A CONSTRAINED FORWARD-BACKWARD ALGORITHM FOR IMAGE RECOVERY PROBLEMS

Nelly Pustelnik, Caroline Chaux and Jean-Christophe Pesquet

Université Paris-Est
Institut Gaspard Monge, UMR CNRS 8049
77454 Marne-la-Vallée Cedex, France
e-mail: {nelly.pustelnik,caroline.chaux,jean-christophe.pesquet}@univ-paris-est.fr

ABSTRACT

In the solution of inverse problems, the objective is often to minimize the sum of two convex functions f and g subject to convex constraints. Recently, many works have been devoted to this problem in the *unconstrained case*, when f is possibly non-smooth and g is differentiable with a Lipschitz-continuous gradient. The use of a non-smooth penalizing function arises in particular in wavelet regularization techniques in connection with sparsity issues. In this paper, we propose a modification of the standard forward-backward algorithm, which allows us to minimize $f+g$ over a convex constraint set C. The effectiveness of the proposed approach is illustrated in an image restoration problem involving signal-dependent noise.

1. INTRODUCTION

In a wide range of signal and image processing problems, one needs to recover original data from observations having some a priori information on the data. The prior knowledge may consist of a model linking the observations to the original data, some probabilistic information on the data (e.g. prior probability distribution) and/or some constraints which are satisfied by these data. Of particular interest are problems where the observations are obtained from the data by a linear operator T and some addition of noise. Restoration problems correspond to the case when T is a convolution operator and denoising problems are obtained when $T = \mathrm{Id}$.

In order to solve such inverse problems, many approaches have been developed since the 1960s. In particular, successfull methods have been proposed which take advantage of a wavelet (or x-let) representation of the data.

In particular, a variational approach can be adopted, aiming at minimizing the sum of two functions f and g over a convex set C in the wavelet domain. The resulting criterion can often be related to Maximum A Posteriori (MAP) estimation as will be shown in Section 5. The first function f may be non-smooth. For example, it may correspond to a ℓ^1-norm so as to promote the sparsity of the wavelet representation of the data. Unlikely, the second function g is often differentiable with a Lipschitz-continuous gradient. It usually corresponds to some "distance" function between the observations and the data transformed by T. The convex set C allows to further model some desirable properties of the solution (e.g. positivity or energy bound). For tractability, in this paper, it is assumed that the functions f and g are in the class $\Gamma_0(\mathscr{H})$ of lower semi-continuous convex functions taking their values in $]-\infty,+\infty]$ (not identically equal to $+\infty$), which are defined on a real separable Hilbert space $\mathscr{H}$. Then, our objective is to solve the following:

Problem 1.1 Let C be a nonempty closed convex subset of $\mathscr{H}$. Let f and g be in $\Gamma_0(\mathscr{H})$, where g is differentiable on $\mathscr{H}$ with a β-Lipschitz continuous gradient for some $\beta \in\,]0,+\infty[$.

$$\text{Find} \qquad \min_{x\in C} f(x)+g(x). \tag{1}$$

Problem 1.1 is obviously equivalent to minimize $f+g+\iota_C$, where ι_C denotes the indicator function of C, *i.e.*,

$$(\forall x \in \mathscr{H}) \quad \iota_C(x) = \begin{cases} 0, & \text{if } x \in C; \\ +\infty, & \text{otherwise.} \end{cases} \tag{2}$$

In the specific case when $C = \mathscr{H}$, this problem has recently received much attention. Forward-backward algorithms (which cover so-called thresholded Landweber algorithms [1]) have been proposed to solve it numerically. The idea was proposed by Daubechies *et al.* in [2] for a quadratic function g and it was formulated in a more general convex analysis framework in [3]. Extensions to frame-based inverse problems can be found in [4] and theoretical connections with thresholding rules are further investigated in [5]. Attention was also paid to the improvement of the convergence speed of the forward-backward algorithm [6]. In [7], an accelerated method was proposed in the specific case of a deconvolution in a Shannon wavelet basis. In [8], a Douglas-Rachford algorithm was introduced to relax the assumption of differentiability of g. Finally, a modification of the forward-backward algorithm for restoration problems involving Poisson data was proposed in [9].

As shown by the thorough discussion in [3], a fundamental tool to analyze Problem 1.1 is the *proximity operator*. This operator was introduced by Moreau in 1962 [10]. The proximity operator of $f \in \Gamma_0(\mathscr{H})$ is $\mathrm{prox}_f : \mathscr{H} \to \mathscr{H}$; $x \mapsto \arg\min_{y\in\mathscr{H}} \frac{1}{2}\|y-x\|^2 + f(y)$. We thus see that prox_{ι_C} reduces to the projection P_C onto the convex set C. Conceptually, by gathering the terms f and ι_C in the previous formulation, the standard form of the forward-backward algorithm could be used. However, this would require a closed form expression of $\mathrm{prox}_{\iota_C+f}$, which is only available in some specific situations.

In Section 2, we recall some properties of the proximity operator. It is shown that in general the proximity operator of $\iota_C + f$ cannot be easily expressed. In Section 3, we subsequently propose an iterative approach to compute this operator for any function $f \in \Gamma_0(\mathscr{H})$. In Section 4, we describe the constrained forward-backward algorithm we introduce in order to solve Problem 1.1. An application of this approach to a wavelet-based restoration problem is described in Section 5. The problem of interest is made difficult by the presence of signal-dependent noise. In this context, we propose to realize a quadratic lower approximation of the considered MAP criterion to further improve the convergence profile of the algorithm.

2. SOME PROPERTIES OF PROXIMITY OPERATORS

As already mentioned, the proximity operator of $\iota_C + f$ plays a key role in our approach.

Some useful results for the calculation of $\operatorname{prox}_{\iota_C+f}$ are first recalled:

Proposition 2.1 [8, Proposition 12] *Let $f \in \Gamma_0(\mathscr{H})$ and let C be a closed convex subset of $\mathscr{H}$ such that $C \cap \operatorname{dom} f \neq \emptyset$.*[1] *Then the following properties hold.*

(i) $(\forall x \in \mathscr{H})$, $\operatorname{prox}_f x \in C \Rightarrow \operatorname{prox}_{\iota_C+f} x = \operatorname{prox}_f x$

(ii) *Suppose that $\mathscr{H} = \mathbb{R}$. Then*

$$\operatorname{prox}_{\iota_C+f} = P_C \circ \operatorname{prox}_f. \tag{3}$$

Note that, the second part of this proposition can be easily generalized, yielding to the following result which appears also an extension of [4]:

Proposition 2.2 *Let $\mathbb{K}$ be a nonempty subset of $\mathbb{N}$, let $(o_k)_{k\in\mathbb{K}}$ be an orthonormal basis of $\mathscr{H}$ and let $(\varphi_k)_{k\in\mathbb{K}}$ be functions in $\Gamma_0(\mathbb{R})$. Set*

$$f\colon \mathscr{H} \to \left]-\infty,+\infty\right] : x \mapsto \sum_{k\in\mathbb{K}} \varphi_k(\langle x \mid o_k\rangle). \tag{4}$$

Let

$$C = \bigcap_{k\in\mathbb{K}} \{x \in \mathscr{H} \mid \langle x \mid o_k\rangle \in C_k\} \tag{5}$$

where $(C_k)_{k\in\mathbb{K}}$ are nonempty closed intervals in $\mathbb{R}$ such that $(\forall k \in \mathbb{K})\ C_k \cap \operatorname{dom}\varphi_k \neq \emptyset$.
Suppose that either $\mathbb{K}$ is finite, or there exists a subset $\mathbb{L}$ of $\mathbb{K}$ such that:

(i) *$\mathbb{K} \smallsetminus \mathbb{L}$ is finite;*

(ii) *$(\forall k \in \mathbb{L})\ \varphi_k \geq \varphi_k(0) = 0$ and $0 \in C_k$.*

Then

$$(\forall x \in \mathscr{H}) \quad \operatorname{prox}_{\iota_C+f} x = \sum_{k\in\mathbb{K}} \pi_k o_k \tag{6}$$

where

$$\pi_k = \begin{cases} \inf C_k & \text{if } \operatorname{prox}_{\varphi_k} \langle x \mid o_k\rangle < \inf C_k \\ \sup C_k & \text{if } \operatorname{prox}_{\varphi_k} \langle x \mid o_k\rangle > \sup C_k \\ \operatorname{prox}_{\varphi_k} \langle x \mid o_k\rangle & \text{otherwise.} \end{cases} \tag{7}$$

A function f (resp. convex C) satisfying (4) (resp. (5)) will be said *separable*. Note that (6) and (7) imply that (3) holds. However, this relation has been proved under the restrictive assumption that both f *and* C are separable. In general, when either f *or* C is not separable, (3) is no longer valid. Let us give two simple counterexamples to illustrate this fact.

[1] The domain of a function $f : \mathscr{H} \to]-\infty,+\infty]$ is $\operatorname{dom} f = \{x \in \mathscr{H} \mid f(x) < +\infty\}$.

Example 2.3 Let $\mathscr{H} = \mathbb{R}^2$ and let f be the function defined by $(\forall x \in \mathbb{R}^2)\ f(x) = \frac{1}{2}x^\top \Lambda x$ with $\Lambda = \begin{pmatrix} 1 & \Lambda_{1,2} \\ \Lambda_{1,2} & \Lambda_{2,2}\end{pmatrix}$ where $\Lambda_{2,2} \geq 0$ and $|\Lambda_{1,2}| \leq \Lambda_{2,2}^{1/2}$. Let $C = [-1,1]^2$. This convex set is obviously separable w.r.t. the canonical basis of $\mathbb{R}^2$. Now, set $x = 2(\Lambda_{1,2}, 1+\Lambda_{2,2})^\top$. After some calculations, we obtain:

- $P_C(\operatorname{prox}_f x) = (0,1)^\top$
- $\operatorname{prox}_{\iota_C+f} x = (\pi,1)^\top$ where

$$\pi = \begin{cases} \frac{\Lambda_{1,2}}{2} & \text{if } \Lambda_{1,2} \in [-2,2] \\ 1 & \text{if } \Lambda_{1,2} > 2 \\ -1 & \text{if } \Lambda_{1,2} < -2. \end{cases} \tag{8}$$

We conclude that (3) is not satisfied as soon as $\Lambda_{1,2} \neq 0$, that is f is not separable.

Example 2.4 Let $\mathscr{H} = \mathbb{R}^2$. Consider the separable function defined by $(\forall x = (\xi_1,\xi_2)^\top \in \mathbb{R}^2)\ f(x) = (1+\Lambda_{1,2})\xi_1^2 + (1-\Lambda_{1,2})\xi_2^2$ where $|\Lambda_{1,2}| \leq 1$. Let us now consider the nonseparable convex set given by

$$C = \{x = (\xi_1,\xi_2)^\top \in \mathbb{R}^2 \mid \max(|\xi_1+\xi_2|, |\xi_1-\xi_2|) \leq \sqrt{2}\}. \tag{9}$$

In this case, it can also be shown that (3) does not hold.

In summary, for an arbitrary function $f \in \Gamma_0(\mathscr{H})$ and an arbitrary closed convex set C, we cannot trust (3) to determine $\operatorname{prox}_{\iota_C+f}$. In the next section, we propose an efficient approach to compute the desired proximity operator in a general setting.

3. COMPUTATION OF $\operatorname{prox}_{\iota_C+f}$

Let us first recall that the Douglas-Rachford algorithm provides an appealing numerical solution to the minimization of the sum of two convex functions f_1 and f_2. More precisely, we have:

Proposition 3.1 (see [8]) *Let f_1 and f_2 in $\Gamma_0(\mathscr{H})$ satisfy $\operatorname{int}(\operatorname{dom} f_1) \cap \operatorname{dom} f_2 \neq \emptyset$.*[2] *Set $z_0 \in \mathscr{H}$ and construct, for all $m \in \mathbb{N}$,*

$$\begin{cases} z_{m+\frac{1}{2}} = \operatorname{prox}_{\kappa f_2}(z_m) \\ z_{m+1} = z_m + \tau\big(\operatorname{prox}_{\kappa f_1}(2z_{m+\frac{1}{2}} - z_m) - z_{m+\frac{1}{2}}\big), \end{cases} \tag{10}$$

where $\kappa > 0$, $\tau \in]0,2[$. Then, $(z_m)_{m\in\mathbb{N}}$ converges weakly to $z \in \mathscr{H}$ such that $\operatorname{prox}_{\kappa f_2}(z)$ is a minimizer of $f_1 + f_2$.

On the other hand, a way to compute $\operatorname{prox}_{\iota_C+f}$ is to come back to its definition: $(\forall x \in \mathscr{H})$,

$$\operatorname{prox}_{\iota_C+f}(x) = \arg\min_{y\in\mathscr{H}} \frac{1}{2}\|y-x\|^2 + \iota_C(y) + f(y). \tag{11}$$

To solve the above minimization problem, we propose to use the Douglas-Rachford algorithm by setting $f_1 = f$ and $f_2 = \frac{1}{2}\|\cdot - x\|^2 + \iota_C$. Note that both $\operatorname{prox}_{\kappa f_1}$ and $\operatorname{prox}_{\kappa f_2}$ with

[2] The interior of $\operatorname{dom} f$ is designated by $\operatorname{int}(\operatorname{dom} f)$.

$\kappa > 0$, must be computed to apply this algorithm. In our case, we have

$$\mathrm{prox}_{\kappa f_1}(z) = \mathrm{prox}_{\kappa f}(z) \quad (12)$$

and

$$\mathrm{prox}_{\kappa f_2}(z) = P_C\Big(\frac{z+\kappa x}{1+\kappa}\Big). \quad (13)$$

The resulting Douglas-Rachford algorithm enjoys the following properties:

Proposition 3.2 *Assume that* $\mathrm{int}(\mathrm{dom} f) \cap C \neq \emptyset$. *Consider the algorithm given by* (10), (12) *and* (13) *with* $\tau = \kappa = 1$ *and* $z_0 = 2\mathrm{prox}_f x - x$. *Then*

(i) $(z_{m+\frac{1}{2}})_{m\in\mathbb{N}}$ *converges strongly to* $\mathrm{prox}_{\iota_C+f} x$;

(ii) $\mathrm{prox}_f x \in C \Rightarrow (\forall m \in \mathbb{N})$, $z_{m+\frac{1}{2}} = \mathrm{prox}_{\iota_C+f} x$.

The second property shows that the proposed algorithm converges in one iteration when $\mathrm{prox}_f x \in C$. This appears quite consistent with Proposition 2.1(i).

4. PROPOSED MINIMIZATION METHOD

4.1 Forward-backward approach

Let us now turn our attention to Problem 1.1 which is subsequently assumed to admit a solution. Such a solution can be computed by the forward-backward algorithm.

Let $x_0 \in \mathscr{H}$ be an initial value. The algorithm constructs a sequence $(x_n)_{n\geq 1}$ by the iteration: for every $n \in \mathbb{N}$,

$$x_{n+1} = x_n + \lambda_n\big(\mathrm{prox}_{\iota_C+\gamma_n f}(x_n - \gamma_n \nabla g(x_n)) + a_n - x_n\big) \quad (14)$$

where $\gamma_n > 0$ is the algorithm step-size, $\lambda_n > 0$ is a relaxation parameter and a_n represents an error allowed in the computation of the proximity operator.

The weak convergence of $(x_n)_{n\in\mathbb{N}}$ to a solution to Problem 1.1 is then guaranteed provided that:

Assumption 4.1

(i) $0 < \inf_{n\in\mathbb{N}} \gamma_n < \sup_{n\in\mathbb{N}} \gamma_n < 2\beta^{-1}$

(ii) $(\forall n \in \mathbb{N})\ \lambda_n \in]0,1]$ and $\inf_{n\in\mathbb{N}} \lambda_n > 0$

(iii) $\sum_{n=0}^{+\infty} \|a_n\| < +\infty$.

More details concerning this algorithm are given in [3, 4] and additional conditions for the strong convergence of the algorithm can be found in [5].

To implement (14), we see however that $\mathrm{prox}_{\iota_C+\gamma_n f}$ needs to be computed, which in general, is a non trivial problem as pointed out in the previous sections.

4.2 Algorithm

Let us summarize the complete form of the algorithm we propose to solve Problem 1.1. In the following, $(\gamma_n)_{n\in\mathbb{N}}$ and $(\lambda_n)_{n\in\mathbb{N}}$ are sequences satisfying Assumption 4.1(i) and (ii).

Algorithm 4.1

① Set $x_0 \in C$.

② Set $x'_n = x_n - \gamma_n \nabla g(x_n)$.

③ Set $z_{n,0} = 2\mathrm{prox}_{\gamma_n f} x'_n - x'_n$.

④ For $m = 0, \dots, M_n$

a) Compute $z_{n,m+\frac{1}{2}} = P_C\Big(\frac{z_{n,m} + x'_n}{2}\Big)$.

b) If $z_{n,m+\frac{1}{2}} = z_{n,m-\frac{1}{2}}$, goto ⑤.

c) Compute $z_{n,m+1} = z_{n,m} + \mathrm{prox}_{\gamma_n f}(2z_{n,m+\frac{1}{2}} - z_{n,m}) - z_{n,m+\frac{1}{2}}$.

⑤ Set $x_{n+1} = x_n + \lambda_n\big(z_{n,m+\frac{1}{2}} - x_n\big)$.

⑥ Goto ②.

We see that Step ④ consists of at most M_n iterations of the Douglas-Rachford algorithm described in the previous section. Steps ② and ⑤ correspond to the forward-backward iteration where the error term is $a_n = z_{n,M_n+\frac{1}{2}} - \mathrm{prox}_{\iota_C+\gamma_n f} x_n$. Due to Proposition 3.2(i), when $\mathrm{int}(\mathrm{dom} f) \cap C \neq \emptyset$, we are guaranteed that, by choosing M_n large enough, Assumption 4.1(iii) can be satisfied, so that Algorithm 4.1 weakly converges to a solution to Problem 1.1. Note however that we do not need to take large values of M_n when n is small.

In addition, since we have chosen x_0 in C and $(\forall n \in \mathbb{N})$ $z_{n,m+\frac{1}{2}}$ also lies in C, it can be deduced that $(\forall n \geq 1)$ $x_n \in C$. Consequently, in Step ②, the gradient of g is only evaluated on C. This means that the assumption of Lipschitz-continuity on the gradient of g is only required on C.

5. APPLICATION TO IMAGE RESTORATION

5.1 Context

We aim at restoring an image $\overline{y}$ in a real separable Hilbert space $\mathscr{G}$ corrupted by a linear operator $T : \mathscr{G} \to \mathscr{G}$ and an additive noise $w \in \mathscr{G}$, having an observation

$$z = T\overline{y} + w. \quad (15)$$

Here, digital images of size $N_1 \times N_2$ are considered and thus $\mathscr{G} = \mathbb{R}^N$ with $N = N_1 N_2$. In addition, T is a convolutive blur and $w = (w^{(i)})_{1\leq i\leq N}$ is a realization of an independent zero-mean Gaussian noise $W = (W^{(i)})_{1\leq i\leq N}$. The variance of each random variable $W^{(i)}$ is signal-dependent and is equal to $\sigma^2((T\overline{y})^{(i)})$ where $T\overline{y} = \big((T\overline{y})^{(i)}\big)_{1\leq i\leq N}$ and

$$(\forall \mu \in [\delta, +\infty[) \qquad \sigma^2(\mu) = \alpha_1 \mu + \alpha_0 \quad (16)$$

with $\alpha_0 \geq 0$, $\alpha_1 > 0$ and $\delta \in]-\alpha_0/\alpha_1, +\infty[$. Hereabove, it has been assumed that $T\overline{y} \in [\delta, +\infty[^N$.

A both simple and efficient prior probabilistic model on the unknown image $\overline{y}$ can be adopted by using a representation of this image in a frame. We thus use a linear representation of the form: $\overline{y} = F^*\overline{x}$, where $F^* : \mathbb{R}^K \to \mathbb{R}^N$ $(K \geq N)$ is a frame synthesis operator. We then assume that the vector $\overline{x}$ of frame coefficients is a realization of a random vector $\overline{X}$ with independent components. Each component $\overline{X}^{(k)}$, $1 \leq k \leq K$, has a probability density $\exp(-\phi_k(\cdot))/\int_{-\infty}^{+\infty} \exp(-\phi_k(\eta))\, d\eta$ where ϕ_k is a finite function in $\Gamma_0(\mathbb{R})$.

In addition, we assume that we have prior information on $\overline{x}$ which can be expressed by the fact that $\overline{x}$ belongs to a nonempty closed convex set C of $\mathbb{R}^K$.

With these assumptions, it can be shown that a MAP estimate of the vector of frame coefficients $\overline{x}$ can be obtained from $z = \big(z^{(i)}\big)_{1\leq i\leq N}$ by minimizing in the Hilbert space $\mathscr{H} = \mathbb{R}^K$ the function $f + g + \iota_C$ where

$$(\forall x = \big(x^{(k)}\big)_{1\leq k\leq K} \in \mathbb{R}^K) \quad f(x) = \sum_{k=1}^{K} \phi_k\big(x^{(k)}\big) \quad (17)$$

and $g = \Psi \circ T \circ F^*$ with

$$(\forall u = (u^{(i)})_{1 \le i \le N} \in \mathbb{R}^N) \quad \Psi(u) = \sum_{i=1}^{N} \psi_i(u^{(i)}) \tag{18}$$

$$(\forall \mu \in \mathbb{R}) \quad \psi_i(\mu) = \begin{cases} \dfrac{(\mu - z^{(i)})^2}{2(\alpha_1 \mu + \alpha_0)} & \text{if } \mu \ge \delta \\ +\infty & \text{otherwise.} \end{cases} \tag{19}$$

5.2 Quadratic extension

The functions f and g in the resulting MAP criterion belong to $\Gamma_0(\mathbb{R}^K)$. If we now investigate the Lipschitz-continuity of the gradient of g, it turns out that g has a Lipschitz-continuous gradient on the domain of g but that the Lipschitz constant β is large when the components of z take large values and/or $\alpha_1\delta + \alpha_0$ is small. As a consequence, if Algorithm 4.1 is employed, only small values of the step-sizes $(\gamma_n)_{n\in\mathbb{N}}$ are allowed, inducing a slow convergence.

To circumvent this problem, we propose to use a lower approximation g_θ of the function g where $\theta > 0$ and $g_\theta = \Psi_\theta \circ T \circ F^*$ with

$$(\forall u = (u^{(i)})_{1 \le i \le N} \in \mathbb{R}^N) \quad \Psi_\theta(u) = \sum_{i=1}^{N} \psi_{\theta,i}(u^{(i)}). \tag{20}$$

The functions $(\psi_{\theta,i})_{1\le i\le N}$ are chosen such that, for every $\mu \in \mathbb{R}$,

$$\psi_{\theta,i}(\mu) = \begin{cases} \dfrac{\theta}{2}\mu^2 + \zeta_{i,1}(\theta)\,\mu + \zeta_{i,0}(\theta) & \text{if } \delta \le \mu < \mu_i(\theta) \\ \psi_i(\mu) & \text{otherwise,} \end{cases} \tag{21}$$

and

$$\alpha_1 \mu_i(\theta) = \left(\theta^{-1}\sigma^4(z^{(i)})\right)^{1/3} - \alpha_0. \tag{22}$$

This means that, when $\sigma^2(z^{(i)}) > (\alpha_0 + \delta)^3\theta$, a quadratic extension of the function ψ_i over $[\delta, \mu_i(\theta)]$ is performed, where the real constants $\zeta_{i,0}(\theta)$ and $\zeta_{i,1}(\theta)$ are adjusted so as to guarantee the continuity of $\psi_{\theta,i}$ and of its first and second order derivatives over $\mathbb{R}_+$. This extension is illustrated in Fig. 1.

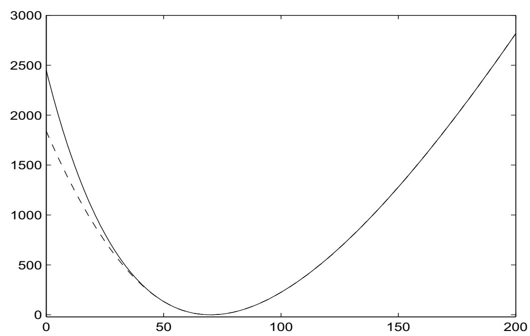


Figure 1: Graph of the function ψ_i (continuous line) and its quadratic extension $\psi_{\theta,i}$ (dashed line) when $\delta = 0$, $\alpha_0 = 1$, $\alpha_1 = 10^{-2}$, $\theta = 0.8$ and $z^{(i)} = 70$.

As a consequence, it can be shown that Ψ_θ is in $\Gamma_0(\mathbb{R}^N)$ and it has a Lipschitz-continuous gradient with constant θ. So, if $f + g_\theta + \iota_C$ is minimized with the proposed algorithm, the parameter θ allows us to control the Lipschitz constant of the gradient of g_θ (i.e. the convergence speed) and, at the same time, the closeness of the approximation to a minimizer of the original MAP criterion. In practice, the choice of this parameter results from a trade-off.

5.3 Simulation results

Here, T is a 7×7 uniform blur with $\|T\| = 1$. A 256×256 phantom image $\overline{y}$ is degraded by T and a signal-dependent additive noise following the model described in Section 5.1.

An orthonormal discrete wavelet representation (Symlets of length 6) has been adopted in this example. A generalized Gaussian prior distribution has been used to model the wavelet coefficients. The potential functions ϕ_k are thus taken of the form $\omega_k|.|^{p_k}$ where $\omega_k > 0$ and $p_k \in \{1, 4/3, 3/2, 2\}$ are subband adaptive. A constraint on the solution can be introduced by choosing $C = \{x \in \mathbb{R}^K | F^*x \in [0,255]^{256\times256}\}$

The relaxation parameter and the stepsize of Algorithm 4.1 are respectively chosen equal to $\lambda_n \equiv 1$ and $\gamma_n \equiv 1.99/\theta$. The error between images y and $\overline{y}$ is defined by $20\log_{10}(\|\overline{y}\|/\|y - \overline{y}\|)$. Results are provided in Fig. 2 where the two scenarii $\alpha_0 > \alpha_1$ and $\alpha_0 < \alpha_1$ have been considered.

It can be noticed that satisfactory results are obtained when the solution is contrained to belong to the convex C. In the case when $\alpha_0 > \alpha_1$ (left column in Fig. 2), this allows to observe a 1dB improvement with respect to the case when the constraint is not activated and visually speaking, dark areas are better restored. In the case when $\alpha_0 < \alpha_1$ (right column in Fig. 2), we observe a smaller quantitative improvement but light areas in the image are better restored. In both cases, the dynamic range of the image is clearly not respected if the convex constraint is not activated.

6. CONCLUSION

In this paper, we have proposed a new algorithm based on convex optimization to solve recovery problem subject to constraints. Our method is based on the insertion of a Douglas-Rachford step in a forward-backward algorithm. A major problem with this approach is the requirement of a differentiable term with a Lipschitz-continuous gradient in the objective function, which is not always satisfied, in particular in the considered example involving signal-dependent noise. Another contribution of this work was to propose a quadratic extension technique allowing to overcome this limitation.

Acknowledgements. We would like to thank Dr Claude Comtat from Service Hospitalier Frédéric Joliot (CEA, Orsay) for providing us the experimental data.

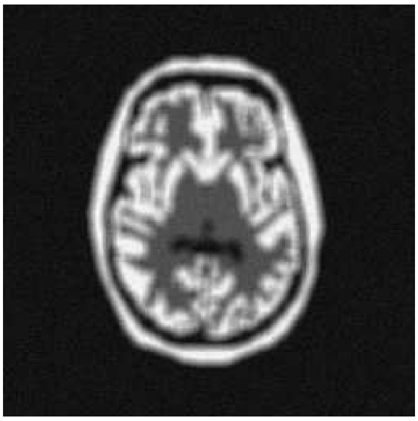
REFERENCES

[1] M. A. T. Figueiredo and R. D. Nowak. An EM algorithm for wavelet-based image restoration. *IEEE Trans. on Image Proc.*, 12(8):906–916, August 2003.

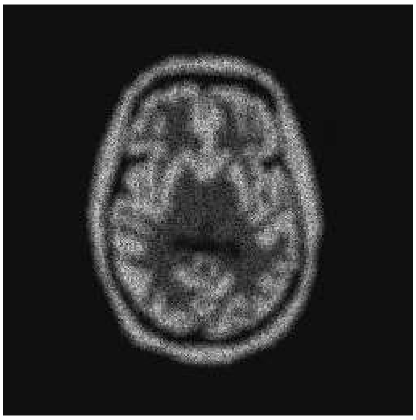
[2] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Applied Math.*, 57:1413–1457, 2004.

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

[4] C. Chaux, P. L. Combettes, J.-C. Pesquet, and V. R. Wajs. A variational formulation for frame-based inverse problems. *Inverse Problems*, 23:1495–1518, June 2007.



Degraded ($\alpha_0 = 25$, $\alpha_1 = 0.01$ and SNR $= 8.29$dB)



Degraded ($\alpha_0 = 4$, $\alpha_1 = 10$ and SNR $= 7.24$dB)

Restored without the convex constraint (SNR=10.9 dB and pixel range $[-123; 413]$)

Restored without the convex constraint (SNR=8.62 dB and pixel range $[-2.51; 448]$)

Restored with ι_C (SNR=12.1 dB)

Restored with ι_C (SNR=8.88 dB)

Figure 2: Restoration results on a 256×256 phantom image ($M_n \equiv 10$, $\theta = 1/\alpha_0$) after at most 10000 iterations.

[5] P. L. Combettes and J.-C. Pesquet. Proximal thresholding algorithm for minimization over orthonormal bases. *SIAM Journal on Optimization*, 18(4):1351–1376, 2007.

[6] J. M. Bioucas-Dias and M. A. T. Figueiredo. A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration. *IEEE Trans. on Image Proc.*, 16(12):2992–3004, 2007.

[7] C. Vonesch and M. Unser. A fast thresholded Landweber algorithm for wavelet-regularized multidimensional deconvolution. *IEEE Trans. on Image Proc.*, 2008. Preprint.

[8] P. L. Combettes and J.-C. Pesquet. A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery. *IEEE Journal of Selected Topics in Signal Processing*, 1(4):564–574, December 2007.

[9] F.-X. Dupé, M. J Fadili, and Starck J.-L. Image deconvolution under Poisson noise using sparse representations and proximal thresholding iteration. In *Proc. Int. Conf. on Acoust., Speech and Sig. Proc.*, 2008. Preprint.

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