

A COMPLEX-VALUED MAJORIZE-MINIMIZE MEMORY GRADIENT METHOD WITH APPLICATION TO PARALLEL MRI

Anisia Florescu¹, Emilie Chouzenoux², Jean-Christophe Pesquet², Philippe Ciuciu³ and Silviu Ciochina¹

¹ Politehnica University of Bucharest, Telecommunications Dept., Romania

² Université Paris-Est, LIGM, UMR CNRS 8049, Champs sur Marne, France

³ CEA, NeuroSpin center, INRIA Saclay, PARIETAL Team, Gif-sur-Yvette, France

ABSTRACT

Complex-valued data are encountered in many application areas of signal and image processing. In the context of optimization of functions of real variables, subspace algorithms have recently attracted much interest, due to their efficiency in solving large-size problems while simultaneously offering theoretical convergence guarantees. The goal of this paper is to show how some of these methods can be successfully extended to the complex case. More precisely, we investigate the properties of the proposed complex-valued Majorize-Minimize Memory Gradient (3MG) algorithm. An important practical application of these results arises for image reconstruction in Parallel Magnetic Resonance Imaging (PMRI). Comparisons with existing optimization methods confirm the good performance of our approach for PMRI reconstruction.

Index Terms— complex-valued signals, optimization, subspace algorithms, descent methods, majorization-minimization, image reconstruction, inverse problems, magnetic resonance imaging, sampling

1. INTRODUCTION

Complex-valued data are ubiquitous in signal and image processing. As emphasized in [1], dealing with complex-valued signals raises a number of challenging theoretical issues, in particular owing to their existing relations with the theory of analytic functions. Problems involving complex-valued signals are often formulated as the search for a solution satisfying some optimality conditions. Since the related optimization problems usually do not have closed form solutions, efforts have been dedicated to the development of specific iterative algorithms for minimizing real-valued functions of complex variables [2, 3]. However, one shortcoming of existing approaches (e.g. interior point methods) is that they may not be very efficient to deal with large-size problems. Another weakness lies in the lack of theoretical convergence guarantees. For example, popular methods such as the nonlinear conjugate gradient algorithm, which may be quite effective

in practice, have only been proved to converge under restrictive assumptions. Proximal splitting [4] and augmented Lagrangian [5] methods offer more flexibility for minimizing possibly nonsmooth objective functions, but they may suffer from slow convergence. In the case of functions of real variables, a recent majorize-minimize (MM) subspace algorithm has been proposed which overcomes these limitations [6, 7]. Note that MM strategies for functions of complex variables were already investigated in [8, 9] but they were restricted to half-quadratic algorithms requiring the inversion of a large-size linear operator, which is not tractable for any acquisition model.

Complex-valued data are involved in several imaging systems such as in Magnetic Resonance Imaging (MRI). Many recent works were directed to the proposal of reconstruction methods for parallel MRI (PMRI) [10, 11, 12, 13, 14]. The objective of PMRI is to reduce the acquisition time while maintaining a good image quality. This is achieved by combining subsampling strategies in the k -space with the use of an array of coils so as to compensate spectral decimation with spatial diversity. Let us emphasize that the design of an appropriate subsampling scheme is strongly related to compressive sensing issues [15, 16]. Among the existing reconstruction approaches based on variational formulations and optimization algorithms, we can mention methods based on iterative soft-thresholding [17] or more elaborate proximal algorithms [11], and augmented Lagrangian techniques [13, 18]. Note that, to the best of our knowledge, MM subspace algorithms have never been used in the context of PMRI.

The organization of the paper is as follows: in Section 2, the addressed optimization problem is formulated in a general manner and our notation is introduced. In Section 3, we recall some classical results about the derivative of real-valued functions of complex variables, which are relevant to this work. Section 4 describes the employed MM strategy in the complex case. The proposed complex-valued 3MG (Majorize-Minimize Memory Gradient) algorithm is studied. In particular, its connections with the algorithm in [6, 7] for minimizing functions of real variables are discussed. Section 5 presents the application of our algorithm to PMRI reconstruction.

Part of this work was supported by PhD Fellowship “Investitii in cercetare-inovare-dezvoltare pentru viitor (DocInvest)”, EC project POS-DRU/107/1.5/S/76813.

2. PROBLEM STATEMENT

In this work, we will consider the following penalized optimization problem:

$$\underset{\mathbf{x} \in \mathbb{C}^N}{\text{minimize}} \quad (F(\mathbf{x}) = \Phi(\mathbf{H}\mathbf{x} - \mathbf{y}) + \Psi(\mathbf{x})), \quad (1)$$

where $\mathbf{H} \neq \mathbf{0}$ is a matrix in $\mathbb{C}^{Q \times N}$, \mathbf{y} is an observation vector in \mathbb{C}^Q , $\Phi: \mathbb{C}^Q \rightarrow \mathbb{R}$, and $\Psi: \mathbb{C}^N \rightarrow \mathbb{R}$. In inverse problems, function Φ usually corresponds to a data-fidelity term and Ψ to a regularization function. We focus on the case when the latter function takes the following form:

$$(\forall \mathbf{x} \in \mathbb{C}^N) \quad \Psi(\mathbf{x}) = \sum_{s=1}^S \psi_s(|\mathbf{v}_s^H \mathbf{x}|) \quad (2)$$

where $|\cdot|$ denotes the complex modulus, $(\cdot)^H$ is the matrix trans-conjugate operation, and, for every $s \in \{1, \dots, S\}$, $\psi_s: \mathbb{R} \rightarrow \mathbb{R}$, and $\mathbf{v}_s \in \mathbb{C}^N$. Note that (2) includes anisotropic total variation and frame analysis penalties as special cases.

Notation: For every vector $\mathbf{x} \in \mathbb{C}^N$, $\mathbf{x}_R \in \mathbb{R}^N$ (resp. $\mathbf{x}_I \in \mathbb{R}^N$) denotes the vector of real (resp. imaginary) parts of the components of \mathbf{x} . Let $\tilde{\mathbf{x}} \in \mathbb{R}^{2N}$ be the ‘‘concatenated’’ vector $\tilde{\mathbf{x}} = [\mathbf{x}_R^\top \ \mathbf{x}_I^\top]^\top$. We define $\tilde{\Psi}$ the function of real variables associated with Ψ , i.e. $(\forall \mathbf{x} \in \mathbb{C}^N) \tilde{\Psi}(\tilde{\mathbf{x}}) = \Psi(\mathbf{x})$. A similar notation will be employed for complex-valued matrices and other functions of complex variables.

3. COMPLEX-VALUED DIFFERENTIAL CALCULUS

Let Θ be a function from \mathbb{C}^N to \mathbb{C} . According to Wirtinger’s calculus [1], the derivative of Θ with respect to the conjugate of its variable is formally defined as

$$(\forall \mathbf{x} \in \mathbb{C}^N) \quad \nabla \Theta(\mathbf{x}) = \frac{1}{2} \left(\frac{\partial \tilde{\Theta}(\tilde{\mathbf{x}})}{\partial \mathbf{x}_R} + \iota \frac{\partial \tilde{\Theta}(\tilde{\mathbf{x}})}{\partial \mathbf{x}_I} \right). \quad (3)$$

Throughout this paper, we suppose that:

Assumption 1.

- (i) $\tilde{\Phi}$ is differentiable.
- (ii) For every $s \in \{1, \dots, S\}$, ψ_s is a differentiable function and $\lim_{t \rightarrow 0, t \neq 0} \dot{\psi}_s(t)/t \in \mathbb{R}$ where $\dot{\psi}_s$ denotes the derivative of ψ_s .

The definition in (3) implies that the derivative of $\Phi(\mathbf{H} \cdot - \mathbf{y})$ at $\mathbf{x} \in \mathbb{C}^N$ is $\mathbf{H}^H \nabla \Phi(\mathbf{H}\mathbf{x} - \mathbf{y})$. Let us now define

$$(\forall a \in \mathbb{R}) \quad \omega_s(a) = \begin{cases} \dot{\psi}_s(a) & \text{if } a \neq 0 \\ \lim_{t \rightarrow 0, t \neq 0} \dot{\psi}_s(t)/t & \text{otherwise.} \end{cases} \quad (4)$$

It can be easily shown that the Wirtinger derivative of Ψ is

$$(\forall \mathbf{x} \in \mathbb{C}^N) \quad \nabla \Psi(\mathbf{x}) = \frac{1}{2} \mathbf{V} \text{Diag}(\mathbf{b}(\mathbf{x})) \mathbf{V}^H \mathbf{x} \quad (5)$$

where $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_S] \in \mathbb{C}^{N \times S}$ and $\mathbf{b}(\mathbf{x}) = (\omega_s(|\mathbf{v}_s^H \mathbf{x}|))_{1 \leq s \leq S}$. So, the complex-valued derivative of F is

$$(\forall \mathbf{x} \in \mathbb{C}^N) \quad \nabla F(\mathbf{x}) = \mathbf{H}^H \nabla \Phi(\mathbf{H}\mathbf{x} - \mathbf{y}) + \nabla \Psi(\mathbf{x}) \quad (6)$$

where the derivative of Ψ is given by (5).

4. QUADRATIC MAJORIZATION

In order to develop an efficient algorithm for solving Problem (1), we introduce the following additional assumption:

Assumption 2.

- (i) Φ has a β -Lipschitz derivative with $\beta \in (0, +\infty)$, i.e.

$$(\forall \mathbf{z} \in \mathbb{C}^Q) (\forall \mathbf{z}' \in \mathbb{C}^Q) \quad \|\nabla \Phi(\mathbf{z}) - \nabla \Phi(\mathbf{z}')\| \leq \beta \|\mathbf{z} - \mathbf{z}'\|. \quad (7)$$

- (ii) For every $s \in \{1, \dots, S\}$, $\psi_s(\sqrt{\cdot})$ is concave on $[0, +\infty)$.

- (iii) For every $s \in \{1, \dots, S\}$, there exists $\bar{\omega}_s \in [0, +\infty)$ such that $(\forall t \in (0, +\infty)) 0 \leq \omega_s(t) \leq \bar{\omega}_s$.

Note that Assumption 2(i) is quite standard and, in particular, it is satisfied for least squares data fidelity terms. Assumptions 2(ii) and 2(iii) hold for a wide class of penalty functions, e.g. ℓ_2 - ℓ_1 convex functions constituting smooth approximations of the ℓ_1 norm [6], or ℓ_2 - ℓ_0 nonconvex functions providing smooth approximations of the ℓ_0 cost [7].

The following property can then be proved:

Proposition 1. Under Assumptions 1 and 2, for every $(\mathbf{x}, \mathbf{x}') \in (\mathbb{C}^N)^2$, $F(\mathbf{x}) \leq \Theta(\mathbf{x}, \mathbf{x}')$, where

$$\Theta(\mathbf{x}, \mathbf{x}') = F(\mathbf{x}') + 2\text{Re} \{ \nabla F(\mathbf{x}')^H (\mathbf{x} - \mathbf{x}') \} + \frac{1}{2} (\mathbf{x} - \mathbf{x}')^H \mathbf{A}(\mathbf{x}') (\mathbf{x} - \mathbf{x}')$$

and $\mathbf{A}(\mathbf{x}') = \mu \mathbf{H}^H \mathbf{H} + \mathbf{V} \text{Diag}(\mathbf{b}(\mathbf{x}')) \mathbf{V}^H$ with $\mu \in [2\beta, +\infty)$.

Subspace algorithms consist of building a sequence $(\mathbf{x}_k)_{k \in \mathbb{N}}$ according to the following iterative scheme:

$$(\forall k \in \mathbb{N}) \quad \mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{D}_k \mathbf{u}_k, \quad (8)$$

where $\mathbf{D}_k \in \mathbb{C}^{N \times M}$ is a subspace search matrix and $\mathbf{u}_k \in \mathbb{C}^M$ is a multivariate step-size minimizing $\mathbf{u} \mapsto F(\mathbf{x}_k + \mathbf{D}_k \mathbf{u})$ over \mathbb{C}^M . The MM strategy replaces the minimization of the original function F over the subspace with successive minimizations of quadratic tangent majorants $\mathbf{u} \mapsto$

$\Theta(\mathbf{x}' + \mathbf{D}_k \mathbf{u}, \mathbf{x}')$ over \mathbb{C}^M for some vector $\mathbf{x}' \in \mathbb{C}^N$. The expression of Θ in Proposition 1 shows that, for a given \mathbf{x}' , an optimal solution is

$$\widehat{\mathbf{u}} = -2(\mathbf{D}_k^H \mathbf{A}(\mathbf{x}') \mathbf{D}_k)^\dagger \mathbf{D}_k^H \nabla F(\mathbf{x}'), \quad (9)$$

where $(\cdot)^\dagger$ denotes the pseudo-inverse operation. The resulting complex-valued MM subspace algorithm for solving Problem (1) reads

$$\left\{ \begin{array}{l} \mathbf{x}_0 \in \mathbb{C}^N, \\ \text{For all } k = 0, \dots \\ \quad \left\{ \begin{array}{l} \mathbf{u}_k^0 = \mathbf{0}, \\ \text{For all } j = 1, \dots, J \\ \quad \left\{ \begin{array}{l} \mathbf{B}_k^{j-1} = \mathbf{D}_k^H \mathbf{A}(\mathbf{x}_k + \mathbf{D}_k \mathbf{u}_k^{j-1}) \mathbf{D}_k, \\ \mathbf{u}_k^j = \mathbf{u}_k^{j-1} - 2(\mathbf{B}_k^{j-1})^\dagger \mathbf{D}_k^H \nabla F(\mathbf{x}_k + \mathbf{D}_k \mathbf{u}_k^{j-1}), \\ \mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{D}_k \mathbf{u}_k^J. \end{array} \right. \end{array} \right. \end{array} \right. \quad (10)$$

When the number M of search directions is small, the computation cost of the multivariate step-size $\widehat{\mathbf{u}}$ in (9) is reduced, thus making the complexity of an iteration of the proposed algorithm quite reasonable. When $M = 2$, a typical choice for the search directions is $\mathbf{D}_k = [-\nabla F(\mathbf{x}_k), \mathbf{x}_k - \mathbf{x}_{k-1}]$ for every $k \in \mathbb{N}$ (by setting $\mathbf{x}_{-1} = \mathbf{0}$), which leads to the so-called MM Memory Gradient algorithm.

Algorithm (10) takes a form similar to the one developed in the real case in [6]. Note however that, since, for every $k \in \mathbb{N}$ and $j \in \{1, \dots, J\}$, the multivariate step-size \mathbf{u}_k^j is complex-valued, Algorithm (10) can be viewed as a way of expressing in a concise manner [7, Algorithm (3.16)] for minimizing F over \mathbb{R}^{2N} , when the subspace search matrix at iteration k reads $\widetilde{\mathbf{D}}_k = \begin{bmatrix} \mathbf{D}_{k,R} & -\mathbf{D}_{k,I} \\ \mathbf{D}_{k,I} & \mathbf{D}_{k,R} \end{bmatrix}$. Therefore, it corresponds to $2M$ search directions in \mathbb{R}^{2N} .

Due to this relation between the complex-valued 3MG algorithm and its real-valued counterpart, the following result can be deduced from [7]:

Proposition 2. *Assume that F is a semi-algebraic function such that $\lim_{\|\mathbf{x}\| \rightarrow +\infty} F(\mathbf{x}) = +\infty$. Under Assumptions 1 and 2, Algorithm (10) generates a sequence $(\mathbf{x}_k)_{k \in \mathbb{N}}$ converging to a critical point of F . Moreover, this sequence has a finite length in the sense that $\sum_{k=0}^{+\infty} \|\mathbf{x}_{k+1} - \mathbf{x}_k\| < +\infty$.*

5. APPLICATION TO PARALLEL MRI

5.1. Model

In parallel MRI, a set of measures $(\mathbf{d}_\ell)_{1 \leq \ell \leq L}$ is acquired from L coils. These measures are related to the original full FOV (Field Of View) image $\bar{\rho} \in \mathbb{C}^K$ (the image being columnwise reshaped as a vector) corresponding to a spin density. More precisely, the observation model reads:

$$(\forall \ell \in \{1, \dots, L\}) \quad \mathbf{d}_\ell = \Sigma \mathbf{F} \mathbf{S}_\ell \bar{\rho} + \mathbf{w}_\ell \quad (11)$$

where $\mathbf{S}_\ell \in \mathbb{C}^{K \times K}$ is a diagonal matrix modelling the sensitivity of the coils, $\mathbf{F} \in \mathbb{C}^{K \times K}$ is a 2D discrete Fourier transform, and $\Sigma \in \{0, 1\}^{\lfloor \frac{K}{R} \rfloor \times K}$ is a subsampling matrix (here, $\lfloor \cdot \rfloor$ designates the rounding operation). The $\lfloor K/R \rfloor$ lines of matrix Σ are thus distinct lines of a $K \times K$ identity matrix, R being the subsampling or acceleration factor. The noise vectors $(\mathbf{w}_\ell)_{1 \leq \ell \leq L}$ are realizations of random vectors $(\mathbf{W}_\ell)_{1 \leq \ell \leq L}$, which can be assumed mutually statistically independent. In addition, for every $\ell \in \{1, \dots, L\}$, \mathbf{W}_ℓ is a circular complex Gaussian vector with zero-mean and covariance matrix Λ_ℓ .

In order to provide an estimate of $\bar{\rho}$, we propose to solve the following optimization problem:

$$\underset{\rho \in \mathbb{E}}{\text{minimize}} \quad \sum_{\ell=1}^L \|\Sigma \mathbf{F} \mathbf{S}_\ell \rho - \mathbf{d}_\ell\|_{\Lambda_\ell^{-1}}^2 + \sum_{s=1}^S \psi_s(|\mathbf{f}_s^H \rho|) \quad (12)$$

where $(\forall \ell \in \{1, \dots, L\}) \|\cdot\|_{\Lambda_\ell^{-1}} = (\cdot)^H \Lambda_\ell^{-1} (\cdot)$, $(\forall s \in \{1, \dots, S\}) \psi_s: \mathbb{R} \rightarrow \mathbb{R}$ and $\mathbf{f}_s \in \mathbb{C}^K$, and \mathbb{E} is a vector subspace corresponding to the range of a matrix $\mathbf{E} \in \mathbb{C}^{K \times N}$ with $N \leq K$. By choosing for $(\mathbf{f}_s)_{1 \leq s \leq S}$ a frame of \mathbb{C}^K (possibly redundant when $S > K$) the above function introduces a so-called frame analysis penalization. The vector space \mathbb{E} serves to incorporate some prior knowledge about the target image. In our case, \mathbf{E} is an interpolation matrix (i.e. the transpose of a subsampling matrix as defined above) the zero lines of which are associated with pixels belonging to the background of the image. Such an area can be identified from the sensitivity matrices.

By setting $\rho = \mathbf{E} \mathbf{x}$ (with $\mathbf{x} \in \mathbb{C}^N$) in Problem (12), this one appears as an instance of Problem (1) where

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_L \end{bmatrix} = \begin{bmatrix} \Lambda_1^{-1/2} \Sigma \mathbf{F} \mathbf{S}_1 \\ \vdots \\ \Lambda_L^{-1/2} \Sigma \mathbf{F} \mathbf{S}_L \end{bmatrix}, \quad \mathbf{E}, \quad \mathbf{y} = \begin{bmatrix} \Lambda_1^{-1/2} \mathbf{d}_1 \\ \vdots \\ \Lambda_L^{-1/2} \mathbf{d}_L \end{bmatrix}, \quad (13)$$

$(\mathbf{v}_s)_{1 \leq s \leq S} = (\mathbf{E}^H \mathbf{f}_s)_{1 \leq s \leq S}$, and Φ is the squared Hermitian norm of \mathbb{C}^Q with $Q = L \lfloor K/R \rfloor$.

5.2. Simulation results

In our experiments, the reconstruction of a sagittal view of a 3D anatomical image is performed from noisy parallel MRI data generated according to Model (11). The reference image $\bar{\rho}$ (see Fig. 1 left) is defined as the reconstruction result from a non-accelerated acquisition ($R = 1$) obtained with a 3 Tesla Siemens Trio magnet having an $L = 32$ -channel receiver coil (no parallel transmission has been used). The data have been acquired using a 3D T1-weighted MP-RAGE pulse sequence. A resampling has been performed in the k -space by zero-filling in order to facilitate the use of fast wavelet decompositions, so leading to a 256×256 image size ($K = 256^2$).

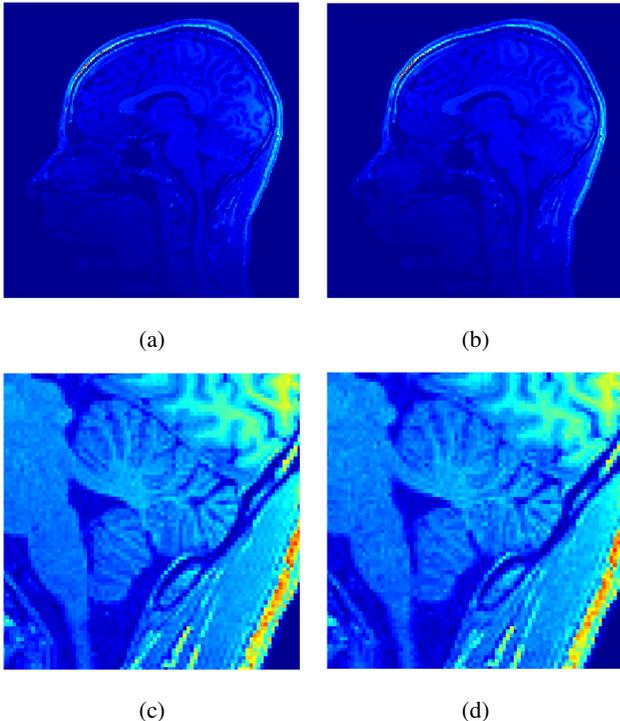


Fig. 1. Moduli of original (a) and reconstructed (b) images, using Poly1 sampling, 3MG algorithm and $\ell_2\text{-}\ell_1$ regularization, SNR = 19.95 dB. Figs (c) and (d) display corresponding zooms centered on the cerebellum area.

Estimates of the sensitivity matrices $(S_\ell)_{1 \leq \ell \leq L}$ are also available. Different sampling patterns with $R = 5$, are considered for Σ , namely regular line subsampling, uniform random, radial, spiral, with π density [19], and polynomial decay of various orders [15]. Finally, a circular complex Gaussian white noise with variance equal to 6×10^9 is added to the data.

Problem (12) is solved by using Algorithm (10) with parameter $J = 1$. The convex $\ell_2\text{-}\ell_1$ penalization function $\psi_s : t \rightarrow \lambda_s(\sqrt{1+t^2/\delta^2} - 1)$ is employed, for every $s \in \{1, \dots, S\}$. In the presented results, $(f_s)_{1 \leq s \leq S}$ ($S = K$) corresponds to an orthonormal wavelet basis using Symmlet filters of length 10, and the decomposition is carried out over 3 resolution levels. For simplicity, the parameters $(\lambda_s)_{1 \leq s \leq S}$ are equal to the same constant λ for the detail coefficients, while they have been set to zero for the approximation ones. The parameters λ and δ are tuned so as to maximize the Signal-to-Noise Ratio (SNR) between the reference image and its reconstructed version.

Table 1 allows us to evaluate the reconstruction performance of 3MG algorithm in terms of SNR for the different sampling patterns. One can observe that sampling strategies based on low-order polynomial distributions as well as on the uniform or π distributions lead to higher quality reconstructed images. The modulus of the reconstructed image for Poly1 sampling strategy is displayed in Fig. 1 (right).

We compare the proposed algorithm with state-of-the-

Sampling pattern	SNR (dB)
Poly1	19.95
Poly2	19.34
Poly3	18.53
Poly4	17.50
Poly5	16.95
Uniform	19.71
Radial	19.20
Spiral	19.17
Regular	18.13
π	19.31

Table 1. SNR values for various subsampling strategies using 3MG algorithm with $\ell_2\text{-}\ell_1$ regularization.

Algorithm	Penalization	SNR (dB)
M+LFBF [20]	ℓ_1	19.95
CPCV [21, 22]	ℓ_1	19.95
ADMM	ℓ_1	19.95
3MG	$\ell_2\text{-}\ell_1$	19.95
3MG	$\ell_2\text{-}\ell_0$ (G)	20.27
3MG	$\ell_2\text{-}\ell_0$ (W)	20.17
3MG	$\ell_2\text{-}\ell_0$ (H)	20.05

Table 2. Reconstruction results for several optimization and regularization strategies, for Poly1 subsampling pattern.

art primal-dual convex optimization methods [21, 20, 22] and the Alternating-Direction Method of Multipliers (ADMM) [5, 18, 13] in terms of computation time for Matlab R2011b codes running on a single-core Intel i7-2620M CPU@2.7 GHz with 8 Gb of RAM, in the case of Poly1 sampling strategy. Although an ℓ_1 penalization is used in the other convex optimization approaches, it is worth noticing that the resulting SNR values reported in Table 2 are identical to those provided by 3MG with $\ell_2\text{-}\ell_1$. Moreover, as illustrated by Fig. 2, the proposed algorithm benefits from a faster convergence. A further advantage of 3MG is that it allows the use of nonconvex penalizations. In Table 2, we also indicate results obtained for some $\ell_2\text{-}\ell_0$ penalizations, namely Geman-McClure (G), Welsh (W) and Hyperbolic tangent (H) potentials (see [7, Sec.2.2]), for which the local convergence of 3MG algorithm is guaranteed. One can observe a quantitative improvement of the reconstruction quality with respect to the convex case.

6. CONCLUSION

In this paper, we have proposed an extension of the 3MG algorithm for the resolution of large-size optimization problems involving functions of complex variables. We have shown that the proposed algorithm is guaranteed to converge under weak assumptions. Its good numerical performance has been demonstrated in the context of complex-valued im-

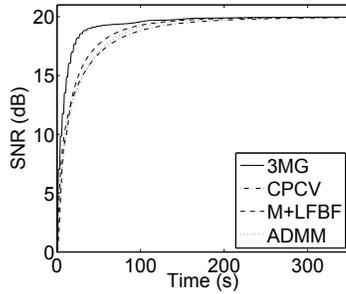


Fig. 2. SNR evolution as a function of computation time using 3MG, M+LFBF [20], CPCV [21, 22] and ADMM.

age reconstruction from real parallel MRI data.

Acknowledgements: The authors would like to thank N. Chauffert and S. Mériaux from NeuroSpin (CEA) for their help in providing the MRI data and the subsampling schemes.

7. REFERENCES

- [1] T. Adali, P. J. Schreier, and L. L. Scharf, “Complex-valued signal processing: the proper way to deal with impropriety,” *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5101–5125, Nov. 2011.
- [2] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, “An interior-point method for large-scale ℓ_1 -regularized least squares,” *IEEE J. Sel. Top. Sign. Proces.*, vol. 1, no. 4, pp. 606–617, Dec. 2007.
- [3] L. Sorber, M. Van Barel, and L. De Lathauwer, “Unconstrained optimization of real functions in complex variables,” *SIAM J. Optim.*, vol. 22, no. 3, pp. 879–898, Jul. 2012.
- [4] P. L. Combettes and J.-C. Pesquet, “Proximal splitting methods in signal processing,” in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, H. H. Bauschke, R. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, Eds., pp. 185–212. Springer-Verlag, New York, 2010.
- [5] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, “An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems,” *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, 2011.
- [6] E. Chouzenoux, J. Idier, and S. Moussaoui, “A majorize-minimize subspace strategy for subspace optimization applied to image restoration,” *IEEE Trans. Image Process.*, vol. 20, no. 18, pp. 1517–1528, Jun. 2011.
- [7] E. Chouzenoux, A. Jezierska, J.-C. Pesquet, and H. Talbot, “A majorize-minimize subspace approach for ℓ_2 - ℓ_0 image regularization,” *SIAM J. Imag. Sci.*, vol. 6, no. 1, pp. 563–591, 2013.
- [8] P. Ciuciu and J. Idier, “A half-quadratic block-coordinate descent method for spectral estimation,” *Signal Process.*, vol. 82, no. 7, pp. 941–959, 2002.
- [9] S. Husse, Y. Goussard, and J. Idier, “Extended forms of Geman Yang algorithm: application to MRI reconstruction,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP’04)*, 17-21 May 2004, vol. 3, pp. 513–516.
- [10] B. Liu, F. M. Seibert, Y. Zou, and L. Ying, “SparseSENSE: randomly-sampled parallel imaging using compressed sensing,” in *Proc. 16th Annual Meeting of ISMRM*, 3-8 May 2008, p. 3154.
- [11] L. Chaâri, J.-C. Pesquet, A. Benazza-Benyahia, and P. Ciuciu, “A wavelet-based regularized reconstruction algorithm for SENSE parallel MRI with applications to neuroimaging,” *Med. Image Anal.*, vol. 15, no. 2, pp. 185–201, Nov. 2011.
- [12] X. Ye, Y. Chen, W. Lin, and F. Huang, “Fast MR image reconstruction for partially parallel imaging with arbitrary k-space trajectories,” *IEEE Trans. Med. Imag.*, vol. 30, no. 3, pp. 575–585, 2011.
- [13] C. Bilen, Y. Wang, and I. W. Selesnick, “High-speed compressed sensing reconstruction in dynamic parallel MRI using augmented Lagrangian and parallel processing,” *IEEE J. Emerging Sel. Top. Circuits Syst.*, vol. 2, no. 3, pp. 370–379, Sep. 2012.
- [14] C. Boyer, P. Ciuciu, P. Weiss, and S. Mériaux, “HYR2PICS: hybrid regularized reconstruction for combined parallel imaging and compressive sensing in MRI,” in *Proc. 9th IEEE Int. Symp. Biomed. Imaging (ISBI’12)*, 2-5 May 2012, pp. 66–69.
- [15] M. Lustig, D. L. Donoho, and J. M. Pauly, “Sparse MRI: The application of compressed sensing for rapid MR imaging,” *Magn. Reson. Med.*, vol. 58, pp. 1182–1195, 2007.
- [16] G. Puy, J. P. Marques, R. Gruetter, J. Thiran, D. Van De Ville, P. Vanderghenst, and Y. Wiaux, “Spread spectrum magnetic resonance imaging,” *IEEE Trans. Med. Imag.*, vol. 31, no. 3, pp. 586–598, Mar. 2012.
- [17] M. Guerquin-Kern, M. Häberlin, K. P. Pruessmann, and M. Unser, “A fast wavelet-based reconstruction method for magnetic resonance imaging,” *IEEE Trans. Med. Imag.*, vol. 30, no. 9, pp. 1649–1660, Sep. 2011.
- [18] J. Aelterman, H. Q. Luong, B. Goossens, A. Pizurica, and W. Philips, “Augmented Lagrangian based reconstruction of non-uniformly sub-Nyquist sampled MRI data,” *Signal Process., Special Issue: Advances in Multirate Filter Bank Structures and Multiscale Representations*, vol. 91, no. 12, pp. 2731–2742, Dec. 2011.
- [19] N. Chauffert, P. Ciuciu, and P. Weiss, “Variable density compressed sensing in MRI. Theoretical vs heuristic sampling strategies,” in *Proc. 10th IEEE Int. Symp. Biomed. Imaging (ISBI’13)*, 7-11 Apr. 2013.
- [20] P. L. Combettes and J.-C. Pesquet, “Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators,” *Set-Valued Var. Anal.*, vol. 20, no. 2, pp. 307–330, Jun. 2012.
- [21] A. Chambolle and T. Pock, “A first-order primal-dual algorithm for convex problems with applications to imaging,” *J. Math. Imaging Vision*, vol. 40, no. 1, pp. 120–145, 2010.
- [22] L. Condat, “A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms,” to appear in *J. Optim. Theory Appl.*, Dec. 2012, hal.archives-ouvertes.fr/hal-00609728.