SPARSE SIGNAL RECOVERY BY ITERATIVE PROXIMAL THRESHOLDING

Patrick L. Combettes[1] *and Jean-Christophe Pesquet*[2]

[1]Laboratoire Jacques-Louis Lions – UMR 7598, Université Pierre et Marie Curie – Paris 6, 75005 Paris, France
[2]Institut Gaspard Monge and UMR 8049, Université de Paris-Est, 77454 Marne la Vallée Cedex 2, France

ABSTRACT

Soft thresholding plays a central role in the various signal processing problems in which the ideal solution is known to possess a sparse decomposition in some orthonormal basis. Using convex-analytical tools, we extend this notion to that of proximal thresholding and investigate its properties. We then propose a versatile convex variational formulation for optimization over orthonormal bases that covers a wide range of problems, and establish the strong convergence of a proximal thresholding algorithm to solve it. Numerical applications to signal recovery are demonstrated.

1. INTRODUCTION

Throughout this paper, $\mathcal{H}$ is a separable infinite-dimensional real Hilbert space with scalar product $\langle \cdot \mid \cdot \rangle$, norm $\|\cdot\|$, and distance d. Moreover, $\Gamma_0(\mathcal{H})$ denotes the class of proper ($\not\equiv +\infty$) lower semicontinuous convex functions from $\mathcal{H}$ to $]-\infty,+\infty]$, and $(e_k)_{k\in\mathbb{N}}$ is an orthonormal basis of $\mathcal{H}$.

The standard signal denoising problem consists of recovering the original form of a signal $\overline{x}\in\mathcal{H}$ from an observation $z=\overline{x}+v$, where $v\in\mathcal{H}$ is the realization of a noise process. If $\overline{x}$ is known to admit a sparse representation in $(e_k)_{k\in\mathbb{N}}$, an estimate x can be constructed by removing the coefficients of smallest magnitude in the representation $(\langle z \mid e_k\rangle)_{k\in\mathbb{N}}$ of z with respect to $(e_k)_{k\in\mathbb{N}}$. A popular method [2, 3, 9, 10, 11] consists of soft thresholding each coefficient $\langle z \mid e_k\rangle$ at some predetermined level $\omega_k\in\,]0,+\infty[$, namely

$$x=\sum_{k\in\mathbb{N}}\operatorname{soft}_{[-\omega_k,\omega_k]}(\langle z \mid e_k\rangle)e_k, \tag{1.1}$$

where (see Fig. 1)

$$\operatorname{soft}_{[-\omega_k,\omega_k]}\colon \xi\mapsto \operatorname{sign}(\xi)\max\{|\xi|-\omega_k,0\}. \tag{1.2}$$

From an optimization point of view, the vector x exhibited in (1.1) is simply the solution to the variational problem

$$\underset{x\in\mathcal{H}}{\text{minimize}}\;\; \frac{1}{2}\|x-z\|^2+\sum_{k\in\mathbb{N}}\omega_k|\langle x \mid e_k\rangle|. \tag{1.3}$$

This formulation has been extended to the more general inverse problems in which the observation assumes the form $z=T\overline{x}+v$, where T is a bounded linear operator from $\mathcal{H}$ to some real Hilbert space $\mathcal{G}$, and where $v\in\mathcal{G}$ is the realization of a noise process. Thus, the variational problem

$$\underset{x\in\mathcal{H}}{\text{minimize}}\;\; \frac{1}{2}\|Tx-z\|^2+\sum_{k\in\mathbb{N}}\omega_k|\langle x \mid e_k\rangle| \tag{1.4}$$

has been considered in several studies, along with the soft thresholding iterations

$$x_{n+1}=\sum_{k\in\mathbb{N}}\operatorname{soft}_{[-\omega_k,\omega_k]}\big(\langle x_n+T^*(z-Tx_n) \mid e_k\rangle\big)e_k \tag{1.5}$$

to solve it (see [7] for background). The strong convergence of this algorithm was first formally established in [8] (in [7], (1.4) was analyzed in a broader framework and extended).

Theorem 1.1 [8, Theorem 3.1] *Suppose that $x_0\in\mathcal{H}$, $\inf_{k\in\mathbb{N}}\omega_k>0$, and $\|T\|<1$. Then the sequence $(x_n)_{n\in\mathbb{N}}$ generated by* (1.5) *converges strongly to a solution to* (1.4)*.*

Various considerations have led researchers to consider alternative thresholding strategies in (1.1); see, e.g., [1, 11, 12]. Is is therefore natural to ask whether such thresholding rules can be used in more general algorithms. This question is significant because the current theory of iterative thresholding [7, 8] can tackle only problems described by the variational formulation (1.4), which offers limited flexibility in the penalization of the coefficients $(\langle x \mid e_k\rangle)_{k\in\mathbb{N}}$ and which is furthermore restricted to standard linear inverse problems. The aim of the present paper is to bring out general answers to these questions. Our analysis will revolve around the following variational formulation, where σ_Ω denotes the support function of a set Ω (see Section 2).

Problem 1.2 Let $\Phi\in\Gamma_0(\mathcal{H})$, let $\mathbb{K}\subset\mathbb{N}$, let $\mathbb{L}=\mathbb{N}\smallsetminus\mathbb{K}$, let $(\Omega_k)_{k\in\mathbb{K}}$ be a sequence of closed intervals in $\mathbb{R}$, and let $(\psi_k)_{k\in\mathbb{N}}$ be a sequence in $\Gamma_0(\mathbb{R})$. The objective is to

$$\underset{x\in\mathcal{H}}{\text{minimize}}\;\; \Phi(x)+\sum_{k\in\mathbb{N}}\psi_k(\langle x \mid e_k\rangle)+\sum_{k\in\mathbb{K}}\sigma_{\Omega_k}(\langle x \mid e_k\rangle), \tag{1.6}$$

under the following technical assumptions:

i) Φ is differentiable on $\mathcal{H}$, $\inf\Phi(\mathcal{H})>-\infty$, and $\nabla\Phi$ is $1/\beta$-Lipschitz continuous for some $\beta\in\,]0,+\infty[$;
ii) for every $k\in\mathbb{N}$, $\psi_k\geq\psi_k(0)=0$;
iii) for every $k\in\mathbb{N}$, ψ_k is differentiable at 0;
iv) the functions $(\psi_k)_{k\in\mathbb{L}}$ are finite and twice differentiable on $\mathbb{R}\smallsetminus\{0\}$, and

$$(\forall\rho\in\,]0,+\infty[)(\exists\,\theta\in\,]0,+\infty[)\quad \inf_{k\in\mathbb{L}}\;\inf_{0<|\xi|\leq\rho}\psi_k''(\xi)\geq\theta;$$

v) the function $\Upsilon_{\mathbb{L}}\colon \ell^2(\mathbb{L})\to\,]-\infty,+\infty]\colon (\xi_k)_{k\in\mathbb{L}}\mapsto \sum_{k\in\mathbb{L}}\psi_k(\xi_k)$ is coercive;
vi) $0\in\operatorname{int}\bigcap_{k\in\mathbb{K}}\Omega_k$.

Problem 1.2 reduces to (1.4) when $\Phi\colon x\mapsto\|Tx-z\|^2/2$, $\mathbb{K}=\mathbb{N}$, and, for every $k\in\mathbb{N}$, $\Omega_k=[-\omega_k,\omega_k]$ and $\psi_k=0$. We shall see (Proposition 4.1) that Problem 1.2 admits at least one solution and that it covers important scenarios (Section 5.1). In addition, it lends itself to the use of a forward-backward splitting strategy (Algorithm 4.3), which consists of alternating a forward (explicit) gradient step on Φ with a backward (implicit) proximal step on

$$\Psi\colon \mathcal{H} \to \,]-\infty,+\infty] : x \mapsto \sum_{k\in\mathbb{N}} \psi_k(\langle x \mid e_k\rangle) + \sum_{k\in\mathbb{K}} \sigma_{\Omega_k}(\langle x \mid e_k\rangle). \tag{1.7}$$

Our main result (Theorem 4.5) concerns the convergence of this algorithm to a solution to Problem 1.2. Another contribution of this paper (Remark 3.3) shows that the function displayed in (1.7) is quite general in the sense that the operators on $\mathcal{H}$ that perform nonexpansive (as required by our convergence analysis) and nonincreasing (as imposed by practical considerations) thresholdings on the closed intervals $(\Omega_k)_{k\in\mathbb{K}}$ of the coefficients $(\langle x \mid e_k\rangle)_{k\in\mathbb{K}}$ of a point $x \in \mathcal{H}$ are precisely those of the form prox_Ψ, i.e., the proximity operator of Ψ. Furthermore (Proposition 3.4 and Lemma 2.4) such an operator can be decomposed as

$$\operatorname{prox}_\Psi\colon \mathcal{H} \to \mathcal{H} : x \mapsto \sum_{k\in\mathbb{K}} \operatorname{prox}_{\psi_k}\big(\operatorname{soft}_{\Omega_k} \langle x \mid e_k\rangle\big) e_k + \sum_{k\in\mathbb{L}} \operatorname{prox}_{\psi_k} \langle x \mid e_k\rangle\, e_k, \tag{1.8}$$

where we define the soft thresholder relative to $\Omega \subset \mathbb{R}$ as

$$\operatorname{soft}_\Omega : \mathbb{R} \to \mathbb{R} : \xi \mapsto \begin{cases} \xi - \underline{\omega}, & \text{if } \xi < \underline{\omega};\\ 0, & \text{if } \xi \in \Omega;\\ \xi - \overline{\omega}, & \text{if } \xi > \overline{\omega}, \end{cases} \quad \text{with } \begin{cases} \underline{\omega} = \inf\Omega,\\ \overline{\omega} = \sup\Omega. \end{cases} \tag{1.9}$$

The remainder of the paper is organized as follows. In Section 2, we provide an account of the theory of proximity operators. In Section 3, we introduce and study the notion of a proximal thresholder. Our algorithm is presented in Section 4. Numerical results are presented in Section 5.

2. PROXIMITY OPERATORS

Let us first introduce some basic notation (for a detailed account of convex analysis, see [14]). Let C be a subset of $\mathcal{H}$. The indicator function of C is

$$\iota_C : x \mapsto \begin{cases} 0, & \text{if } x \in C;\\ +\infty, & \text{if } x \notin C, \end{cases} \tag{2.1}$$

its support function is $\sigma_C : u \mapsto \sup_{x\in C} \langle x \mid u\rangle$, and its distance function is $d_C : x \mapsto \inf \|C - x\|$.

Example 2.1 *Let $\Omega \subset \mathbb{R}$ be a nonempty closed interval, let $\underline{\omega} = \inf\Omega$, let $\overline{\omega} = \sup\Omega$, and let $\xi \in \mathbb{R}$. Then*

$$\sigma_\Omega(\xi) = \begin{cases} \underline{\omega}\xi, & \text{if } \xi < 0;\\ 0, & \text{if } \xi = 0;\\ \overline{\omega}\xi, & \text{if } \xi > 0. \end{cases} \tag{2.2}$$

We now provide some some essential facts on proximity operators and refer the reader to [7] for complements.

Definition 2.2 Let $f \in \Gamma_0(\mathcal{H})$. The proximity operator of f is the operator $\operatorname{prox}_f : \mathcal{H} \to \mathcal{H}$ which maps every $x \in \mathcal{H}$ to the unique minimizer of the function $y \mapsto f(y) + \|x - y\|^2/2$.

Lemma 2.3 *Let $f \in \Gamma_0(\mathcal{H})$. Then the following hold.*

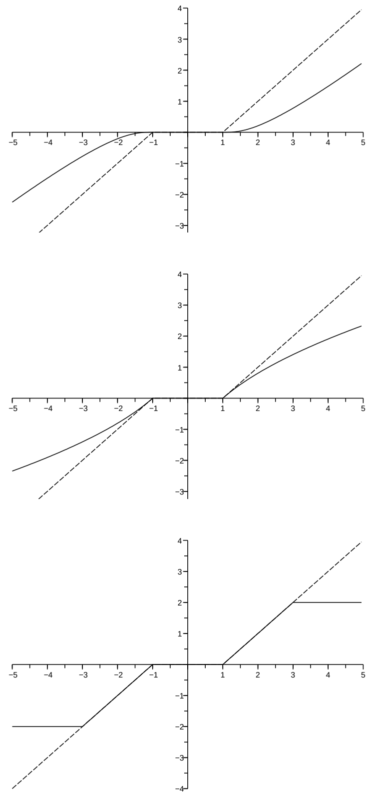

Figure 1: $\Omega = [-1,1]$. Graphs of $\operatorname{soft}_\Omega$ (dashed line) and prox_ϕ in Proposition 3.4 with $\psi = |\cdot|^{4/3}$ (top), $\psi = 0.1|\cdot|^3$ (center), and $\psi = \iota_{[-2,2]}$ (bottom).

i) $(\forall x \in \mathcal{H})\ x \in \operatorname{Argmin} f \Leftrightarrow \operatorname{prox}_f x = x$.

ii) prox_f *is nonexpansive:* $(\forall (x,y) \in \mathcal{H}^2)\ \|\operatorname{prox}_f x - \operatorname{prox}_f y\| \le \|x - y\|$.

Lemma 2.4 [7, Example 2.19] *Let $(b_k)_{k\in\mathbb{N}}$ be an orthonormal basis of $\mathcal{H}$ and let*

$$f : \mathcal{H} \to \,]-\infty,+\infty] : x \mapsto \sum_{k\in\mathbb{N}} \phi_k(\langle x \mid b_k\rangle), \tag{2.3}$$

where $(\forall k \in \mathbb{N})\ \Gamma_0(\mathbb{R}) \ni \phi_k \ge \phi_k(0) = 0$. Then $f \in \Gamma_0(\mathcal{H})$ and $(\forall x \in \mathcal{H})\ \operatorname{prox}_f x = \sum_{k\in\mathbb{N}} \operatorname{prox}_{\phi_k} \langle x \mid b_k\rangle\, b_k$.

Proposition 2.5 [5] *Let $R\colon \mathbb{R} \to \mathbb{R}$. Then R is the proximity operator of a function in $\Gamma_0(\mathbb{R})$ if and only if it is nonexpansive and nondecreasing.*

Let us now provide an important example (see Fig. 1 in the case when $\Omega = [-1,1]$). Additional examples can be found in [5] and [7].

Example 2.6 *Let $\Omega \subset \mathbb{R}$ be a nonempty closed interval, let $\underline{\omega} = \inf\Omega$, let $\overline{\omega} = \sup\Omega$, and let $\xi \in \mathbb{R}$. Then $\operatorname{prox}_{\sigma_\Omega} \xi = \operatorname{soft}_\Omega \xi$, where $\operatorname{soft}_\Omega$ is the soft thresholder defined in* (1.9).

3. PROXIMAL THRESHOLDING

The standard soft thresholder of (1.1) was seen in Example 2.6 to be a proximity operator. As such, it possesses attractive properties (see Lemma 2.3) in terms of iterative methods [7]. This remark motivates the following definition.

Definition 3.1 Let $R\colon \mathbb{R} \to \mathbb{R}$ and let $\Omega \subset \mathbb{R}$ be a nonempty closed interval. Then R is a *proximal thresholder* on Ω if there exists a function $\phi \in \Gamma_0(\mathbb{R})$ such that

$$R = \operatorname{prox}_\phi \quad \text{and} \quad (\forall x \in \mathbb{R}) \;\; Rx = 0 \;\Leftrightarrow\; x \in \Omega. \tag{3.1}$$

The following theorem characterizes all the functions $\phi \in \Gamma_0(\mathbb{R})$ for which prox_ϕ is a proximal thresholder.

Theorem 3.2 [5] *Let $\phi \in \Gamma_0(\mathbb{R})$ and let $\Omega \subset \mathbb{R}$ be a nonempty closed interval. Then the following are equivalent.*

i) prox_ϕ *is a proximal thresholder on* Ω.
ii) $\phi = \psi + \sigma_\Omega$, *where* $\psi \in \Gamma_0(\mathbb{R})$ *is differentiable at* 0 *and* $\psi'(0) = 0$.

Remark 3.3 A standard requirement for thresholders on $\mathbb{R}$ is that they be nondecreasing functions [1, 11, 12]. On the other hand, nonexpansivity is a key property to establish the convergence of iterative methods [7]. As seen in Proposition 2.5 and Definition 3.1, the nondecreasing and nonexpansive functions $R\colon \mathbb{R} \to \mathbb{R}$ that vanish only on a closed interval $\Omega \subset \mathbb{R}$ coincide with the proximal thresholders on Ω. Hence, it follows from Theorem 3.2 and Lemma 2.4 that the operators that perform a componentwise nondecreasing and nonexpansive thresholding on $(\Omega_k)_{k\in\mathbb{K}}$ of those coefficients of the decomposition in $(e_k)_{k\in\mathbb{N}}$ indexed by $\mathbb{K}$ are precisely of the form prox_Ψ, where Ψ is as in (1.7).

Next, we provide a convenient decomposition rule for implementing proximal thresholders (see Fig. 1).

Proposition 3.4 [5] *Let $\phi = \psi + \sigma_\Omega$, where $\psi \in \Gamma_0(\mathbb{R})$ and $\Omega \subset \mathbb{R}$ is a nonempty closed interval. Suppose that ψ is differentiable at* 0 *with* $\psi'(0) = 0$. *Then* $\operatorname{prox}_\phi = \operatorname{prox}_\psi \circ \operatorname{soft}_\Omega$.

4. ITERATIVE PROXIMAL THRESHOLDING

Let us start with some basic properties of Problem 1.2.

Proposition 4.1 [5] *Problem 1.2 has at least one solution.*

Proposition 4.2 [5] *Let Ψ be as in* (1.7)*, let $x \in \mathscr{H}$, and let $\gamma \in \left]0,+\infty\right[$. Then x is a solution to Problem 1.2 if and only if $x = \operatorname{prox}_{\gamma\Psi}(x - \gamma\nabla\Phi(x))$.*

Here is our algorithm for solving Problem 1.2.

Algorithm 4.3 Fix $x_0 \in \mathscr{H}$ and set, for every $n \in \mathbb{N}$,

$$\begin{aligned} x_{n+1} = \sum_{k\in\mathbb{K}} \Big(\operatorname{prox}_{\gamma_n\psi_k}\big(\operatorname{soft}_{\gamma_n\Omega_k}\langle x_n - \gamma_n\nabla\Phi(x_n) \mid e_k\rangle\big)\Big)e_k \\ + \sum_{k\in\mathbb{L}} \Big(\operatorname{prox}_{\gamma_n\psi_k}\langle x_n - \gamma_n\nabla\Phi(x_n) \mid e_k\rangle\Big)e_k, \end{aligned} \tag{4.1}$$

where $(\gamma_n)_{n\in\mathbb{N}}$ is a sequence in $\left]0,+\infty\right[$ such that $\inf_{n\in\mathbb{N}}\gamma_n > 0$ and $\sup_{n\in\mathbb{N}}\gamma_n < 2\beta$.

Remark 4.4 In Algorithm 4.3, the set $\mathbb{K}$ contains the indices of those coefficients of the decomposition in $(e_k)_{k\in\mathbb{N}}$ that are thresholded. The parameters γ_n provide added flexibility and can be used to improve the convergence profile. Finally, the operator $\operatorname{soft}_{\gamma_n\Omega_k}$ is given explicitly in (1.9).

Theorem 4.5 [5] *Every sequence generated by Algorithm 4.3 converges strongly to a solution to Problem 1.2.*

Remark 4.6 Let T be a nonzero bounded linear operator from $\mathscr{H}$ to a real Hilbert space $\mathscr{G}$, let $z \in \mathscr{G}$, and let τ and ω be in $\left]0,+\infty\right[$. Specializing Theorem 4.5 to the case when $\Phi\colon x \mapsto \|Tx - z\|^2/2$ and either

- $\mathbb{K} = \varnothing$ and $(\forall k \in \mathbb{L})\;\psi_k = \tau_k|\cdot|^p$, where $p \in \left]1,2\right]$ and $\tau_k \in \left[\tau,+\infty\right[$; or
- $\mathbb{L} = \varnothing$ and $(\forall k \in \mathbb{K})\;\psi_k = 0$ and $\Omega_k = [-\omega_k, \omega_k]$, where $\omega_k \in \left[\omega,+\infty\right[$,

yields [7, Corollary 5.19]. If we further impose $\|T\| < 1$ and $\gamma_n \equiv 1$, we obtain [8, Theorem 3.1].

5. APPLICATIONS TO SPARSE SIGNAL RECOVERY

5.1 A special case of Problem 1.2

In certain problems, q noisy linear observations are available, say $z_i = T_i\overline{x} + v_i$ $(1 \le i \le q)$, which leads to the weighted least-squares data fidelity term $x \mapsto \sum_{i=1}^q \mu_i\|T_ix - z_i\|^2$. Furthermore, signal recovery problems are typically accompanied with convex constraints that confine $\overline{x}$ to some convex sets $(S_i)_{1\le i\le m}$. These constraints can be aggregated via the cost function $x \mapsto \sum_{i=1}^m \vartheta_i d_{S_i}^2(x)$ [4]. On the other hand, a common approach to penalize the coefficients of an orthonormal basis decomposition is to use power functions, e.g., [1, 3, 8]. Moreover, we aim at promoting sparsity of a solution $x \in \mathscr{H}$ with respect to $(e_k)_{k\in\mathbb{N}}$ in the sense that, for every k in $\mathbb{K}$, we wish to set to 0 the coefficient $\langle x \mid e_k\rangle$ if it lies in the interval Ω_k. Altogether, these considerations suggest the following formulation.

Problem 5.1 For every $i \in \{1,\ldots,q\}$, let $\mu_i \in \left]0,+\infty\right[$, let T_i be a nonzero bounded linear operator from $\mathscr{H}$ to a real Hilbert space $\mathscr{G}_i$, and let $z_i \in \mathscr{G}_i$. For every $i \in \{1,\ldots,m\}$, let $\vartheta_i \in \left]0,+\infty\right[$ and let S_i be a nonempty closed and convex subset of $\mathscr{H}$. Furthermore, let $(p_{k,l})_{0\le l\le L_k}$ be distinct numbers in $\left]1,+\infty\right[$, let $(\tau_{k,l})_{0\le l\le L_k}$ be in $\left[0,+\infty\right[$, and let $l_k \in \{0,\ldots,L_k\}$ satisfy $p_{k,l_k} = \min_{0\le l\le L_k} p_{k,l}$, where $(L_k)_{k\in\mathbb{N}}$ is a sequence in $\mathbb{N}$. Finally, let $\mathbb{K} \subset \mathbb{N}$, let $\mathbb{L} = \mathbb{N} \smallsetminus \mathbb{K}$, and let $(\Omega_k)_{k\in\mathbb{K}}$ be closed intervals in $\mathbb{R}$. The objective is to

$$\begin{aligned} \underset{x\in\mathscr{H}}{\text{minimize}} \;\; \frac{1}{2}\sum_{i=1}^q \mu_i\|T_ix - z_i\|^2 + \frac{1}{2}\sum_{i=1}^m \vartheta_i d_{S_i}^2(x) \\ + \sum_{k\in\mathbb{N}}\sum_{l=0}^{L_k} \tau_{k,l}|\langle x \mid e_k\rangle|^{p_{k,l}} + \sum_{k\in\mathbb{K}} \sigma_{\Omega_k}(\langle x \mid e_k\rangle), \end{aligned} \tag{5.1}$$

under the following assumptions: $\inf_{k\in\mathbb{L}} \tau_{k,l_k} > 0$; $\inf_{k\in\mathbb{L}} p_{k,l_k} > 1$; $\sup_{k\in\mathbb{L}} p_{k,l_k} \le 2$; $0 \in \operatorname{int}\bigcap_{k\in\mathbb{K}}\Omega_k$.

Proposition 5.2 [5] *Problem 5.1 is a special case of Problem 1.2.*

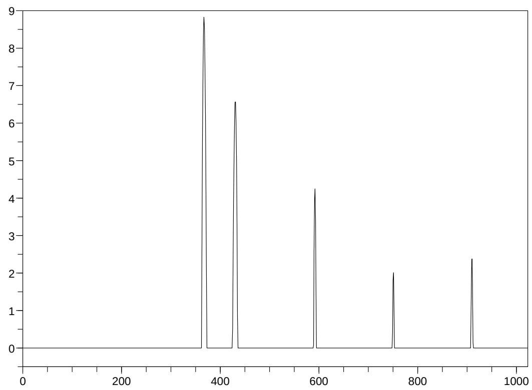

Figure 2: Original signal – Example 1.

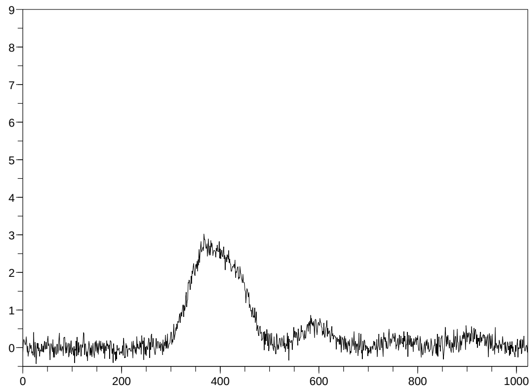

Figure 3: Degraded signal – Example 1.

5.2 First example

Our first example concerns the simulated X-ray fluorescence spectrum $\overline{x}$ displayed in Fig. 2. The measured signal z shown in Fig. 3 has undergone blurring by the limited resolution of the spectrometer and further corruption by addition of noise. In the underlying Hilbert space $\mathscr{H} = \ell^2(\mathbb{N})$, this process is modeled by $z = T\overline{x} + v$, where $T\colon \mathscr{H} \to \mathscr{H}$ is the operator of convolution with a truncated Gaussian kernel. The noise samples are uncorrelated and drawn from a $\mathscr{N}(0, 0.0225)$ Gaussian population. The original signal $\overline{x}$ has support $\{0, \ldots, N-1\}$ ($N = 1024$), takes on nonnegative values, and possesses a sparse structure. These features can be promoted in Problem 5.1 by letting $(e_k)_{k\in\mathbb{N}}$ be the canonical orthonormal basis of $\mathscr{H}$, and setting $\mathbb{K} = \mathbb{N}$, $\tau_{k,l} \equiv 0$, and

$$(\forall k \in \mathbb{N}) \quad \Omega_k = \begin{cases}]-\infty, \omega], & \text{if } 0 \leq k \leq N-1; \\ \mathbb{R}, & \text{otherwise}, \end{cases} \tag{5.2}$$

where the one-sided thresholding level is set to $\omega = 0.01$. Using the methodology described in [13], the above information about the noise can be used to construct the constraint sets $S_1 = \big\{x \in \mathscr{H} \;\big|\; \|Tx - z\| \leq \delta_1\big\}$ and $S_2 = \bigcap_{l=1}^{N-1} \big\{x \in \mathscr{H} \;\big|\; |\widehat{Tx}(l/N) - \widehat{z}(l/N)| \leq \delta_2\big\}$, where $\widehat{a}\colon \nu \mapsto \sum_{k=0}^{+\infty} \langle a \mid e_k\rangle \exp(-\imath 2\pi k\nu)$ designates the Fourier transform of $a \in \mathscr{H}$. The bounds δ_1 and δ_2 have been determined so as to guarantee that $\overline{x}$ lies in S_1 and in S_2 with a 99 percent confidence level (see [13] for details). Finally, we set $q = 0$ and $\vartheta_1 = \vartheta_2 = 1$ in (5.1). The solution produced by Algorithm 4.3 is shown in Fig. 4. It is of much better quality than the restorations obtained in [6, 13] via alternative methods.

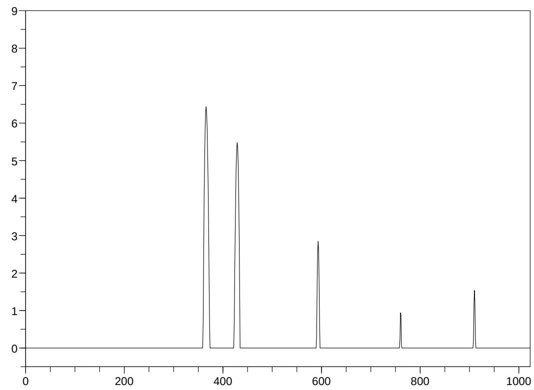

Figure 4: Signal restored by Algorithm 4.3 – Example 1.

5.3 Second example

We provide a wavelet deconvolution example in $\mathscr{H} = \mathbb{L}^2(\mathbb{R})$. The original signal $\overline{x}$ is the classical "bumps" signal displayed in Fig. 5. The degraded version shown in Fig. 6 is $z_1 = T_1\overline{x} + v_1$, where T_1 models convolution with a uniform kernel and v_1 is a realization of a zero-mean white Gaussian noise. The basis $(e_k)_{k\in\mathbb{N}}$ is an orthonormal wavelet symlet basis with 8 vanishing moments. Such wavelet bases are known to provide sparse representations for a wide class of signals such as this standard test signal. Note that there exists a strong connection between Problem 5.1 and maximum *a posteriori* techniques for estimating $\overline{x}$ in the presence of white Gaussian noise. In particular, setting $q = 1$, $m = 0$, $\mathbb{K} = \varnothing$ and $L_k \equiv 0$, and using suitably subband-adapted values of $p_{k,0}$ and $\tau_{k,0}$ amounts to fitting an appropriate generalized Gaussian prior distribution to the wavelet coefficients in each subband [1]. Such a statistical modeling is commonly used in wavelet-based estimation, where values of $p_{k,0}$ close to 2 may provide a good model at coarse resolution levels, whereas values close to 1 should be used at finer resolutions.

The setting of the more general model we adopt here is the following: in Problem 5.1, $\mathbb{K}$ and $\mathbb{L}$ are the index sets of the detail and approximation coefficients [10], respectively, and

- $(\forall k \in \mathbb{K})$ $\Omega_k = [-0.0023, 0.0023]$, $L_k = 1$, $(p_{k,0}, p_{k,1}) = (2, 4)$, $(\tau_{k,0}, \tau_{k,1}) = (0.0052, 0.0001)$.
- $(\forall k \in \mathbb{L})$ $L_k = 0$, $p_{k,0} = 2$, $\tau_{k,0} = 0.00083$.

In addition, we set $q = 1$, $\mu_1 = 1$, $m = 1$, $\vartheta_1 = 1$, and $S_1 = \big\{x \in \mathscr{H} \;\big|\; x \geq 0\big\}$. The solution x produced by Algorithm 4.3 is shown in Fig. 7. The estimation error is $\|x - \overline{x}\| = 8.33$. For comparison, the signal $\widetilde{x}$ restored via (1.4) with Algorithm (1.5) is displayed in Fig. 8. In Problem 5.1, this corresponds to $q = 1$, $m = 0$, $\mathbb{K} = \mathbb{N}$, $\tau_{k,l} \equiv 0$, $\omega_k \equiv 2.9$ for the detail coefficients, and $\omega_k \equiv 0.0062$ for the approximation coefficients. This setup yields a worse error of $\|\widetilde{x} - \overline{x}\| = 14.14$. These results have been obtained with a discrete implementation of the wavelet decomposition over 4 resolution levels using 2048 signal samples [10].

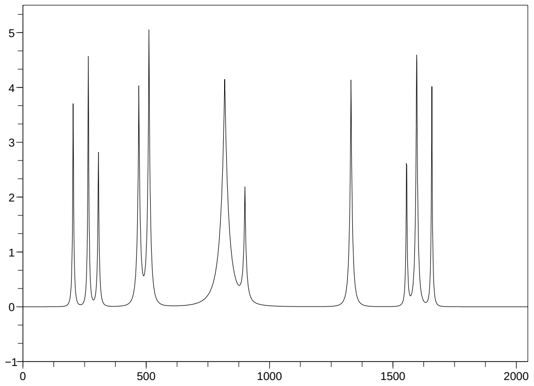

Figure 5: Original signal – Example 2.

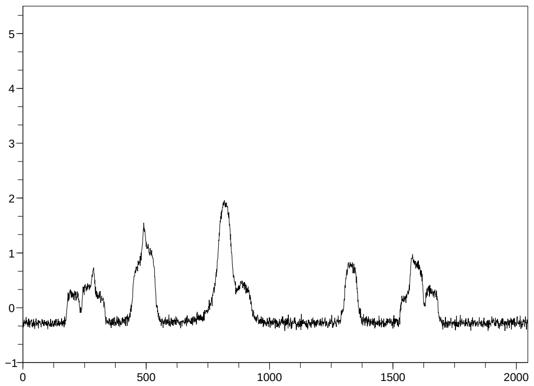

Figure 6: Degraded signal – Example 2.

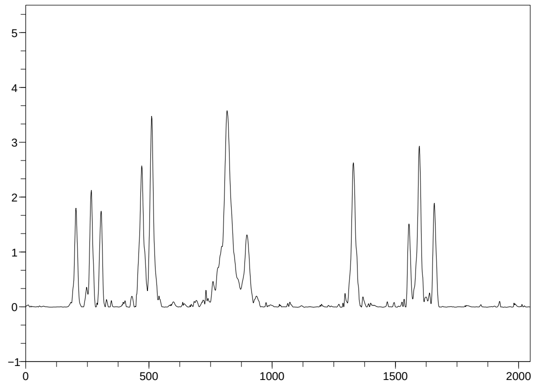

Figure 7: Signal restored by Algorithm 4.3 – Example 2.

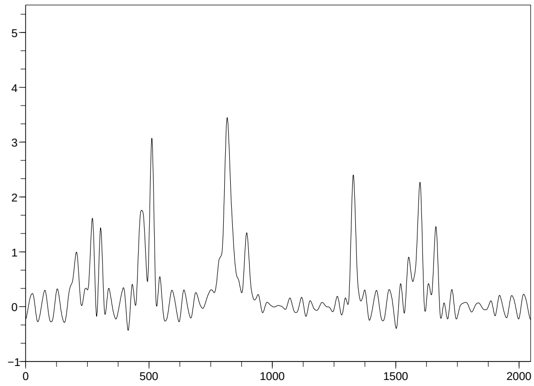

Figure 8: Signal restored by solving (1.4) – Example 2.

REFERENCES

[1] A. Antoniadis, D. Leporini, and J.-C. Pesquet, "Wavelet thresholding for some classes of non-Gaussian noise," *Statist. Neerlandica*, vol. 56, pp. 434–453, 2002.

[2] S. Bacchelli and S. Papi, "Filtered wavelet thresholding methods," *J. Comput. Appl. Math.*, vol. 164/165, pp. 39–52, 2004.

[3] 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.

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

[5] P. L. Combettes and J.-C. Pesquet, "Proximal thresholding algorithm for minimization over orthonormal bases," *SIAM J. Optim.*, to appear.

[6] P. L. Combettes and H. J. Trussell, "Method of successive projections for finding a common point of sets in metric spaces," *J. Optim. Theory Appl.*, vol. 67, pp. 487–507, 1990.

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

[8] 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.

[9] D. L. Donoho, I. M. Johnstone, G. Kerkyacharian, and D. Picard, "Wavelet shrinkage: Asymptopia?," *J. R. Statist. Soc. B.*, vol. 57, pp. 301–369, 1995.

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

[11] G. Steidl *et al.*, "On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and SIDEs," *SIAM J. Numer. Anal.*, vol. 42, pp. 686–713, 2004.

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

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

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