5, Corollary 3.22(i)] that

$$\begin{aligned}
\|T_n x_n - x_n\|^2 &= \big\|z_n - x_n + \gamma_n(x_n)\big(\mathrm{proj}_{C_{\mathrm{i}(n)}}(S_n z_n) - z_n\big)\big\|^2 \\
&= \|z_n - x_n\|^2 + 2\gamma_n(x_n)\langle \mathrm{proj}_{C_{\mathrm{i}(n)}} x_n - x_n \mid \mathrm{proj}_{C_{\mathrm{i}(n)}}(S_n z_n) - \mathrm{proj}_{C_{\mathrm{i}(n)}} x_n\rangle \\
&\quad + |\gamma_n(x_n)|^2 \big\|\mathrm{proj}_{C_{\mathrm{i}(n)}}(S_n z_n) - z_n\big\|^2 \\
&= \|z_n - x_n\|^2 + |\gamma_n(x_n)|^2 \big\|\mathrm{proj}_{C_{\mathrm{i}(n)}}(S_n z_n) - z_n\big\|^2 \\
&\geqslant \|z_n - x_n\|^2 + \varepsilon^2 \big\|\mathrm{proj}_{C_{\mathrm{i}(n)}}(S_n z_n) - z_n\big\|^2.
\end{aligned} \tag{3.25}$$

Since (3.16) implies that

$$\varnothing \neq C \subset \bigcap_{n\in\mathbb{N}} \mathrm{Fix}\, T_n, \tag{3.26}$$

we derive from (3.24) and Lemma 3.8(i) that $(x_n)_{n\in\mathbb{N}}$ is well defined. Furthermore, (3.25) and Lemma 3.8(iii) guarantee that

$$\sum_{n\in\mathbb{N}} \|z_n - x_n\|^2 < +\infty \tag{3.27}$$

and

$$\sum_{n\in\mathbb{N}} \|\mathrm{proj}_{C_{\mathrm{i}(n)}}(S_n z_n) - z_n\|^2 < +\infty. \tag{3.28}$$

Finally, in view of (3.26) and Lemma 3.8(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], (3.27), and (3.28), showing this inclusion can be done by following the same steps as in the proof of [6, Theorem 3.3(vi)]. □

**Remark 3.12** Condition [a] in Theorem 3.11 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 3.11 is discussed in [6, Section 3.4], where concrete scenarios that satisfy it are described.

## 4 Fixed point model and algorithm for Problem 1.1

To solve Problem 1.1, we are going to reformulate it as an instance of Problem 3.1. To this end, let us set

$$(\forall k \in K) \quad C_k = \big\{x \in \mathcal{H} \;\big|\; F_k x = p_k\big\} \quad \text{and} \quad T_k = p_k + \mathrm{Id} - F_k. \tag{4.1}$$

Then it follows from (1.9) that

$$(\forall k \in K) \quad T_k \text{ is firmly nonexpansive and } \operatorname{Fix} T_k = C_k. \tag{4.2}$$

We therefore deduce from Lemma 3.7 that $(C_k)_{k\in K}$ are closed convex subsets of $\mathcal{H}$. Thus, upon setting $I = J \cup K$, we recast Problem 1.1 is an instantiation of Problem 3.1. This leads us to the following solution method based on Algorithm 3.9.

**Proposition 4.1** *In the setting of Problem 1.1, let $Q$ be as in Proposition 3.2, 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*

$$\begin{array}{l}
\textit{for } n = 0,1,\ldots\\
\left\lfloor
\begin{array}{l}
\textit{take } \mathrm{i}(n) \in I'\\
z_n = \operatorname{proj}_{C_{\mathrm{i}(n)}} x_n\\
\textit{take a nonempty finite set } I_n \subset J \cup K\\
\textit{for every } i \in I_n\\
\left\lfloor
\begin{array}{l}
\textit{if } i \in J\\
\left\lfloor
\begin{array}{l}
\textit{take } T_{i,n} \in \mathfrak{T} \textit{ such that } \operatorname{Fix} T_{i,n} = C_i\\
a_{i,n} = T_{i,n} z_n
\end{array}
\right.\\
\textit{if } i \in K\\
\left\lfloor
\begin{array}{l}
a_{i,n} = p_i + z_n - F_i z_n
\end{array}
\right.\\
\theta_{i,n} = \|a_{i,n} - z_n\|^2
\end{array}
\right.\\
\textit{take } j_n \in I_n \textit{ such that } \theta_{j_n,n} = \max_{i\in I_n} \theta_{i,n}\\
\textit{take } \{\omega_{i,n}\}_{i\in I_n} \subset [0,1] \textit{ such that } \sum_{i\in I_n} \omega_{i,n} = 1 \textit{ and } \omega_{j_n,n} \geqslant \varepsilon\\
I_n^+ = \big\{ i \in I_n \;\big|\; \omega_{i,n} > 0 \big\}\\
\theta_n = \sum_{i\in I_n^+} \omega_{i,n} \theta_{i,n}\\
\textit{if } \theta_n = 0\\
\left\lfloor
\begin{array}{l}
t_n = z_n
\end{array}
\right.\\
\textit{else}\\
\left\lfloor
\begin{array}{l}
d_n = \sum_{i\in I_n^+} \omega_{i,n} a_{i,n}\\
y_n = \operatorname{proj}_{C_{\mathrm{i}(n)}} d_n - z_n\\
\textit{take } \lambda_n \in \big[\varepsilon\theta_n/\|d_n - z_n\|^2, \theta_n/\|y_n\|^2\big]\\
t_n = z_n + \lambda_n y_n
\end{array}
\right.\\
x_{n+1} = Q(x_0, x_n, t_n).
\end{array}
\right.
\end{array} \tag{4.3}$$

*Suppose that condition [a] in Theorem 3.11 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 $\mathbb{N}$, (3.10) holds.*

*Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to the solution to Problem 1.1.*

*Proof*. Let us bring into play (4.1) and (4.2). As discussed above, Problem 1.1 is an instance of Problem 3.1, where $I = J \cup K$. Now set

$$(\forall k \in K)(\forall n \in \mathbb{N}) \quad T_{k,n} = T_k = p_k + \operatorname{Id} - F_k. \tag{4.4}$$

Then (3.8) reduces to (4.3) and, in view of condition [c] above, to conclude via Theorem 3.11, it suffices to check that condition [b] in Theorem 3.11 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 $(\forall n \in \mathbb{N})\ u_n = \text{proj}_{C_{\mathrm{i}(r_n)}} x_{r_n}$. Suppose that $u_n \rightharpoonup x \in \mathcal{H}$ and that $T_{k,r_n}u_n - u_n \to 0$. Then (4.4) yields $T_k u_n - u_n \to 0$ and, since $T_k$ is nonexpansive by (4.2), it follows from Browder's demiclosedness principle [5, Corollary 4.28] that $x \in \text{Fix}\, T_k = C_k$, which concludes the proof. ☐

As was mentioned in Remark 3.10(iv) and will be illustrated in Section 5, exploiting the presence of affine subspaces typically leads to faster convergence. Problem 1.1 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 4.1 leads to the following implementation.

**Corollary 4.2** *In the setting of Problem 1.1, let $Q$ be as in Proposition 3.2, and fix $\varepsilon \in\, ]0,1[$. Iterate*

$$
\begin{array}{l}
\textit{for}\ n = 0,1,\ldots\\
\left\lfloor
\begin{array}{l}
\textit{take a nonempty finite set}\ I_n \subset J \cup K\\
\textit{for every}\ i \in I_n\\
\left\lfloor
\begin{array}{l}
\textit{if}\ i \in J\\
\left\lfloor
\begin{array}{l}
\textit{take}\ T_{i,n} \in \mathfrak{T}\ \textit{such that}\ \text{Fix}\, T_{i,n} = C_i\\
a_{i,n} = T_{i,n}x_n
\end{array}
\right.\\
\textit{if}\ i \in K\\
\left\lfloor\ a_{i,n} = p_i + x_n - F_i x_n \right.\\
\theta_{i,n} = \|a_{i,n} - x_n\|^2
\end{array}
\right.\\
\textit{take}\ j_n \in I_n\ \textit{such that}\ \theta_{j_n,n} = \max_{i\in I_n}\theta_{i,n}\\
\textit{take}\ \{\omega_{i,n}\}_{i\in I_n} \subset [0,1]\ \textit{such that}\ \sum_{i\in I_n}\omega_{i,n} = 1\ \textit{and}\ \omega_{j_n,n} \geqslant \varepsilon\\
I_n^+ = \big\{ i \in I_n \;\big|\; \omega_{i,n} > 0 \big\}\\
\theta_n = \sum_{i\in I_n^+}\omega_{i,n}\theta_{i,n}\\
\textit{if}\ \theta_n = 0\\
\left\lfloor\ t_n = x_n \right.\\
\textit{else}\\
\left\lfloor
\begin{array}{l}
y_n = \sum_{i\in I_n^+}\omega_{i,n}a_{i,n} - x_n\\
\textit{take}\ \lambda_n \in \big[\varepsilon\theta_n/\|y_n\|^2, \theta_n/\|y_n\|^2\big]\\
t_n = x_n + \lambda_n y_n
\end{array}
\right.\\
x_{n+1} = Q(x_0, x_n, t_n).
\end{array}
\right.
\end{array}
\tag{4.5}
$$

*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 $\mathbb{N}$,*

$$
\Big[ i \in \bigcap_{n\in\mathbb{N}} I_{r_n},\ x_{r_n} \rightharpoonup x,\ \textit{and}\ T_{i,r_n}x_{r_n} - x_{r_n} \to 0 \Big] \quad \Rightarrow \quad x \in C_i. \tag{4.6}
$$

*Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to the solution to Problem 1.1.*

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

(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 2.11(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 $k \in K$, $q_k = \text{proj}_D(L_k\overline{x})$, 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

$$\text{minimize}\;\; \|x\| \quad \text{subject to} \quad x \in C_1 \cap C_2 \quad \text{and} \quad (\forall k \in K) \quad \text{proj}_D(L_kx) = q_k. \tag{5.1}$$

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 $F_k = \|L_k\|^{-2}L_k^*\circ\text{proj}_D\circ L_k$. For every $k \in K$, applying Proposition 2.8 with $\mathbb{I} = \{k\}$, $\mathcal{G}_k = \mathcal{G}$, $\beta_k = 1$, and $Q_k = \text{proj}_D$ shows that $p_k$ is the proximal point of $\overline{x}$ relative to $F_k$ and, for every $x \in \mathcal{H}$, $F_kx = p_k \Leftrightarrow \text{proj}_D(L_kx) = q_k$. We therefore arrive at an instance of Problem 1.1 which is equivalent to (5.1), namely

$$\text{minimize}\;\; \|x\| \quad \text{subject to} \quad x \in C_1 \cap C_2 \quad \text{and} \quad (\forall k \in K) \quad F_kx = p_k. \tag{5.2}$$

With an eye towards algorithm (4.3), 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 $T_{2,n} = \text{sproj}_{C_2}$ of (3.7) (see [19] for its computation) since the direct projector is hard to implement. The fact that condition [c] in Proposition 4.1 is satisfied follows from [5, Proposition 29.41(vi)(a)]. We solve (5.2) with algorithm (4.3) to obtain the solution $x_\infty$ shown in Figure 2 (see [33, Algorithm 8.1.1] for the computation of $\text{proj}_D$).

To demonstrate the benefits of exploiting the presence of affine subspaces in algorithm (4.3), we show in Figure 3 the approximate solution it generates after $1000$ iterations. For the sake of comparison, we display in Figure 4 the approximate solution generated by algorithm (4.5) after $1000$ iterations. The following parameters are used:

- **Algorithm (4.3):** For every $n \in \mathbb{N}$, $\text{i}(n) = 1$, and whenever $\theta_n \neq 0$,

$$\lambda_n = \begin{cases} \dfrac{\theta_n}{2\|y_n\|^2}, & \text{if } n \equiv 0 \mod 3;\\[2mm] \dfrac{\theta_n}{\|y_n\|^2}, & \text{if } n \not\equiv 0 \mod 3. \end{cases} \tag{5.3}$$

  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 3.11 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 (4.5):** Iteration $n \in \mathbb{N}$ is executed with the same relaxation scheme (5.3) as in algorithm (4.3), 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 (5.2), Figures 3 and 4 demonstrate qualitatively that algorithm (4.3) yields faster convergence to the solution $x_\infty$ than algorithm (4.5). This is confirmed quantitatively by the error plots of Figure 5.

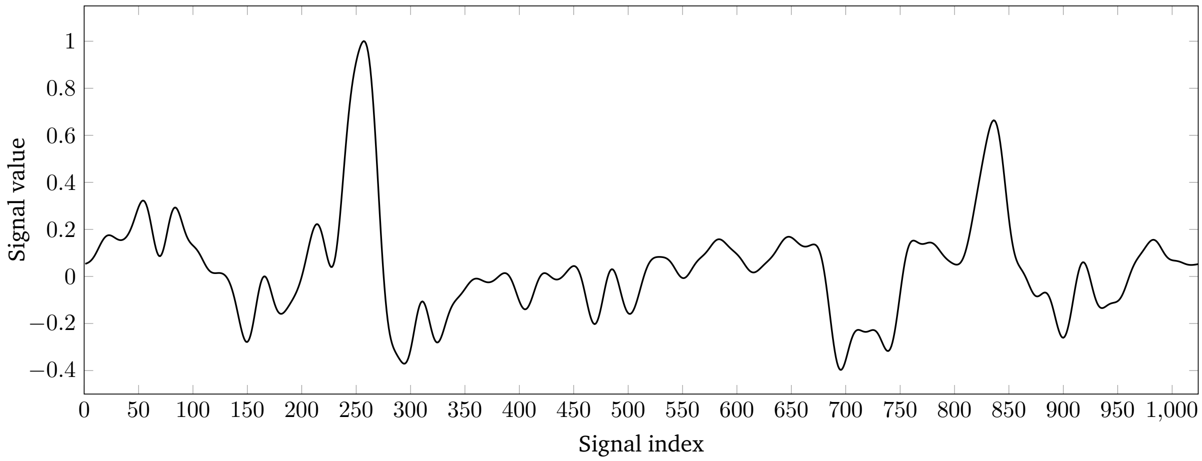


Figure 1: Original signal $\overline{x}$.

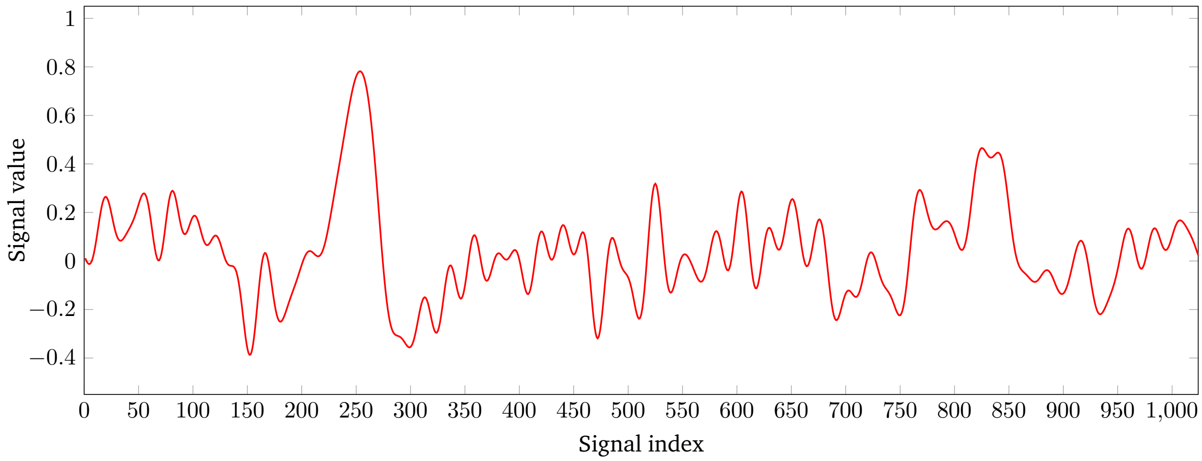


Figure 2: Solution $x_\infty$ to (5.1).

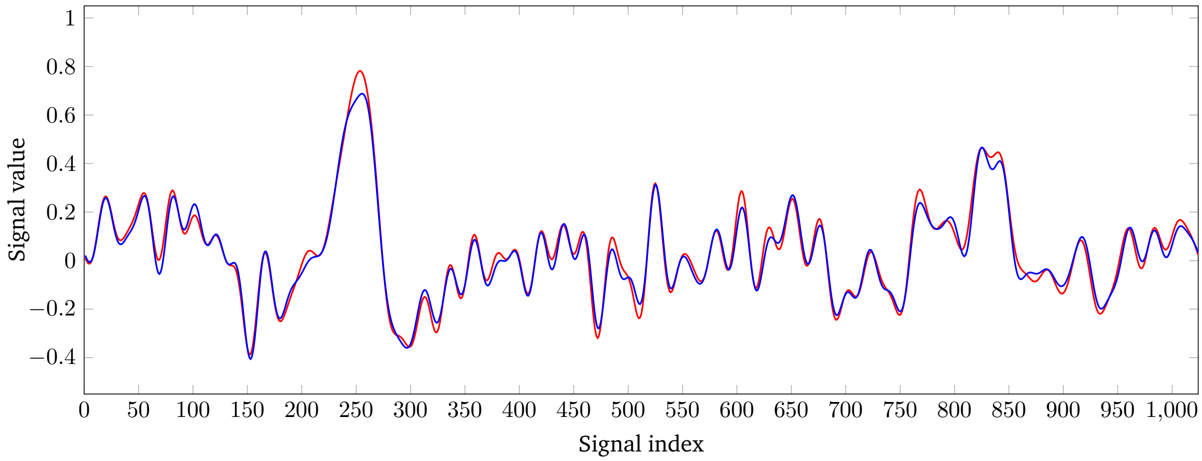


Figure 3: Solution $x_\infty$ (red) and the approximate recovery obtained with $1000$ iterations of algorithm (4.3), which exploits affine constraints (blue).

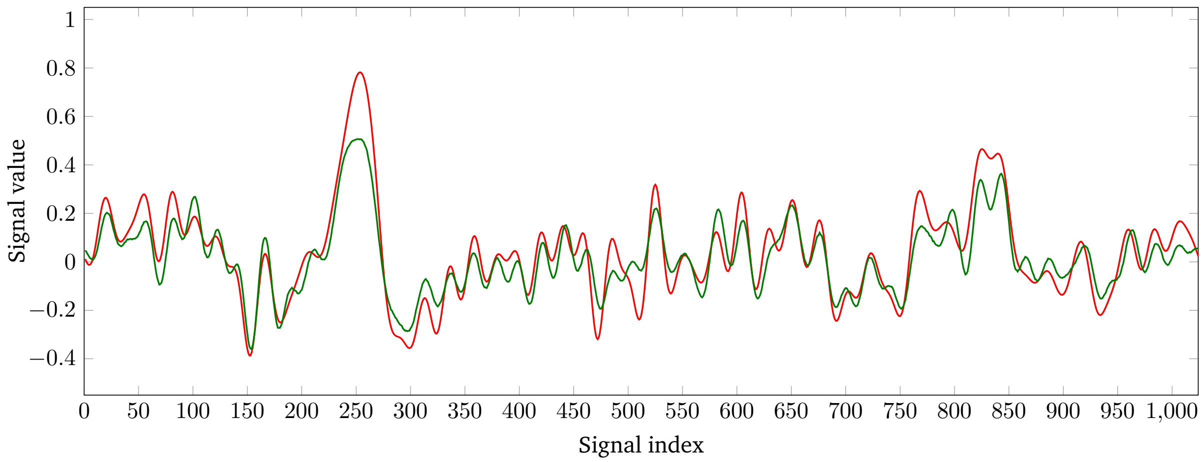


Figure 4: Solution $x_\infty$ (red) and the approximate recovery obtained with $1000$ iterations of algorithm (4.5), which does not exploit affine constraints (green).

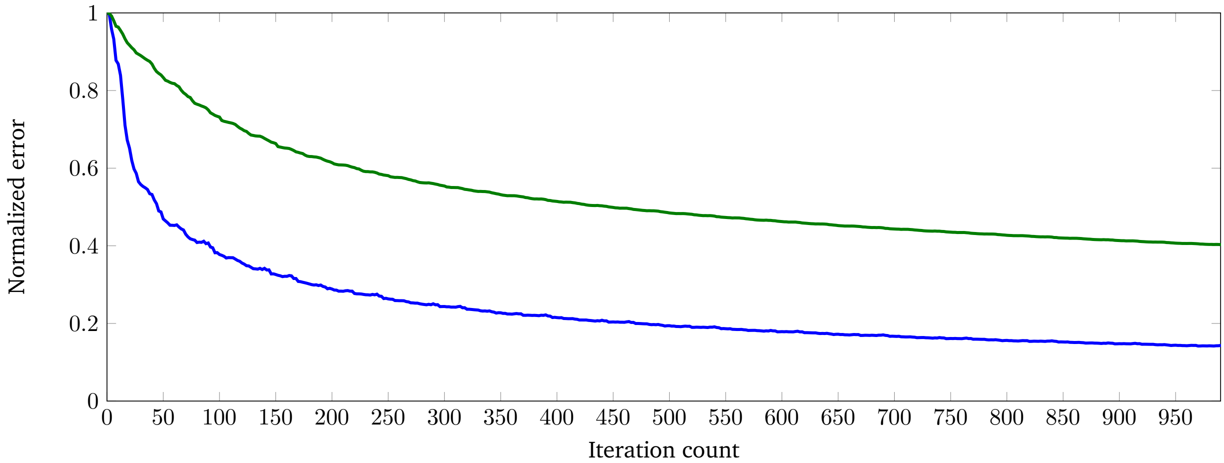


Figure 5: Normalized error $\|x_n - x_\infty\|/\|x_0 - x_\infty\|$ versus iteration count $n \in \{0, \ldots, 1000\}$ for algorithm (4.3) (blue) and algorithm (4.5) (green).

# 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.