A DECOMPOSITION METHOD FOR NONSMOOTH CONVEX VARIATIONAL SIGNAL RECOVERY

Heinz H. Bauschke,[1] *Patrick L. Combettes,*[2] *and Jean-Christophe Pesquet*[3]

[1]Mathematics, University of British Columbia Okanagan, Kelowna, BC V1V 1V7, Canada
[2]Laboratoire Jacques-Louis Lions, UMR-CNRS 7598, Université Pierre et Marie Curie – Paris 6, 75005 Paris, France
[3]Institut Gaspard Monge and UMR–CNRS 8049, Université de Marne la Vallée, 77454 Marne la Vallée Cedex 2, France

ABSTRACT

Under consideration is the large body of signal recovery problems that can be formulated as the problem of minimizing the sum of two (not necessarily smooth) proper lower semicontinuous convex functions in a real Hilbert space. This generic problem is analyzed and a decomposition method is proposed to solve it. The convergence of the method, which is based on an extension of the Douglas-Rachford algorithm for monotone operators splitting, is established under general conditions. Various signal recovery applications are discussed and numerical results are provided.

1. INTRODUCTION

A wide array of methodological approaches have been proposed to solve signal recovery problems based on various physical, statistical, and numerical considerations, as well as on certain heuristic beliefs. Mathematically, though, signal recovery problems are most commonly posed as optimization problems and typically solved on a case-by-case basis by *ad hoc* algorithms. In [5], it was shown that a number of apparently unrelated problems fitted the following simple variational format in a real Hilbert space $\mathcal{H}$.

Problem 1 Let $f_1\colon \mathcal{H} \to \left]-\infty,+\infty\right]$ and $f_2\colon \mathcal{H} \to \mathbb{R}$ be two proper lower semicontinuous convex functions such that f_2 is differentiable on $\mathcal{H}$ with a Lipschitz continuous gradient. The objective is to minimize $f_1 + f_2$ over $\mathcal{H}$.

Problem 1 was shown to cover a variety of signal recovery formulations, including constrained least-squares problems, multiresolution sparse regularization problems, Fourier regularization problems, geometry/texture image decomposition problems, hard-constrained inconsistent feasibility problems, split feasibility problems, as well as certain maximum *a posteriori* problems [5]. Investigating this generic formulation therefore made it possible to derive existence, uniqueness, characterization, and stability results in a unified and standardized fashion. Moreover, a relaxed forward-backward algorithm was proposed to solve Problem 1, which was shown to capture, extend, and provide a simplified analysis for a variety of existing iterative methods, such as the projected Landweber method, the alternating projection method, or the iterative thresholding method recently proposed in [6] (see [5] for details).

Despite its relatively broad scope, Problem 1 fails to cover the important situations in which f_2 is differentiable with a non-Lipschitz gradient, or not finite everywhere, or when it is not differentiable at all. The latter situation includes for instance the problem of minimizing the total variation of a signal over a convex set, the problem of minimizing the sum of two set-distance functions, problems in which both functions are maxima of convex functions, and Tykhonov-like problems with L^1 norms. The objectives of the present paper are to extend Problem 1 by relaxing the assumptions on f_2 to a mere standard qualification condition, to analyze the properties of the resulting nonsmooth convex optimization problem, and to propose an iterative decomposition method based on recent developments [3] on monotone operator splitting. In Section 2, we define our notation and provide the necessary mathematical background. The recovery problem is formulated, analyzed, and illustrated in Section 3. In Section 4, an algorithm is proposed to solve it. Finally, a numerical application to wavelet-based signal recovery is presented in Section 5.

2. NOTATION AND THEORETICAL TOOLS

Throughout this paper, $\mathcal{H}$ is a real Hilbert space with scalar product $\langle \cdot \mid \cdot \rangle$, norm $\|\cdot\|$, and distance d.

2.1. Convex analysis [9]

The *indicator function* of a subset C of $\mathcal{H}$ is

$$\iota_C \colon x \mapsto \begin{cases} 0, & \text{if } x \in C; \\ +\infty, & \text{if } x \notin C, \end{cases} \tag{1}$$

and the *distance* from a point $x \in \mathcal{H}$ to C is $d_C(x) = \inf\|x - C\|$; if C is also closed and convex then, for every

1-4244-0469-X/06/$20.00 ©2006 IEEE

$x \in \mathcal{H}$, there exists a unique point $P_C x \in C$ such that $\|x - P_C x\| = d_C(x)$; $P_C x$ is the *projection* of x onto C.

The *domain* of a function $f\colon \mathcal{H} \to \,]-\infty, +\infty]$ is $\operatorname{dom} f = \{x \in \mathcal{H} \mid f(x) < +\infty\}$. $\Gamma_0(\mathcal{H})$ is the class of all lower semicontinuous convex functions from $\mathcal{H}$ to $]-\infty, +\infty]$ that are proper in the sense that their domain is nonempty. Now let $f \in \Gamma_0(\mathcal{H})$. The *conjugate* of f is the function $f^* \in \Gamma_0(\mathcal{H})$ defined by

$$(\forall u \in \mathcal{H}) \;\; f^*(u) = \sup_{x \in \mathcal{H}} \langle x \mid u \rangle - f(x) \tag{2}$$

and $f^{**} = f$. The *subdifferential* of f at $x \in \mathcal{H}$ is the set

$$\partial f(x) = \big\{u \in \mathcal{H} \,\big|\, (\forall y \in \mathcal{H}) \;\; \langle y - x \mid u \rangle + f(x) \leq f(y)\big\}. \tag{3}$$

If f is Gâteaux differentiable at x with gradient $\nabla f(x)$, then $\partial f(x) = \{\nabla f(x)\}$.

Lemma 2 *Take f_1 and f_2 in $\Gamma_0(\mathcal{H})$ such that*

$$\bigcup_{\gamma > 0} \big\{\gamma\,(x_1 - x_2) \,\big|\, x_1 \in \operatorname{dom} f_1, x_2 \in \operatorname{dom} f_2\big\} \quad \text{is a closed vector subspace}, \tag{4}$$

and let $x \in \mathcal{H}$. Then x minimizes $f_1 + f_2$ if and only if $0 \in \partial f_1(x) + \partial f_2(x)$.

Remark 3 Condition (4) is satisfied in each of the following cases.

(i) f_1 or f_2 is finite.

(ii) $\operatorname{dom} f_1 \cap \operatorname{int} \operatorname{dom} f_2 \neq \varnothing$ or $\operatorname{dom} f_2 \cap \operatorname{int} \operatorname{dom} f_1 \neq \varnothing$.

(iii) $\dim \mathcal{H} < +\infty$ and the relative interiors of $\operatorname{dom} f_1$ and $\operatorname{dom} f_2$ have a nonempty intersection.

2.2. Proximity operators [5]

Let $f \in \Gamma_0(\mathcal{H})$. Then, for every $x \in \mathcal{H}$, the function $y \mapsto f(y) + \|x - y\|^2/2$ admits a unique minimizer denoted by $\operatorname{prox}_f x$. The *proximity operator* of f is

$$\operatorname{prox}_f\colon \mathcal{H} \to \mathcal{H}\colon x \mapsto \underset{y \in \mathcal{H}}{\arg\min}\; f(y) + \frac{1}{2}\|x - y\|^2. \tag{5}$$

Lemma 4 *Let $f \in \Gamma_0(\mathcal{H})$. Then*

(i) $(\forall \gamma \in \,]0, +\infty[)(\forall x \in \mathcal{H})$
$$x = \operatorname{prox}_{\gamma f} x + \gamma \operatorname{prox}_{f^*/\gamma}(x/\gamma).$$

(ii) $(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \;\; \|\operatorname{prox}_f x - \operatorname{prox}_f y\|^2$
$$\leq \|x - y\|^2 - \|\operatorname{prox}_{f^*} x - \operatorname{prox}_{f^*} y\|^2.$$

(iii) *If f is even, then prox_f is odd.*

Example 5 Set $f = \iota_C$, where C is a nonempty closed convex subset of $\mathcal{H}$. Then $\operatorname{prox}_f = P_C$. Hence, proximity operators generalize the notion of a projection operator.

Example 6 Let C be a nonempty closed convex subset of $\mathcal{H}$, let $\gamma \in \,]0, +\infty[$, and let $x \in \mathcal{H}$. Then

$$\operatorname{prox}_{\gamma d_C} x = \begin{cases} x + \dfrac{\gamma}{d_C(x)}(P_C x - x), & \text{if } d_C(x) > \gamma; \\ P_C x, & \text{if } d_C(x) \leq \gamma. \end{cases}$$

Example 7 Let K be a (finite of infinite) subset of $\mathbb{N}$, let $(e_k)_{k \in K}$ be an orthonormal basis of $\mathcal{H}$, let $(\phi_k)_{k \in K}$ be functions in $\Gamma_0(\mathbb{R})$ such that

$$(\forall k \in K) \;\; \phi_k \geq 0 \quad \text{and} \quad \phi_k(0) = 0, \tag{6}$$

let $f\colon \mathcal{H} \to \,]-\infty, +\infty] : x \mapsto \sum_{k \in K} \phi_k(\langle x \mid e_k \rangle)$, and let $x \in \mathcal{H}$. Then $f \in \Gamma_0(\mathcal{H})$ and $\operatorname{prox}_f x = \sum_{k \in K} \pi_k e_k$, where $\pi_k = \operatorname{prox}_{\phi_k} \langle x \mid e_k \rangle$ is the unique solution to the inclusion $\langle x \mid e_k \rangle - \pi_k \in \partial \phi_k(\pi_k)$.

Example 8 Let K be a (finite of infinite) subset of $\mathbb{N}$, let $(e_k)_{k \in K}$ be an orthonormal basis of $\mathcal{H}$, let $(p_k)_{k \in K}$ be a sequence in $[1, +\infty[$, let $(\omega_k)_{k \in K}$ be a sequence in $]0, +\infty[$, let $f\colon \mathcal{H} \to \,]-\infty, +\infty] : x \mapsto \sum_{k \in K} \omega_k |\langle x \mid e_k \rangle|^{p_k}$, and let $x \in \mathcal{H}$. Then $\operatorname{prox}_f x = \sum_{k \in K} \pi_k e_k$ where, for every $k \in K$, π_k is the unique solution to

$$\begin{cases} \xi_k \in \pi_k + \omega_k\big(\operatorname{sgn}(\pi_k) + (1 - |\operatorname{sgn}(\pi_k)|)[-1, 1]\big), \\ \hfill \text{if } p_k = 1; \\ \xi_k = p_k \omega_k |\pi_k|^{p_k - 1} \operatorname{sgn}(\pi_k) + \pi_k, \quad \text{if } p_k > 1, \end{cases}$$

where $\xi_k = \langle x \mid e_k \rangle$; in particular, π_k is given by

$$\begin{cases} \operatorname{sgn}(\xi_k) \max\{|\xi_k| - \omega_k, 0\}, & \text{if } p_k = 1; \\ \xi_k - \dfrac{4\omega_k}{3 \cdot 2^{1/3}}\Big((\eta_k + \xi_k)^{1/3} - (\eta_k - \xi_k)^{1/3}\Big), & \\ \quad \text{where } \eta_k = \sqrt{\xi_k^2 + 256\omega_k^3/729}, & \text{if } p_k = \dfrac{4}{3}; \\ \xi_k + \dfrac{9\omega_k^2 \operatorname{sgn}(\xi_k)}{8}\left(1 - \sqrt{1 + \dfrac{16|\xi_k|}{9\omega_k^2}}\right), & \text{if } p_k = \dfrac{3}{2}; \\ \xi_k/(1 + 2\omega_k), & \text{if } p_k = 2. \end{cases}$$

3. PROBLEM FORMULATION

Problem 9 Let f_1 and f_2 be two functions in $\Gamma_0(\mathcal{H})$ which satisfy (4). The objective is to

$$\underset{x \in \mathcal{H}}{\operatorname{minimize}} \;\; f_1(x) + f_2(x). \tag{7}$$

Proposition 10 [2]

(i) *Existence: Problem 9 possesses at least one solution if* $\lim_{\|x\|\to+\infty} f_1(x) + f_2(x) = +\infty$.

(ii) *Uniqueness: Problem 9 possesses at most one solution if* $f_1 + f_2$ *is strictly convex, as is the case when* f_1 *or* f_2 *is strictly convex.*

(iii) *Characterization: Let* $x \in \mathcal{H}$ *and* $\gamma \in \,]0, +\infty[$, *and set* $\operatorname{rprox}_f = 2\operatorname{prox}_f - \operatorname{Id}$. *Then the following statements are equivalent.*

(a) x *solves Problem 9.*

(b) $x = \operatorname{prox}_{\gamma f_2} y$, *where* $y = \operatorname{rprox}_{\gamma f_1} \operatorname{rprox}_{\gamma f_2} y$.

The only restriction imposed in Problem 9 is the relatively mild qualification condition (4). It therefore follows from Remark 3(i) that Problem 9 subsumes Problem 1. Thus, the examples of signal recovery problems discussed in [5] are covered by Problem 9. Here are concrete scenarios which are covered by Problem 9, but not by Problem 1.

Example 11 $\mathcal{H}$ is either or $\mathbb{R}^N$ or $H^1(\Omega)$, where Ω is an open bounded domain of $\mathbb{R}^m$, f_1 is the L^1 norm or the indicator function of a nonempty closed convex set, and f_2 is the total variation. This follows from [4, Proposition 1].

Example 12 C_1 and C_2 are nonempty closed convex sets, $\alpha > 0$, $1 \leq p < +\infty$, $f_1 = \alpha d_{C_1}^p$, and $f_2 = d_{C_2}$. Problem 9 then extends the standard convex feasibility problem, which corresponds to the case when $C_1 \cap C_2 \neq \varnothing$ [8].

Example 13 $\mathcal{H} = \mathrm{L}^2(\Omega)$, where Ω is an open bounded domain of $\mathbb{R}^m$, and the observed data assume the form $z = L\overline{x} + w$, where L is a bounded linear operator from $\mathcal{H}$ to a Hilbert space $\mathcal{G}$ and $w \in \mathcal{G}$ is additive noise. Moreover, $f_1 \colon x \mapsto \|Lx - z\|_{\mathrm{L}^1}$ and $f_2 = \alpha\|\cdot\|_{\mathrm{L}^1}$, with $\alpha > 0$.

4. ALGORITHM

Theorem 14 [2] *Let* $\gamma \in \,]0, +\infty[$, *let* $(\lambda_n)_{n\in\mathbb{N}}$ *be a sequence in* $]0, 2[$, *and let* $(a_n)_{n\in\mathbb{N}}$ *and* $(b_n)_{n\in\mathbb{N}}$ *be sequences in* $\mathcal{H}$. *Suppose that Problem 9 admits at least one solution,* $\sum_{n\in\mathbb{N}} \lambda_n(2 - \lambda_n) = +\infty$, *and* $\sum_{n\in\mathbb{N}} \lambda_n(\|a_n\| + \|b_n\|) < +\infty$. *Take* $x_0 \in \mathcal{H}$ *and set, for every* $n \in \mathbb{N}$,

$$\begin{cases} x_{n+\frac{1}{2}} = \operatorname{prox}_{\gamma f_2} x_n + b_n \\ x_{n+1} = x_n + \\ \qquad \lambda_n\Big(\operatorname{prox}_{\gamma f_1}\big(2x_{n+\frac{1}{2}} - x_n\big) + a_n - x_{n+\frac{1}{2}}\Big). \end{cases} \tag{8}$$

Then $(x_n)_{n\in\mathbb{N}}$ *converges weakly to some point* $y \in \mathcal{H}$ *and* $x = \operatorname{prox}_{\gamma f_2} y$ *is a solution to Problem 9.*

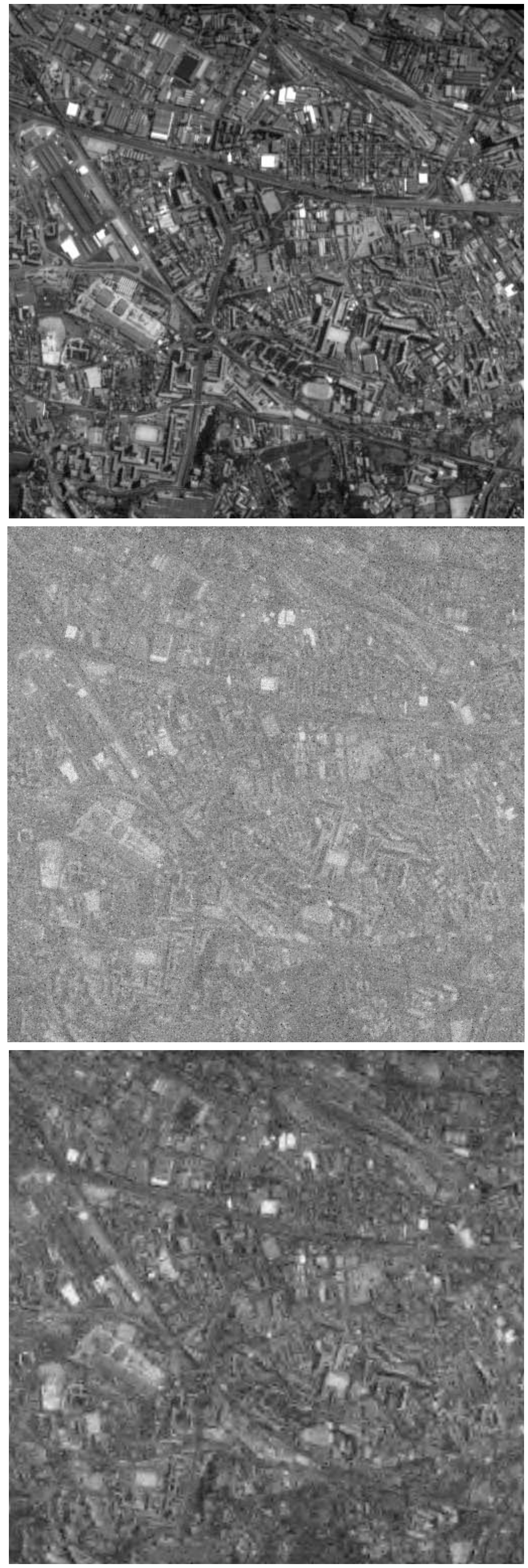

Fig. 1. Top: original image; middle: noisy image; bottom: denoised image.

In (8), the minimization problem (7) is decomposed into two main steps: the current iterate is x_n and the function f_2 is first utilized to compute $x_{n+\frac{1}{2}}$; the function f_1 is then utilized to produce the update x_{n+1}. Note that the algorithm allows for the inexact implementation of these two proximal steps via the incorporation of errors b_n and a_n. Moreover, a variable relaxation parameter λ_n gives added flexibility. It is important to note that the solution to Problem 9 is obtained as the image under $\mathrm{prox}_{\gamma f_2}$ of the weak limit of $(x_n)_{n\in\mathbb{N}}$ and that, in general, little is known about the asymptotic behavior of $\big(\mathrm{prox}_{\gamma f_2} x_n\big)_{n\in\mathbb{N}}$ unless $\mathrm{prox}_{\gamma f_2}$ is weakly continuous. This rather stringent assumption is notably satisfied when $\dim \mathcal{H} < +\infty$ (by continuity of $\mathrm{prox}_{\gamma f_2}$, see Lemma 4(ii)); we can then formulate a sharper convergence result, which is immediately relevant in numerical computations with discretized data.

Corollary 15 *If* $\dim \mathcal{H} < +\infty$ *in Theorem 14, then* $\big(\mathrm{prox}_{\gamma f_2} x_n\big)_{n\in\mathbb{N}}$ *converges to a solution to Problem 9.*

5. NUMERICAL RESULTS

We consider the denoising of an $N \times N$ image, where $N = 512$. The underlying Hilbert space $\mathcal{H}$ is the Euclidean space $\mathbb{R}^{N^2}$. The original image shown in Fig. 1 (top) has been corrupted by addition of i.i.d. zero-mean Laplacian noise. The degraded image $z = (z^{i,j})_{1\le i,j\le N}$ can be seen in Fig. 1 (middle). The image-to-noise ratio is 5.95 dB.

The denoising is performed by solving Problem 9 where

$$f_1(x) = f_1\big((x^{i,j})_{1\le i,j\le N}\big) = \sum_{i=1}^{N}\sum_{j=1}^{N} |x^{i,j} - z^{i,j}| \qquad (9)$$

and f_2 is as in Example 8 with $K = \{1,\ldots,N^2\}$. In our case, $(e_k)_{k\in K}$ is chosen to be a two-dimensional separable orthonormal wavelet basis. More precisely, we use a dyadic wavelet decomposition with symlet filters of length 8, carried out over 4 resolution levels. Moreover, p_k takes its values in $\{1, 4/3, 3/2, 2\}$ and, for each subband, an adapted value of (ω_k, p_k) is selected. Such a problem formulation is closely related to a maximum *a posteriori* approach using i.i.d. generalized Gaussian prior distributions for the wavelet coefficients (see [1] for more details about this statistical model). A similar approach has been adopted in [6] for restoration problems involving Gaussian noise. Let us emphasize that, due to the nondifferentiability of f_1 and f_2, neither the algorithms proposed in [6] nor those of [5] are applicable. The denoised image obtained by using Algorithm (8) is shown in Fig. 1 (bottom). The relative mean square error with respect to the original image is 14.15 dB. The normalized error $\|\mathrm{prox}_{\gamma f_2} x_n - \mathrm{prox}_{\gamma f_2} x_\infty\| / \|\mathrm{prox}_{\gamma f_2} x_\infty\|$ is plotted in Fig. 2.

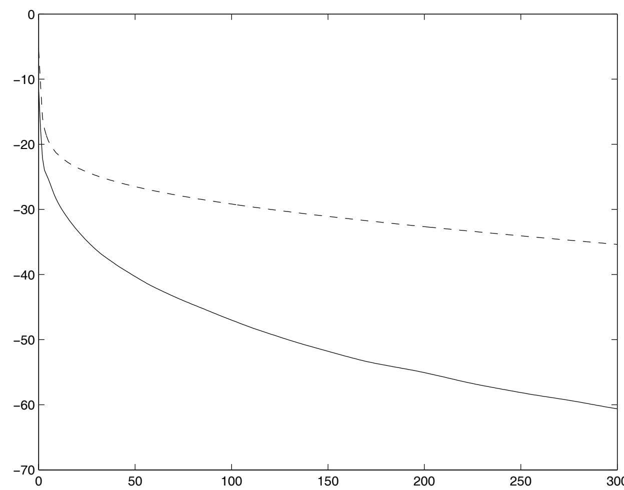

Fig. 2. Normalized error in dB: Proposed method with $\gamma = 50$ and $\lambda_n \equiv 1$ (solid line); comparison with the standard subgradient method [7, Section 2.2] (dashed line).

6. 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] H. H. Bauschke, P. L. Combettes, and J.-C. Pesquet, "Nonsmooth variational signal recovery via operator splitting." Technical report, 2005.

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

[4] 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.

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

[6] 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.

[7] N. Z. Shor, *Minimization Methods for Non-Differentiable Functions.* New York: Springer-Verlag, 1985.

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

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