Combined Analysis-L1 and Total Variation ADMM with Applications to MEG Brain Imaging and Signal Reconstruction

Rui Gao, Filip Tronarp, and Simo Särkkä
Dept. Electrical Engineering and Automation, Aalto University
{rui.gao, filip.tronarp, simo.sarkka}@aalto.fi

***Abstract*—In this article, we propose an efficient method for solving analysis-l1-TV regularization problems with a multi-step alternating direction method of multipliers (ADMM) approach as the fast solver. Additionally, we apply it to a real-data magnetoencephalography (MEG) brain imaging problem as well as to signal reconstruction. In our approach, the inverse problem arising in MEG or signal reconstruction is formulated as an optimization problem which we regularize using a combination of analysis-l1 prior together with a total variation (TV) regularization term. We then formulate an optimization algorithm based on ADMM which can effectively be used to solve the optimization problems. The performance of the algorithm is illustrated in practical scenarios.**

***Index Terms*—Analysis-l1-TV-norm, total variation (TV), alternating direction method of multipliers (ADMM), magnetoencephalography (MEG), image reconstruction**

I. Introduction

Magnetoencephalography (MEG, see, e.g., [1]) is a non-invasive brain imaging method which can be used to observe the brain activity by measuring the magnetic field generated by the neurons in the brain. The reconstruction of the brain activity from the MEG sensor measurements requires solving an inverse problem [2], which refers to the ill-posed problem of reconstructing the unknown quantity (in MEG, the brain activity on the cortex [1]) from a limited number of measurements. One way to solve such inverse problem is to optimize a regularized cost function which combines a data-dependent term with an ℓ_1/ℓ_2 or total variation penalty term [1], [3]–[8].

Another application where an inverse problem arises is image reconstruction, where in addition to the classical ℓ_2, it is commonplace to use ℓ_1-norm [9], total variation (TV) norm [10], or their combinations [11]. In particular, the TV regularization can be used to introduce gradient sparsity, which is commonly occurring in natural images [12]. It is also possible to combine the regularization terms leading to ℓ_1-TV and ℓ_2-TV based regularization methods [9], [12]–[16]. They have the ability to preserve edge information and allow for sharp discontinuities in the solution.

Recently, a new family of *analysis*-ℓ_1-TV regularization methods has been proposed to achieve high-quality reconstruction, especially in case of severe corruption [17]–[19]. Assuming an analysis sparse prior, the methods minimize a cost function with combined analysis ℓ_1-norm and TV regularization terms. Such analysis sparse methods could provide more flexibility in modeling of signals than *synthesis sparsity* methods [20]. Despite good performance that can be obtained, the methods remain computationally difficult due to the non-linearity and non-differentiability [17]. Although some optimization methods have been proposed to solve these difficulties, for example, templates for first-order conic solvers (TFOCS) [21], the primal-dual Newton conjugate gradient method (pdNCG) [22], and alternating direction methods [16], there is still a need for efficient and effective methods for reconstructing high-dimensional signals.

The main contribution of this article is the development of an efficient method for solving analysis-ℓ_1-TV regularization problems with a multi-step ADMM approach [23] as the fast solver, which is demonstrated in an MEG inverse problem with real data, image reconstruction, as well as one-dimensional signal reconstruction. The advantages of the proposed approach are low number of iterations needed for convergence and good convergence speed in terms of CPU time. We empirically compare the performance of the method against the TFOCS and pdNCG solvers. The results demonstrate that our method has a superior convergence speed in terms of number of iterations and CPU time.

II. Problem Formulation

Let the observed signal $\mathbf{y} \in \mathbb{R}^{n_y}$ and the unknown source signal $\mathbf{x} \in \mathbb{R}^{n_x}$, where we typically have $n_y \leq n_x$. The sources and measurements are related according to

$$\mathbf{y} = \mathbf{M}\mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{M}$ is the matrix representation of a linear operator and $\mathbf{n}$ denotes the observation noise that, without loss of generality, is taken to have an identity covariance matrix. Since the inverse of the operator $\mathbf{M}$ is typically ill-conditioned (e.g. in the case of blurring filter in image reconstruction), or it does not even exist (e.g. a forward-operator in MEG source reconstruction), Eq. (1) defines a non-trivial inverse problem. However, the prior knowledge about the unknown signal $\mathbf{x}$ makes its solution possible.

Here, we exploit the analysis-sparsity structure of the signal in some transform domain, and reconstruct $\mathbf{x}$ from the

Thanks to Academy of Finland for funding.

ISBN 978-90-827970-1-5 © EURASIP 2018

observation model in (1) by solving an optimization problem. Specifically, given an analysis operator $\mathbf{\Omega} \in \mathbb{R}^{n_x \times n_x}$, $\mathbf{\Omega x}$ is assumed to be cosparse, that is, $\mathbf{\Omega x}$ has a small number of non-zero entries [17]. Similar to synthesis sparsity [20], a well-studied analysis-ℓ_1-sparse approach is [17]–[19]:

$$\min_{\mathbf{x}} \ \|\mathbf{y}-\mathbf{M}\mathbf{x}\|_2^2, \quad \text{such that} \quad \|\mathbf{\Omega x}\|_1 \le T, \tag{2}$$

where T is a tolerance error, and $\|\cdot\|_1$ and $\|\cdot\|_2$ are the ℓ_1 and ℓ_2 norms, respectively. The operator $\mathbf{\Omega}$ can be generated, for example, by the discrete Fourier transform [9], tight frame [12], wavelet [24], or learned methods [18], [19]. In addition, we use a TV-term to promote gradient sparsity, which allows for sharp discontinuities in the original signal. These considerations suggest that $\mathbf{x}$ can be estimated by the following problem:

$$\mathbf{x} = \operatorname*{argmin}_{\mathbf{x}} \ \frac{\mu}{2}\|\mathbf{y}-\mathbf{M}\mathbf{x}\|_2^2 + \lambda\|\mathbf{\Omega x}\|_1 + \|\mathbf{x}\|_{\text{tv}}, \tag{3}$$

where $\|\mathbf{x}\|_{\text{tv}}$ is the discrete TV-norm, and μ and $\lambda > 0$ are tuning parameters. Due to the nonlinearity and non-differentiability, it is difficult to compute (3). In the following sections, we will propose an efficient algorithm to solve (3) by the multi-step ADMM (m-ADMM) algorithm .

III. ADMM for Signal Reconstruction

This section presents a new method for solving the combined analysis-ℓ_1-TV model based on the standard ADMM algorithm [23], [25], [26]. ADMM is a versatile algorithm for solving high-dimensional signal processing problems efficiently. It decomposes a large global problem into a series of smaller local sub-problems, and uses the local solutions to compute the globally optimal solution.

Our solution is a multi-step extension of the ADMM method, which iteratively updates the multiple primal variables and the corresponding dual variables of the augmented Lagrangian function of (3). Specifically, we introduce the auxiliary variables $\mathbf{v}$ and $\mathbf{w}$, and then reformulate (3) in ADMM form as

$$\begin{aligned} &\min_{\mathbf{x},\mathbf{v},\mathbf{w}} \ \frac{\mu}{2}\|\mathbf{y}-\mathbf{M}\mathbf{x}\|_2^2 + \lambda\|\mathbf{v}\|_1 + \|\mathbf{w}\|_2, \\ &\text{such that } \mathbf{v} = \mathbf{\Omega x}, \ \ \mathbf{w}_i = D_i\mathbf{x}, \ \ i = 1, \cdots, n_y, \end{aligned} \tag{4}$$

where $\mathbf{w} = [\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_{n_y}]$, and $D_i\mathbf{x} \in \mathbb{R}^d$ represents the first-order finite difference of $\mathbf{x}$ at i:th component in d different directions.

Let $\mathcal{L}_\rho(\mathbf{x},\mathbf{v},\mathbf{w};\boldsymbol{\eta},\boldsymbol{\zeta})$ be the augmented Lagrangian function of (4) which is defined as follows

$$\begin{aligned} \mathcal{L}_\rho(\mathbf{x},\mathbf{v},\mathbf{w};\boldsymbol{\eta},\boldsymbol{\zeta}) \triangleq{}& \frac{\mu}{2}\|\mathbf{y}-\mathbf{M}\mathbf{x}\|_2^2 + \lambda\|\mathbf{v}\|_1 + \|\mathbf{w}\|_2 \\ &+ \boldsymbol{\eta}^\top(\mathbf{v}-\mathbf{\Omega x}) + \frac{\rho_1}{2}\|\mathbf{v}-\mathbf{\Omega x}\|_2^2 \\ &+ \boldsymbol{\zeta}^\top(\mathbf{w}-\mathbf{D}\mathbf{x}) + \frac{\rho_2}{2}\|\mathbf{w}-\mathbf{D}\mathbf{x}\|_2^2, \end{aligned} \tag{5}$$

where $\mathbf{D} = [D_1, D_2, \cdots, D_{n_y}]$, $\boldsymbol{\eta} \in \mathbb{R}^{n_x}$ and $\boldsymbol{\zeta} \in \mathbb{R}^{d \times n_y}$ are the Lagrange multipliers, and $\rho_1, \rho_2 > 0$ are penalty parameters. Since the variables in $\mathcal{L}_\rho(\mathbf{x},\mathbf{v},\mathbf{w};\boldsymbol{\eta},\boldsymbol{\zeta})$ are coupled together, it is difficult to solve them simultaneously. Therefore, we apply a method that separates (5) into local sub-problems.

In each iteration, the function (5) is minimized over $\mathbf{x}$, $\mathbf{v}$, and $\mathbf{w}$ separately, one after the other, followed by two dual updates for $\boldsymbol{\eta}$ and $\boldsymbol{\zeta}$. For example, starting at $\mathbf{x} = \mathbf{x}^k$, $\mathbf{v} = \mathbf{v}^k$, $\mathbf{w} = \mathbf{w}^k$, $\boldsymbol{\eta} = \boldsymbol{\eta}^k$ and $\boldsymbol{\zeta} = \boldsymbol{\zeta}^k$:

$$\mathbf{x}^{k+1} = \operatorname*{argmin}_{\mathbf{x}} \ \mathcal{L}_\rho(\mathbf{x},\mathbf{v}^k,\mathbf{w}^k;\boldsymbol{\eta}^k,\boldsymbol{\zeta}^k) \tag{6a}$$

$$\mathbf{v}^{k+1} = \operatorname*{argmin}_{\mathbf{v}} \ \mathcal{L}_\rho(\mathbf{x}^{k+1},\mathbf{v},\mathbf{w}^k;\boldsymbol{\eta}^k,\boldsymbol{\zeta}^k) \tag{6b}$$

$$\mathbf{w}^{k+1} = \operatorname*{argmin}_{\mathbf{w}} \ \mathcal{L}_\rho(\mathbf{x}^{k+1},\mathbf{v}^{k+1},\mathbf{w};\boldsymbol{\eta}^k,\boldsymbol{\zeta}^k) \tag{6c}$$

$$\boldsymbol{\eta}^{k+1} = \boldsymbol{\eta}^k + \rho_1(\mathbf{\Omega}\mathbf{x}^{k+1} - \mathbf{v}^{k+1}) \tag{6d}$$

$$\boldsymbol{\zeta}^{k+1} = \boldsymbol{\zeta}^k + \rho_2(\mathbf{D}\mathbf{x}^{k+1} - \mathbf{w}^{k+1}). \tag{6e}$$

The efficiency of m-ADMM relies on the fact that its sub-problems have exact solutions or can be solved effectively to high accuracy. Explicit solutions to the above sub-problems are the following.

1) Solving $\mathbf{x}$ *sub-problem*. The minimization in (6a) is a quadratic optimization problem. It can be analytically solved by direct dense or sparse matrix computations. Moreover, under periodic boundary conditions, the computations can utilize fast transforms such as the fast Fourier transform (FFT) [15] or wavelet transform [24]. However, when the primal variable $\mathbf{x}$ is high-dimensional, the optimization problem (6a) is costly to solve because the matrix computations $\mathbf{\Omega}^\top\mathbf{\Omega}$, $\mathbf{D}^\top\mathbf{D}$ and $\mathbf{M}^\top\mathbf{M}$ are involved in the quadratic penalty terms. Thus, the problem (6a) is usually redefined to avoid the expensive computations. One approach is to linearize the quadratic term with adding a proximal term [27]. Here, we introduce a proximal term $\frac{1}{2}\|\mathbf{x}-\mathbf{x}^k\|_{\mathbf{G}}^2$ with $\mathbf{G} := \frac{1}{\delta}\mathbf{I} - (\rho_1\mathbf{\Omega}^\top\mathbf{\Omega} + \rho_2\mathbf{D}^\top\mathbf{D} + \mu\mathbf{M}^\top\mathbf{M})$ so that no matrix inversion is required. We use only the first-order optimality condition and effectively computes the matrix-vector multiplications. Then, the solution can be seen as linearizing the quadratic penalty term using the gradient at $\mathbf{x}^k$ plus a proximal term $\frac{1}{2\delta}\|\mathbf{x}-\mathbf{x}^k\|_2^2$.

Typically, the primal variable $\mathbf{x}$ in (6a) is obtained by solving the following problem:

$$\begin{aligned} \mathbf{x}^{k+1} &= \operatorname*{argmin}_{\mathbf{x}} \ \mathcal{L}_\rho(\mathbf{x},\mathbf{v}^k,\mathbf{w}^k;\boldsymbol{\eta}^k,\boldsymbol{\zeta}^k) + \frac{1}{2}\left\|\mathbf{x}-\mathbf{x}^k\right\|_{\mathbf{G}}^2 \\ &= \operatorname*{argmin}_{\mathbf{x}} \frac{\mu}{2}\|\mathbf{y}-\mathbf{M}\mathbf{x}\|_2^2 + \frac{\rho_1}{2}\left\|\mathbf{v}^k - \mathbf{\Omega x} + \frac{\boldsymbol{\eta}^k}{\rho_1}\right\|_2^2 \\ &\quad + \frac{\rho_2}{2}\left\|\mathbf{w}^k - \mathbf{D}\mathbf{x} + \frac{\boldsymbol{\zeta}^k}{\rho_2}\right\|_2^2 + \frac{1}{2}(\mathbf{x}-\mathbf{x}^k)^\top\mathbf{G}(\mathbf{x}-\mathbf{x}^k), \end{aligned} \tag{7}$$

which is quadratic in $\mathbf{x}$ and therefore

$$\begin{aligned} \mathbf{x}^{k+1} &= \mathbf{x}^k - \delta(\mathbf{\Omega}^\top\boldsymbol{\eta}^k + \mathbf{D}^\top\boldsymbol{\zeta}^k) + \delta\rho_1\mathbf{\Omega}^\top(\mathbf{\Omega}\mathbf{x}^k - \mathbf{v}^k) \\ &\quad + \delta\rho_2\mathbf{D}^\top(\mathbf{D}\mathbf{x}^k - \mathbf{w}^k) + \delta\mu\mathbf{M}^\top(\mathbf{y} - \mathbf{M}\mathbf{x}^k). \end{aligned} \tag{8}$$

2) Solving $\mathbf{v}, \mathbf{w}$ *sub-problems*. Fixing $\mathbf{x}^k$, $\boldsymbol{\eta}^k$ and $\boldsymbol{\zeta}^k$, we can sequentially update the variables $\mathbf{v}^{k+1}$ in (6b) and $\mathbf{w}^{k+1}$ in (6c). The updating of $\mathbf{v}^{k+1}$ is expressed as

$$\mathbf{v}^{k+1} = \operatorname*{argmin}_{\mathbf{v}} \ \lambda\|\mathbf{v}\|_1 + \frac{\rho_1}{2}\left\|\mathbf{v} - \mathbf{\Omega}\mathbf{x}^{k+1} + \frac{\boldsymbol{\eta}^k}{\rho_1}\right\|_2^2, \tag{9}$$

ISBN 978-90-827970-1-5 © EURASIP 2018

which is solved by the one-dimensional shrinkage [9], [12]

$$\mathbf{v}^{k+1} = \text{sign}(\mathbf{e}^k) \circ \max(|\mathbf{e}^k| - \lambda/\rho_1, 0), \tag{10}$$

where $\mathbf{e}^k = \mathbf{\Omega}\mathbf{x}^{k+1} + \boldsymbol{\eta}^k/\rho_1$, sign represents the signum function, and $\circ$ is the pointwise product.

Similarly, the sub-problem for $\mathbf{w}^{k+1}$ can be reduced to the following minimization problem:

$$\mathbf{w}^{k+1} = \underset{\mathbf{w}}{\text{argmin}} \ \|\mathbf{w}\|_2 + \frac{\rho_2}{2}\left\|\mathbf{w} - \mathbf{D}\mathbf{x}^{k+1} + \frac{\boldsymbol{\zeta}^k}{\rho_2}\right\|_2^2 \tag{11}$$

and the update of $\mathbf{w}^{k+1}$ is given explicitly by multi-dimensional shrinkage [9], [12]

$$\mathbf{w}^{k+1} = \max\left\{\left\|\mathbf{t}^k\right\|_2 - \frac{1}{\rho_2}, 0\right\}\frac{\mathbf{t}^k}{\left\|\mathbf{t}^k\right\|_2} \tag{12}$$

with $\mathbf{t}^k = \mathbf{D}\mathbf{x}^{k+1} + \boldsymbol{\zeta}^k/\rho_2$.

3) Solving dual sub-problems. Update the dual variables $\boldsymbol{\eta}$, $\boldsymbol{\zeta}$ via (6d) and (6e), respectively.

We terminate the iteration when a stopping criterion is satisfied, for example, when the magnitude of the primal and dual residuals are below a given, application specific, threshold [23]. The m-ADMM method is summarized in Algorithm 1.

Algorithm 1 Combined analysis-ℓ_1-TV method by m-ADMM

Input: $\mathbf{y}$, μ, λ, ρ_1,ρ_2

1: Initialization: Start from initial values $\mathbf{x}^0, \mathbf{v}^0, \mathbf{w}^0, \boldsymbol{\eta}^0, \boldsymbol{\zeta}^0$. Set $k = 0$, $\delta \in (0, \frac{1}{\rho_1\|\mathbf{\Omega}\|^2 + \rho_2\|\mathbf{D}\|^2 + \mu\|\mathbf{M}\|^2})$.
2: **repeat**
3: update $\mathbf{x}^{k+1}$ by computing (8)
4: update $\mathbf{v}^{k+1}$ by computing (10)
5: update $\mathbf{w}^{k+1}$ by computing (12)
6: update $\boldsymbol{\eta}^{k+1}$ by computing (6d)
7: update $\boldsymbol{\zeta}^{k+1}$ by computing (6e)
8: $k = k + 1$
9: **until** (*stopping criterion is satisfied*)

Output: $\mathbf{x}$

IV. Experimental Results

In this section, we demonstrate the efficiency of m-ADMM in three applications: one-dimensional signal reconstruction, image reconstruction, and MEG brain imaging.

A. One-dimensional signal reconstruction

In this first experiment, we consider the recovery of a piecewise continuous signal. Fig. 1(a) shows the synthetic clean signal $\mathbf{x}_{\text{true}}$. The corrupted signal $\mathbf{y}$ (see Fig. 1(b)) is obtained using a Gaussian blurring operator $\mathbf{M}$ with standard deviation 0.1 and an additive white Gaussian noise with standard deviation $\sigma = 0.01$. The goal is to estimate the signal from the corrupted signal. We set $\rho_1 = 1$, $\rho_2 = 4$, $\mu = 1$, and $\lambda = 0.5$ in the experiment. In particular, we solve every $\mathbf{x}$-iteration by direct matrix computations, because n_x is small ($n_y = n_x = 1001$). The result is shown in Fig. 1.

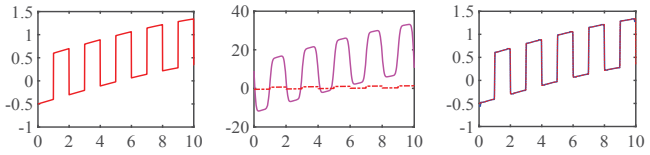

Figure 1. One-dimensional signal reconstruction experiment: (a) the synthetic clean signal $\mathbf{x}_{\text{true}}$, (b) measured signal $\mathbf{y}$, and (c) reconstructed signal (blue solid line). The red dashed line is the ground truth.

For further comparison, we demonstrate the efficiency of the proposed method against two state-of-the-art methods: pdNCG [22] and TFOCS [21]. TFOCS solves the problem (3) by introducing an extra regularization term and pdNCG uses the pseudo-Huber as the proximal term of $\|\cdot\|_1$–norm. Although the two methods introduce different parameters and terms, the objective of these is to solve the combined analysis-ℓ_1-TV problem (3). We used the same parameters $\mu = 1$ and $\lambda = 0.5$ in all the methods to make a fair comparison of their performance. Other parameter settings of the methods were tuned by trial to give the fastest convergence. Moreover, there were no prerequisite conditions on matrices $\mathbf{M}$ and $\mathbf{\Omega}$. All the experiments were performed on a PC running Intel 2.66 GHz quad-core CPU.

Fig. 2 shows the performance of three different methods, where the relative error $\left\|\mathbf{x}^k - \mathbf{x}_{\text{true}}\right\|_2 / \left\|\mathbf{x}_{\text{true}}\right\|_2$ versus the number k of iterations are plotted. As can been seen in Fig. 2, all the methods can reconstruct the clean signal from the measured signal $\mathbf{y}$ and converge to a similar solution in two noise levels ($\sigma \in \{0.01, 1\}$). The proposed method has superior convergence properties compared to pdNCG and TFOCS solvers. Table I summarizes the running time of each solver and the required iteration counts. It should be noted that pdNCG solver uses around 22.7 seconds to achieve 11 iterations, and our method takes 20 iterations to complete the run in about 3.4 seconds.

Table I
Run-times and iteration counts of the methods.

Methods	pdNCG	TFOCS	m-ADMM
Iteration count	11	200	20
CPU time (sec)	22.7	16.34	3.4

B. Application to image reconstruction

In this experiment, we consider the problem of image reconstruction of a blurry and noisy image. The corrupted image $\mathbf{y}$ (see Fig. 3 (a)) is generated by an operator $\mathbf{M}$, corresponding to a Gaussian blur of size 7×7 and standard deviation 5, and adding Gaussian noise with standard deviation 10, from the original 'pepper' image[1] $\mathbf{x}$.

We initialize the variables $\mathbf{v}^0, \mathbf{w}^0, \boldsymbol{\eta}^0, \boldsymbol{\zeta}^0$ to be zero. The parameter μ aims at controlling the smoothness. The large value of μ may give some visually appealing image reconstruction [9].

[1] http://sipi.usc.edu/database/database.php

ISBN 978-90-827970-1-5 © EURASIP 2018

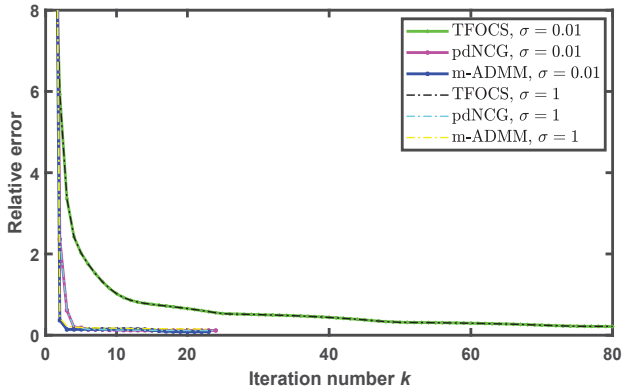

Figure 2. Relative error versus iteration number k for three mentioned methods in different noisy cases.

We set $\mu = 60$ in the our experiment. The values of penalty parameters ρ_1 and ρ_2 balance the influence of the primal and dual residuals [23]. The parameters ρ_1 and ρ_2 are set by first selecting the ratio, ρ_1/ρ_2, which influences the convergence in practice, and then adjusting one of the parameters, which in turn fixes the remaining one.

To restore $\mathbf{x}$, we choose the dual-tree complex wavelet transform [28] as $\mathbf{\Omega}$, which also satisfies the tight frame condition with periodic boundary conditions [24], [29]. The other parameters of our model are set as: $\rho_1 = 3$, $\rho_2 = 10$, $\lambda = 0.8$. We use the index peak signal to noise ratio (PSNR) and the structural similarity (SSIM) index [30] to measure the image quality quantitatively. Fig. 3 (b) and (c) show the restored results from fast TV regularization method [16] with 100 iterations and our proposed method with 30 iterations, respectively. In order to better visualize the results and their comparison, the magnified details of the images are shown in Fig. 3 (d), (e) and (f). As shown in Fig. 3, the proposed method keeps more structural information and produces a better image reconstruction than the other approach.

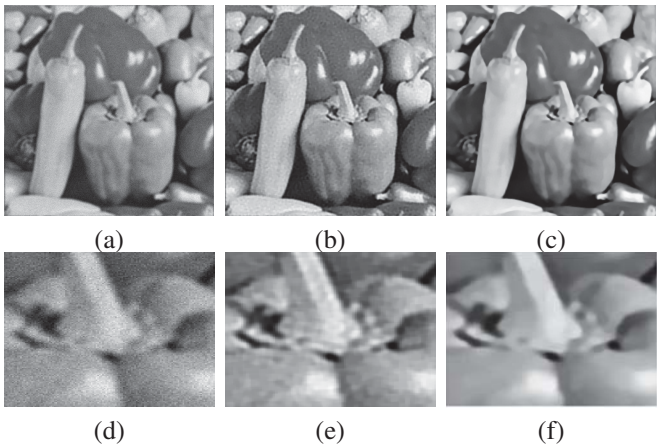

Figure 3. Results of image reconstruction (a) corrupted image, PSNR = 24.49dB, SSIM = 0.49; (b) restored image by fast TV regularization method with $k = 100$, PSNR = 27.59dB, SSIM = 0.76; (c) restored image by the proposed method with $k = 30$, PSNR = 28.55dB, SSIM = 0.86; The magnified details of the images in (d), (e) and (f).

C. Source reconstruction in MEG

In this section, we apply our method to a real-data MEG brain imaging problem $\mathbf{y} = \mathbf{M}\mathbf{x} + \boldsymbol{\varepsilon}$, where $\mathbf{y} \in \mathbb{R}^{n_y}$ is a whitened measurement, $\mathbf{x} \in \mathbb{R}^{n_x}$, $n_y < n_x$, $\boldsymbol{\varepsilon}$ is a measurement error with identity covariance, and $\mathbf{M}$ is a whitened, cortically constrained, lead-field matrix. Furthermore, the number of MEG sensors is, $n_y = 305$, which measure the electromagnetic activity of, $n_x = 7498$, sources.

In order to promote time-continuity, two measurement instants are considered at a time, $\mathbf{y}(t) \triangleq (\mathbf{y}_t^\top, \mathbf{y}_{t-1}^\top)^\top$, corresponding to source activities $\mathbf{x}(t) \triangleq (\mathbf{x}_t^\top, \mathbf{x}_{t-1}^\top)^\top$, and the estimate of $\mathbf{x}_t$, $\hat{\mathbf{x}}_t$, is obtained by taking the appropriate component of the solution of the following optimization problem

$$\underset{\mathbf{x}(t)\in\mathbb{R}^{2n_x}}{\arg\min} \frac{1}{2}||\mathbf{y}(t)-\mathbf{I}_2\otimes\mathbf{M}\mathbf{x}(t)||_2^2+\lambda||\mathbf{I}_2\otimes\mathbf{\Omega}\mathbf{x}(t)||_1+||\mathbf{D}\mathbf{x}(t)||_2$$

where $\otimes$ is Kronecker's product, $\mathbf{\Omega}$ is a matrix of depth-weights [31], and $\mathbf{D}\mathbf{x}(t) = \mathbf{x}_t - \mathbf{x}_{t-1}$. The parameter, λ, is selected in correspondence to the recommendations of [31].

The proposed scheme uses $\rho_1 = \rho_2 = \sqrt{2.5}/2$, and is compared to the depth-weighted minimum norm [31] solution as well as to the depth-weighted minimum current solution [3] on the sample data set from from MNE software package [6], where a healthy volunteer is subjected to an auditory stimuli.

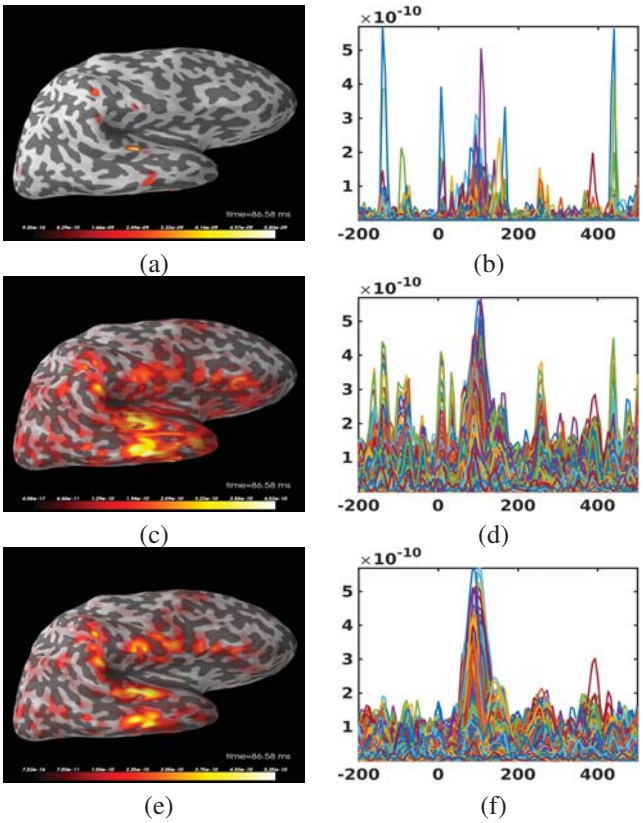

Figure 4. The source current image of the right hemisphere at the time of activity and source time courses for ℓ_1 (a,b), ℓ_2 (c,d), and m-ADMM (e,f) methods, respectively.

ISBN 978-90-827970-1-5 © EURASIP 2018

While all three methods reproduce activity in the superior temporal gyrus, where the primary auditory cortex is located, as can be seen from Fig. 4, the benefit of the proposed method is that it obtains a sparser solution than the ℓ_2 method, while in contrast to both the ℓ_1 and ℓ_2 methods, producing vastly smoother time courses, which is neuroscientifically preferable [5].

V. Conclusion

In this article, we have presented an efficient method for solving analysis-ℓ_1-TV regularization problems with a multi-step ADMM approach as the fast solver. The approach is derived from ADMM method, and it enables high-dimensional optimization problems to be solved by decomposing them to local sub-problems. The proposed approach is applied to one-dimensional signal reconstruction, image reconstruction, and MEG source reconstruction. Experimental results demonstrate the effectiveness of the proposed method in combined analysis-ℓ_1-TV regularization.

As shown by the results, the advantages of our method are the low number of iterations needed for convergence and good convergence speed in terms of CPU time. Other advantages are the conceptual simplicity and parallelizability that would allow for fast parallel implementations [32], for example, using GPU computations.

VI. Acknowledgements

The authors would like to thank Sara Sommariva and Narayan Subramaniyam for their many helpful comments.

References

[1] M. Hämäläinen, R. Hari, R. J. Ilmoniemi, J. Knuutila, and O. V. Lounasmaa, "Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain," *Rev. Mod. Phys.*, vol. 65, no. 2, pp. 413, Apr. 1993.

[2] J. Kaipio and E. Somersalo, *Statistical and computational inverse problems*, Appl. Math. Sci. Springer, 2005.

[3] I. F. Gorodnitskya, J. S. George, and B. D. Rao, "Neuromagnetic source imaging with FOCUSS: a recursive weighted minimum norm algorithm," *Electroencephalogr. Clin. Neurophysiol.*, vol. 95, no. 4, pp. 231–251, Oct. 1995.

[4] W. Ou, M. S. Hämäläinen, and P. Golland, "A distributed spatio-temporal EEG/MEG inverse solver," *NeuroImage*, vol. 44, no. 3, pp. 932–946, 2009.

[5] A. Gramfort, M. Kowalsk, and M. Hämäläinen, "Mixed-norm estimates for the M/EEG inverse problem using accelerated gradient methods," *Phts. Med. Bio.*, vol. 57, no. 7, pp. 1937–1961, 2012.

[6] A. Gramfort, M. Luessiand, E. Larsonand, D. A. Engemann, D. Strohmeierand, C. Brodbeckand, L. Parkkonenand, and M. S. Hämäläinen, "MNE software for processing MEG and EEG data," *NeuroImage*, vol. 86, pp. 446–460, Feb. 2014.

[7] G. Adde, M. Clerc, and R. Keriven, "Imaging methods for MEG/EEG inverse problem," *Int. J. Bioelectromagnetism*, vol. 7, no. 2, pp. 111–114, 2005.

[8] H. Becker, L. Albera, P. Comon, R. Gribonval, and I. Merlet, "Fast, variation-based methods for the analysis of extended brain sources," in *Proc. IEEE 22nd Int. Conf. European Signal Process.* IEEE, 2014, pp. 41–45.

[9] M. Almeida and M. Figueiredo, "Deconvolving images with unknown boundaries using the alternating direction method of multipliers," *IEEE Trans. Image Proc.*, vol. 22, no. 8, pp. 3074–3086, Aug. 2013.

[10] S. Yang, J. Wang, W. Fan, X. Zhang, P. Wonka, and J. Ye, "An efficient ADMM algorithm for multidimensional anisotropic total variation regularization problems," in *Proc. ACM 19th SIGKDD Int. Conf. Knowl. Disc. and Data Min.*, Chicago, Illinois, USA, Aug. 2013.

[11] M. Zhang, C. Desrosiers, and C. Zhang, "Effective compressive sensing via reweighted total variation and weighted nuclear norm regularization," in *Proc. IEEE 42st Int. Conf. Acoustics, Speech and Signal Process.*, New Orleans, LA, USA, Mar. 2017, pp. 1802–1806.

[12] R. H. Chan, M. Tao, and X. Yuan, "Constrained total variation deblurring models and fast algorithms based on alternating direction method of multipliers," *SIAM J. Imag. Sci.*, vol. 6, no. 1, pp. 680–697, 2013.

[13] A. Beck, S. Sabach, and M. Teboulle, "An alternating semiproximal method for nonconvex regularized structured total least squares problems," *SIAM J. Matrix Anal. & Appl.*, vol. 7, no. 3, pp. 1129–1150, 2016.

[14] T. Goldstein and S. Osher, "The split bregman algorithm for L1 regularized problems," *SIAM J. Imag. Sci.*, vol. 2, no. 2, pp. 323–343, 2009.

[15] J. F. Yang, Y. Zhang W. Yin, and Y. Wang, "A fast algorithm for edge-preserving variational multichannel image restoration," *SIAM J. Imag. Sci.*, vol. 2, no. 2, pp. 569–592, 2009.

[16] S. H. Chan, R. Khoshabeh, K.B. Gibson, P.E. Gill, and T.Q. Nguyen, "An augmented Lagrangian method for total variation video restoration," *IEEE Trans. Image Process.*, vol. 20, no. 11, pp. 3097–3111, 2011.

[17] M. Elad, P. Milanfar, and R. Rubinstein, "Analysis versus synthesis in signal priors," *Inv. Probl.*, vol. 23, no. 3, pp. 947–968, Sept. 2007.

[18] S. Hawe, M. Kleinsteuber, and K. Diepold, "Analysis operator learning and its application to image reconstruction," *IEEE Trans. Image Process.*, vol. 22, no. 6, pp. 2138–2150, 2013.

[19] R. Gao, S. A. Vorobyov, and H. Zhao, "Image fusion with cosparse analysis operator," *IEEE Signal Process. Lett.*, vol. 24, no. 7, pp. 943–947, July 2017.

[20] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[21] S. R. Becker, E. J. Candés, and M. C. Grant, "Templates for convex cone problems with applications to sparse signal recovery," *Math. Program. Comput.*, vol. 3, no. 3, pp. 165–218, Sept. 2011.

[22] I. Dassios, K. Fountoulakis, and J. Gondzio, "A preconditioner for a primal-dual Newton conjugate gradients method for compressed sensing problems," *SIAM J. Sci. Comput.*, vol. 37, no. 6, pp. A2783–A2812, July 2015.

[23] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning.*, vol. 3, no. 1, pp. 1–122, 2011.

[24] A. Parekh and I. W. Selesnick, "Convex denoising using non-convex tight frame regularization," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1786–1790, May 2015.

[25] J. Eckstein and D. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Math. Program.*, vol. 55, no. 3, pp. 293–318, Apr. 1992.

[26] C. Chen, B. He, X. Yuan, and Y. Ye, "The direct extension of ADMM for multi-block convex minimization problems is not necessarily convergent," *Math. Program.*, vol. 155, no. 1, pp. 57–79, Jan. 2016.

[27] N. Parikh and S. Boyd, "Proximal algorithms," *Foundations and Trends in Optimization*, vol. 1, no. 3, pp. 123–231, 2014.

[28] I. W. Selesnick, R. G. Baraniuk, and N. C. Kingsbury, "The dual-tree complex wavelet transform," *IEEE Signal Proc. Mag.*, vol. 22, no. 6, pp. 123 –151, Dec. 2005.

[29] M. Kabanava, H. Rauhut, and H. Zhang, "Robust analysis L1-recovery from Gaussian measurements and total variation minimization," *Eur. J. Appl. Math.*, vol. 26, no. 6, pp. 917–929, Dec. 2015.

[30] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[31] F. Lin, T. Witzel, S. P. Ahlfors, S. M. Stufflebeam, J. W. Belliveau, and M. S. Hämäläinen, "Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates," *NeuroImage*, vol. 31, no. 1, pp. 160–171, 2006.

[32] T. Lin, S. Ma, and S. Zhang, "Iteration complexity analysis of multi-block ADMM for a family of convex minimization without strong convexity," *J. Sci. Comput.*, vol. 69, no. 1, pp. 52–81, Oct. 2016.

ISBN 978-90-827970-1-5 © EURASIP 2018