Minimization of a sparsity promoting criterion for the recovery of complex-valued signals

Lotfi Chaâri[1,2], Jean-Christophe Pesquet[1], Amel Benazza-Benyahia[3] and Philippe Ciuciu[2]

[1] Université Paris-Est, IGM and UMR-CNRS 8049,
77454 Marne-la-Vallée cedex, France
{chaari,pesquet}@univ-mlv.fr

[2] CEA/DSV/I^{2}BM/Neurospin
CEA Saclay, Bbt. 145, Point Courrier 156, 91191 Gif-sur-Yvette cedex, France
philippe.ciuciu@cea.fr

[3] URISA, SUP'COM, Cité Technologique des Communications, 2083, Tunisia
benazza.amel@supcom.rnu.tn

***Abstract*—Ill-conditioned inverse problems are often encountered in signal/image processing. In this respect, convex objective functions including a sparsity promoting penalty term can be used. However, most of the existing optimization algorithms were developed for real-valued signals. In this paper, we are interested in complex-valued data. More precisely, we consider a class of penalty functions for which the associated regularized minimization problem can be solved numerically by a forward-backward algorithm. Functions within this class can be used to promote the sparsity of the solution. An application to parallel Magnetic Resonance Imaging (pMRI) reconstruction where complex-valued images are reconstructed is considered.**

Keywords: Regularization, penalization, sparsity, frames, convex optimization, wavelets, parallel MRI.

I. Introduction

To recover a vector $\overline{x} \in \mathbb{R}^K$ from the observation of a noisy signal $z \in \mathbb{R}^L$ through a linear operator T, the optimization of penalized convex criteria has been widely investigated in the regularization of ill-posed inverse problems. This arises in particular when $L < K$, which makes the inverse problem under-determined and the solution not uniquely defined. Generally, a regularization parameter within the penalized criterion balances the characteristics of the solution between its closeness to the observed data and its regularity. Under a Bayesian framework, the regularization parameter can be interpreted as a hyper-parameter of some prior distribution used to model the objective signal.
In the regularization setting, an ℓ^1 norm penalty term can be used in order to promote the sparsity of the solution. Most of the related studies deal with real-valued signals corrupted with Gaussian noise [1, 2, 3]. However, some applications like spectral analysis [4, 5, 6] and parallel MRI reconstruction [7], deal with complex-valued signals. The work in [8] also addresses the complex-valued case.

In the real case, efficient optimization algorithms have been recently developed in [3, 7, 9, 10] relying on the concept of proximity operators. In this paper, we show how the forward-backward algorithm can be applied to the complex case for a class of penalty functions which promote sparsity. The considered penalty functions allow us to introduce dependencies between the real and imaginary parts of complex-valued data.

The outline of this paper is as follows. In the next section we will detail the proposed optimization algorithm. Section III is devoted to the application of the proposed approach to parallel MRI reconstruction and the obtained results are also presented, before concluding in Section IV.

II. Optimization algorithm

Let F be the linear frame analysis operator defined as:

$$\begin{aligned} F\colon \mathbb{C}^L &\to \mathbb{C}^K \\ y &\mapsto (\langle y \mid \boldsymbol{e}_k \rangle)_{1\le k\le K}, \end{aligned} \tag{1}$$

where $\langle \cdot \mid \cdot \rangle$ is the standard Euclidean inner product. Let also S be a linear operator which maps the objective complex-valued signal $\overline{y} \in \mathbb{C}^L$ to the observed one $z \in \mathbb{C}^L$.
Of particular interest is the case when we want to recover $\overline{x} = (\overline{\xi}_k)_{1\le k\le K}$, the vector of coefficients of $\overline{y}$ in a frame $(\boldsymbol{e}_k)_{1\le k\le K}$.
We have then:

$$\overline{y} = \sum_{k=1}^{K} \overline{\xi}_k \boldsymbol{e}_k. \tag{2}$$

By setting $T = SF^*$ where F^* is the adjoint synthesis operator, the frame coefficients of $\overline{y}$ can be recovered by minimizing the following criterion:

$$\forall x \in \mathbb{C}^K,\ J(x) = f(Tx - z) + \alpha g(x). \tag{3}$$

In what follows, we will set $\alpha = 1$ and use the following penalty term:

$$g(x) = \sum_{k=1}^{K} \omega_k |\xi_k|^{p_k} \tag{4}$$

where $|\cdot|$ denotes the complex modulus, $(\omega_k)_{1\le k\le K} \in [0, +\infty[^K$ and $(p_k)_{1\le k\le K} \in [1, +\infty[^K$.

To find the optimal solution minimizing the (non necessarily differentiable) criterion in (3), we will employ a generalization of the forward-backward algorithm proposed in [3, 7, 9, 10]. The goal of the algorithm is to iteratively compute the solution, by making use of the concept of proximity operators which was found to be fundamental in a number of recent works in convex optimization. Before presenting the generalized version of the optimization algorithm, we recall the following definition.

Definition II.1 [11]
Let $\Gamma_0(\mathcal{H})$ be the class of proper lower semicontinuous convex functions from a separable Hilbert space $\mathcal{H}$ to $]-\infty,+\infty[$ and let $\varphi \in \Gamma_0(\mathcal{H})$.
For every $\mathsf{x} \in \mathcal{H}$, the function $\varphi + \|.-\mathsf{x}\|^2/2$ achieves its infimum at a unique point denoted by $\mathrm{prox}_\varphi \mathsf{x}$. The operator $\mathrm{prox}_\varphi\colon \mathcal{H} \to \mathcal{H}$ is the proximity operator of φ.

For the functions of a complex variable defined by

$$\begin{aligned} \phi_k\colon \mathbb{C} &\to \mathbb{R} \\ \xi &\mapsto \omega_k |\xi|^{p_k} \end{aligned} \tag{5}$$

with $k \in \{1,\dots,K\}$, it can be deduced from [12, Prop. 2.8], that the definition of the proximity operator can be extended as follows:

- if $p_k = 1$, then

$$\forall \xi \in \mathbb{C}, \qquad \mathrm{prox}_{\phi_k}(\xi) = \begin{cases} \Big(1 - \dfrac{\omega_k}{|\xi|}\Big)\xi & \text{if } |\xi| > \omega_k \\ 0 & \text{otherwise;} \end{cases}$$

- if $p_k > 1$, then

$$\forall \xi \in \mathbb{C}, \qquad \mathrm{prox}_{\phi_k}(\xi) = \begin{cases} \Big(1 - \dfrac{\nu_k(\xi)}{|\xi|}\Big)\xi & \text{if } \xi \neq 0 \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

where $\nu_k(\xi)$ is the unique number in $[0,+\infty[$ such that $\nu_k(\xi) + \big(\nu_k(\xi)/(\omega_k p_k)\big)^{1/(p_k-1)} = |\xi|$.

We therefore observe that, when $p_k = 1$, a bivariate proximal thresholder [13] is obtained, which can be employed to enforce sparsity. Fig. 2 illustrates an example of original complex-valued signal (left) and the thresholded one (right) using the considered proximity operator with $p_k = 1$ and $\omega_k = 8$.
It can be noted that complex-valued values ξ such that $|\xi| \leq \omega_k$ are set to zero, whereas the moduli of the others are attenuated.

Provided that f is a differentiable convex function with β-Lipschitz gradient ∇f, the employed optimization algorithm can then be summarized as follows.

Initialize with some $x_0 \in \mathbb{C}^K$, fix the relaxation parameter $\lambda \in]0,1]$ and the step-size parameter $\gamma \in]0, \frac{2}{\beta}[$. Set $n = 1$.
For $n \in \mathbb{N}^*$ **do**
  For $k \in \{1,\dots,K\}$ **do**
    $\xi_{n+1,k} = \xi_{n,k} + \lambda\big(\mathrm{prox}_{\gamma\phi_k}(\xi_{n,k} - \gamma(T^*\nabla f(Tx_n - z))_k) - \xi_{n,k}\big)$
  end For
end For
return $x_n = (\xi_{n,k})_{1\leq k\leq K}$

Algorithm 1: Forward-backward

III. Application to parallel MRI (pMRI) reconstruction

A. pMRI basics

Parallel Magnetic Resonance Imaging [14] is a fast acquisition technique to reduce global imaging time in MRI, which is particularly useful in functional MRI (fMRI) to limit the distortion artifacts and signal losses. To this end, a number L of receiver coils with complementary spatial sensitivities $(s_\ell)_{1\leq \ell \leq L}$ are employed to acquire L MRI signals at the same time. The received signal by a given coil ℓ corresponds to the Fourier transform of the desired 2D field $\overline{y}$ weighted by the corresponding coil sensitivity profile s_ℓ. For the sake of simplicity, a regular Cartesian sampling is used during the acquisition process. However, in parallel MRI, sub-sampling of the frequency (i.e. k-space) domain is performed along the phase encoding direction according to a fixed reduction factor R, which involves a time acquisition R times shorter than with conventional MRI. Hence, only reduced size images are acquired. Fig. 1 shows the difference between conventional and parallel MRI acquisitions, where a reduction factor $R = 2$ is applied.
However, the higher the reduction factor is, the more aliased the images are. Indeed, because of the sampling under the Nyquist rate, registred data suffer from aliasing artifacts in the image domain, which increase with the reduction factor. The challenge here is to unfold the received images by exploiting the complementarity between the sensitivity profiles of the used coils, and to reconstruct a non-aliased full Field of View (FoV) image.
In what follows, we will describe the basic-SENSE method as proposed by Pruessmann in 1999, as well as the WT regularized reconstruction using the proposed algorithm. Comparison between the performance of these two methods will be made on a real data set of T1-weighted anatomical images acquired at 1.5 Tesla magnetic field with $R = 4$. This value of the reduction factor is considered as high for such a low magnetic field.

B. The SENSE method

In clinical routines, the most customary technique is the so-called Sensitivity Encoding (SENSE) [14] method. This method operates in the spatial domain, and the acquisition process is modelled by:

$$z = S\overline{y} + n, \tag{7}$$

where z is the complex-valued received signal, S is the sensitivity linear operator, $\overline{y}$ is the objective image and n is a zero-mean additive Gaussian noise with between-coil correlation matrix $\mathbf{\Psi}$. The SENSE approach corresponds to a weighted least squares estimation procedure which computes the following estimator:

$$\widehat{y}_{\mathrm{WLS}} = [S^{\mathsf{H}}\mathbf{\Psi}^{-1}S]^{-1}S^{\mathsf{H}}\mathbf{\Psi}^{-1}z \tag{8}$$

where $(\cdot)^{\mathsf{H}}$ stands for the transposed complex conjugate.
A reconstructed image using this method is displayed in Fig. 3. We notice that this image suffers from aliasing artifacts with curves having very high or very low intensities. Based on the nature of these artifacts, we propose to regularize this reconstruction problem in the Wavelet Transform (WT) domain to have better spatial and frequency localizations of these artifacts.

C. The regularized approach

The Maximum A posteriori (MAP) criterion under a Bayesian framework has been adopted to estimate $\widehat{x} = (\widehat{\xi_k})_{1\leq k\leq K}$, the vector of frame coefficients of $\overline{y}$, based on the observation vector z. F^* is then applied to recover the estimate $\widehat{y}$ in the image domain.
The following 2D sparsity-promoting prior, for which the potential function is given by (5), is used to model the frame coefficients:

$$\forall \xi \in \mathbb{C}, \quad h_{\omega_k,p_k}(\xi) = C_h e^{-\omega_k|\xi|^{p_k}}, \tag{9}$$

where C_h is a normalization constant, $\omega_k \in \mathbb{R}_+^*$ and $p_k \in [1, +\infty[$.
Since a zero-mean circular Gaussian noise corrupts the acquisition process, the likelihood of the data can be expressed as:

$$\mathcal{L}(z \mid y) \propto \exp\left(- \parallel z - Sy \parallel^2_{\mathbf{\Psi}^{-1}}\right). \tag{10}$$

where $\|\cdot\|_{\mathbf{\Psi}^{-1}} = \sqrt{(\cdot)^{\mathsf{H}}\mathbf{\Psi}^{-1}(\cdot)}$ is a norm on $\mathbb{C}^L$.
Combining Equations (9) and (10) leads to the following MAP criterion to be minimized:

$$J(x) = \parallel z - Tx \parallel^2_{\mathbf{\Psi}^{-1}} + \sum_{k=1}^{K} \omega_k|\xi_k|^{p_k}, \tag{11}$$

where $f = \parallel z - T\cdot \parallel^2_{\mathbf{\Psi}^{-1}}$ is a differentiable convex function with β-Lipschitz gradient.

Using a decomposition onto a Symmlet wavelet basis of length 8 over 3 resolution levels and the proposed algorithm to minimize the criterion J, the obtained regularized image is given in Fig. 4. It can be observed that aliasing artifacts are considerably smoothed when compared with the SENSE reconstruction in Fig. 3, but not completely removed: some of them still exist because they were extremely large. By evaluating the Signal to Noise Ratio ($\mathrm{SNR} = 20\log_{10}\left(\frac{\|y_{\mathrm{ref}}\|_2}{\|y_{\mathrm{ref}}-\widehat{y}\|_2}\right)$) for the reconstructed images, where y_{ref} is a reference image acquired with conventional MRI acquisition and $\widehat{y}$ is the reconstructed image, it turns out that the proposed algorithm gives an SNR improvement of 0.76 dB with respect to the basic SENSE reconstruction. Note that such a value is considered as significant in pMRI reconstruction. To illustrate the convergence of our algorithm, Fig. 5 gives the evolution of J versus the iteration number. It is clear that after about 30 iterations, convergence has been reached, and the estimate $\widehat{y}$ corresponds to a good estimation of the optimal MAP solution.

IV. Conclusion

We have proposed an algorithm to minimize sparsity promoting criteria for the regularization of inverse problems involving complex-valued signals. Application to parallel MRI reconstruction shows that this algorithm reduces aliasing artifacts in the reconstructed image compared with standard reconstruction techniques. The considered method can be applied to other classes of inverse problems as well.

References

[1] J.-J. Fuchs, "On sparse representations in arbitrary redundant bases," *IEEE Trans. Inform. Theory*, vol. 50, no. 6, pp. 1341–1344, Jun. 2004.

[2] J.-J. Fuchs, "Recovery of exact sparse representations in the presence of bounded noise," *IEEE Trans. Inform. Theory*, vol. 51, no. 10, pp. 3601–3608, Oct. 2005.

[3] C. Chaux, P. Combettes, J.-C. Pesquet and V.R Wajs, "A variational formulation for frame-based inverse problems," *Inverse Problems*, vol. 23, pp. 1495–1518, Aug. 2007.

[4] Ph. Ciuciu, J. Idier and J.-F. Giovannelli, "Regularized estimation of mixed spectra using a circular Gibbs-Markov model," *IEEE Trans. Sig. Proc.*, vol. 49, no. 10, pp. 2202–2213, Oct. 2001.

[5] Ph. Ciuciu and J. Idier, "A Half-Quadratic block-coordinate descent method for spectral estimation," *IEEE Trans. Sig. Proc.*, vol. 82, no. 7, pp. 941–959, Jul. 2002.

[6] S. Bourguignon, H. Carfantan and J. Idier, "A bparsity-based method for the estimation of spectral lines from irregularly sampled data," *IEEE Sel. Top. Sig. Proc.*, vol. 1, pp. 575–585, Dec. 2007.

[7] L. Chaâri, J.-C. Pesquet, A. Benazza-Benyahia and Ph. Ciuciu, "Autocalibrated parallel MRI reconstruction in the wavelet domain," *IEEE International Symposium on Biomedical Imaging*, pp. 756–759, Paris, France, 14-17 May, 2008.

[8] J.J. Fuchs, "Convergence of a sparse representations algorithm applicable to real or complex data," *IEEE J. Select. Top. Sig. Proc.*, vol. 1, pp. 598–605, 2007.

[9] 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.

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

[11] J.-J. Moreau, "Proximité et dualité dans un espace hilbertien," *Bull. de la Soc. math. de France*, vol. 93, pp. 273–299, 1965.

[12] P. L. Combettes and J.-C. Pesquet, "A proximal decomposition method for solving convex variational inverse problems", *Inverse Problems*, vol. 24, 27 p., Dec. 2008.

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

[14] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: sensitivity encoding for fast MRI," *Magnetic Resonance in Medicine*, vol. 42, no. 5, pp. 952–962, July 1999.

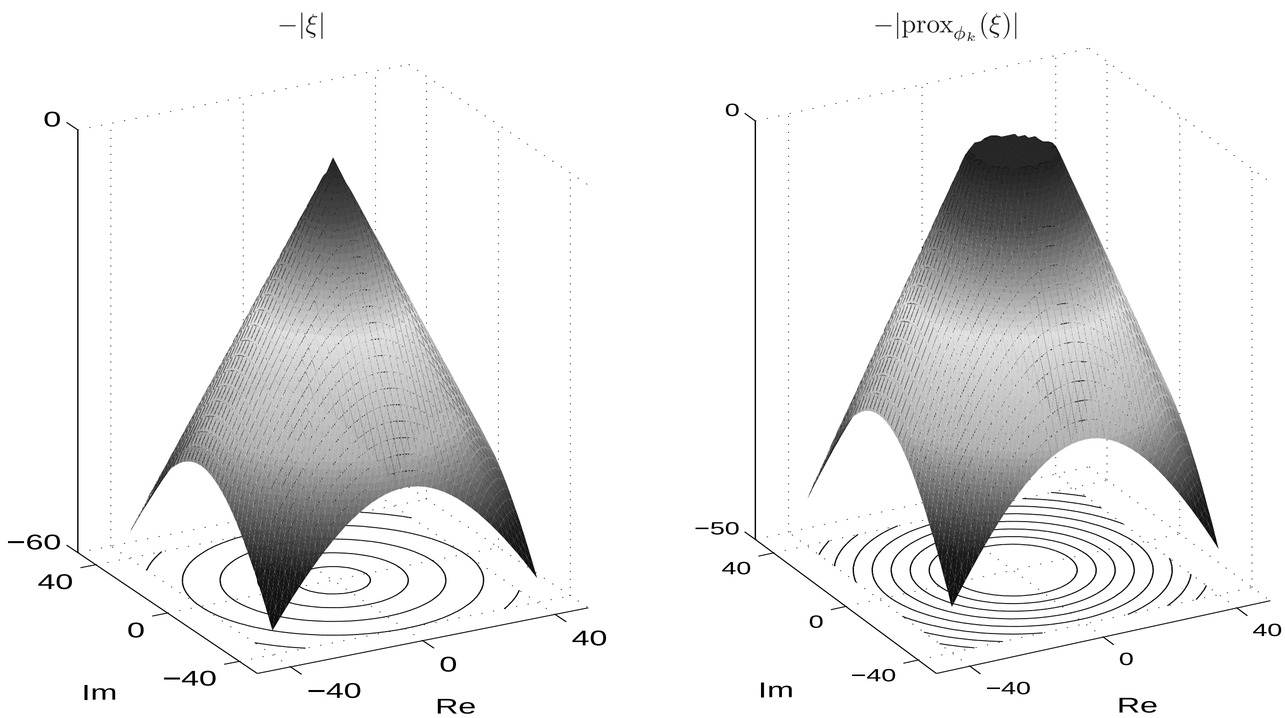

Fig. 2. Original signal (left) and thresholded one (right) with $p_k = 1$ and $\omega_k = 8$.

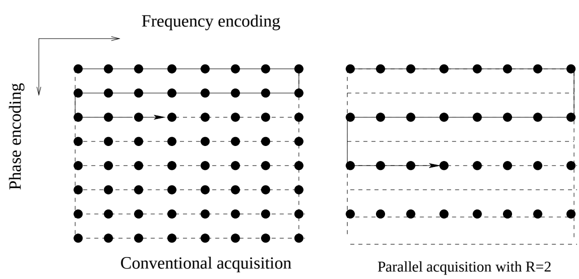

Fig. 1. Sampling of the k-space.

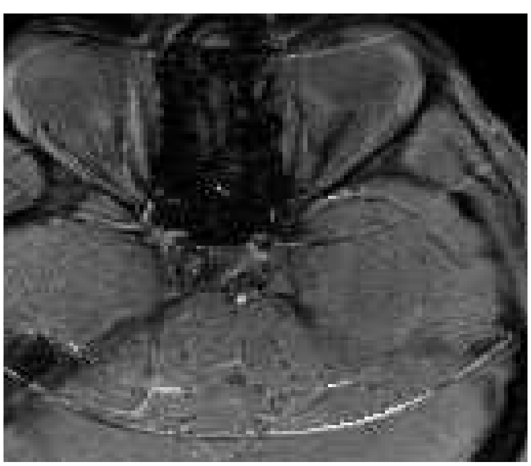

Fig. 3. Reconstructed image using SENSE ($\mathrm{SNR} = 13.74$ dB).

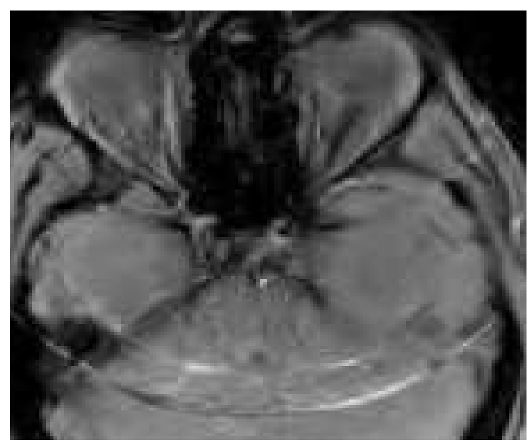

Fig. 4. Reconstructed image using the proposed algorithm ($\mathrm{SNR} = 14.50$ dB).

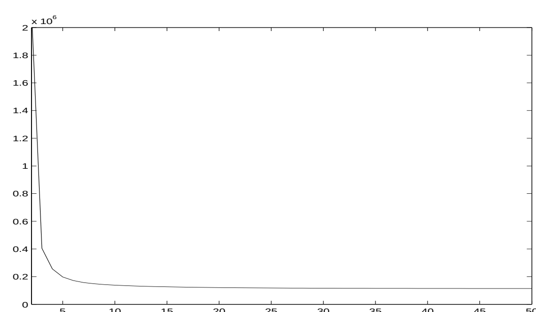

Fig. 5. Convergence curve of the proposed algorithm w.r.t the iteration number.