

Sparse Bayesian Image Restoration with Linear Operator Uncertainties with Application to EEG Signal Recovery

Lotfi Chaari, Hadj Batatia and Jean-Yves Tournet

Abstract—Sparse signal/image recovery is a challenging topic that has captured a great interest during the last decades, especially in the biomedical field. Many techniques generally try to regularize the considered ill-posed inverse problem by defining appropriate priors for the target signal/image. The target regularization problem can then be solved either in a variational or Bayesian context. However, a little interest has been devoted to the uncertainties about the linear operator, which can drastically alter the reconstruction quality. In this paper, we propose a novel method for signal/image recovery that accounts and corrects the linear operator imprecisions. The proposed approach relies on a Bayesian formulation which is applied to EEG signal recovery. Our results show the promising potential of the proposed method compared to other regularization techniques which do not account for any error affecting the linear operator.

Index Terms—Sparse restoration, Hierarchical Bayesian Model, MCMC, linear operator, EEG/MEG

I. INTRODUCTION

In biomedical engineering, solving inverse problems is a challenging issue especially when these problems are ill-posed. Indeed, most of the medical imaging modalities lead to solving some inverse problem in order to recover the target signal or image. The associated observation model generally involves an observation operator which may be *a priori* estimated. However, the knowledge of this observation operator is generally affected by errors or uncertainties which can drastically decrease the reconstruction quality. This problem is for example widely known in many medical signal and image processing applications such as parallel magnetic resonance imaging [1]. These applications require the analysis of signals such as electrocardiograms, electroencephalograms, magnetoencephalograms, magnetic resonance imaging (MRI) or positron emission tomography (PET). In this paper, we focus on Electroencephalographic/Magnetoencephalographic (EEG/MEG) source localization. The standard observation model associated with EEG or MEG signals involves a linear operator called the lead field matrix, which makes the link between the sources of the electrical signal within the brain and the observations. In most of the cases, a realistic calculation of this operator [2] is subject to some imprecisions. These imprecisions make the source localization from dynamic EEG/MEG recordings difficult

Part of this work has been conducted within the image processing trimester of the CIMI Labex. This work has also been performed in the frame of the Hypanema ANR Project (ANR-12-BS03-003) and the STIC AmSud project DynBrain.

University of Toulouse, IRIT - INP-ENSEEIH, France
firstname.lastname@enseeiht.fr

and require to solve an ill-posed inverse problem. In addition, another source of artifacts may interfere with the useful signal such as power grid noise, eye movements, heart beat or muscle activities. For all these reasons, regularization is generally essential to regularize the inverse problem and improve the accuracy of source localization. In this context, a number of recent works have been developed during the last two decades to handle the regularization problem for EEG/MEG signals [3–6]. However, all these works do not account for any eventual imperfection resulting from the estimation of the lead field matrix. These methods try to correct the second source of artifacts (power grid noise, eye movements,...) by using appropriate prior information in the reconstruction model. Due to the nature of EEG/MEG signals, sparse priors have for instance been widely investigated [4, 5].

In this paper, we propose a new method for sparse regularization of EEG/MEG signals accounting for observation operator errors. The proposed method is developed in a Bayesian framework by using appropriate priors for the observation operator as well as for the target signal. This method provides estimation for both the target signal and the linear operator of the observation model. In addition, the other model parameters and hyperparameters are automatically estimated from the data. A hierarchical Bayesian model is built based on the likelihood and the used priors. A Markov Chain Monte Carlo (MCMC) [7] algorithm is then used to draw samples according to the target distribution and derive accurate estimates afterward.

The rest of the paper is organized as follows. Section II presents the proposed Bayesian method as well as its corresponding estimators. The new approach is then validated in the context of EEG signal restoration in Section III. Finally, conclusions and perspectives are drawn in Section IV.

II. UNSUPERVISED SPARSE BAYESIAN REGULARIZATION

A. Problem statement

Let $\mathbf{x} \in \mathbb{R}^{M \times T}$ be our target signal (source amplitudes in EEG/MEG), which is measured by $\mathbf{y} \in \mathbb{R}^{P \times T}$ through a linear observation operator $\mathcal{H} \in \mathbb{R}^{P \times M}$ (lead field matrix), where M , P and T denote the number of dipoles, sources, and time points, respectively. Accounting for the additive acquisition noise \mathbf{n} , the observation model we are interested in can be written as

$$\mathbf{y} = \mathcal{H}\mathbf{x} + \mathbf{n}. \quad (1)$$

Due to all mentioned sources of artifacts (see Section I), the inverse problem in Eq. (1) is ill-posed. We propose to adopt a sparse regularization strategy for estimating the unknown signal \mathbf{x} via a Bayesian framework. Indeed, sparsity promoting priors have been widely used for static regularization problems, but also for dynamic ones where an additional temporal dimension has to be accounted for [5, 8]. Specifically, we propose to extend the Bayesian sparse regularization method recently developed in [9] by considering uncertainties associated with the linear operator \mathcal{H} . Indeed, since EEG signals are sparse in the original domain, using sparsity promoting priors allows fine details and transitions in the target signal to be preserved. In this paper, we use a Bernoulli-Laplace prior for the target signal and a Gaussian prior for the linear operator (lead field matrix). The following sub-section describes the adopted hierarchical Bayesian model.

B. Hierarchical Bayesian model

In a probabilistic setting, \mathbf{x} , \mathbf{y} and \mathcal{H} are assumed to be realizations of random vectors X , Y and H . Our goal will be to characterize the posterior probability distribution of (X, H) given Y , by considering some parametric probabilistic model and by estimating the associated hyperparameters.

1) Likelihood:

Under the assumption of additive iid Gaussian noise of variance σ_n^2 , the likelihood associated with the observation model in Eq. (1) can be expressed as follows

$$f(\mathbf{y}|\mathbf{x}, \mathcal{H}, \sigma_n^2) = \left(\frac{1}{2\pi\sigma_n^2}\right)^{PT/2} \exp\left(-\frac{\|\mathbf{y} - \mathcal{H}\mathbf{x}\|_2^2}{2\sigma_n^2}\right) \quad (2)$$

where $\|\cdot\|_2$ denotes the Euclidean norm.

2) Priors:

In the proposed model, the unknown parameter vector to be estimated will be denoted by $\boldsymbol{\theta} = \{\mathbf{x}, \mathcal{H}, \sigma_n^2\}$. In what follows, we introduce appropriate priors for the unknown model parameters.

Prior for \mathbf{x}

To promote the sparsity of the target image, we propose as in [9, 10] to consider a Bernoulli-Laplace prior for x_i ($i = 1, \dots, MT$), defined by

$$f(x_i|\omega, \lambda) = (1 - \omega)\delta(x_i) + \frac{\omega}{2\lambda} \exp\left(-\frac{|x_i|}{\lambda}\right) \quad (3)$$

where $\delta(\cdot)$ is the Dirac delta function, $\lambda > 0$ is the parameter of the Laplace distribution, and w is a weight belonging to $[0, 1]$. Assuming the coefficients x_i are *a priori* independent, the prior distribution for \mathbf{x} is

$$f(\mathbf{x}|\omega, \lambda) = \prod_{i=1}^{MT} f(x_i|\omega, \lambda). \quad (4)$$

Prior for σ_n^2

To guarantee the positivity of σ_n^2 and keep this prior non-informative, we use here a Jeffrey's prior defined as

$$f(\sigma_n^2) \propto \frac{1}{\sigma_n^2} \mathbf{1}_{\mathbb{R}^+}(\sigma_n^2) \quad (5)$$

where $\mathbf{1}_{\mathbb{R}^+}$ is the indicator function on \mathbb{R}^+ , i.e., $\mathbf{1}_{\mathbb{R}^+}(\xi) = 1$ if $\xi \in \mathbb{R}^+$ and 0 otherwise, and \propto means "proportional to".

Prior for \mathcal{H}

To account for sensitivity map estimation errors, we adopt here a Gaussian prior for the target sensitivity operator \mathcal{H} with diagonal covariance matrix $\sigma_h^2 \mathbf{I}$ and mean $\overline{\mathcal{H}}$

$$\begin{aligned} f(\mathcal{H}|\overline{\mathcal{H}}, \sigma_h^2) &\propto \left(\frac{1}{2\pi\sigma_h^2}\right)^{PM/2} \exp\left(-\frac{\|\mathcal{H} - \overline{\mathcal{H}}\|_2^2}{2\sigma_h^2}\right) \\ &\propto \prod_{l=1}^{PM} \left(\frac{1}{2\pi\sigma_h^2}\right)^{1/2} \exp\left(-\frac{(h_l - \overline{h}_l)^2}{2\sigma_h^2}\right). \end{aligned} \quad (6)$$

This prior allows us to model sensitivity errors as a Gaussian perturbation of \mathcal{H} with respect to some mean $\overline{\mathcal{H}}$. In what follows, $\overline{\mathcal{H}}$ will be assumed to be known while σ_h^2 will be estimated.

3) Hyperpriors:

The hyperparameters λ and ω might be fixed a priori using information about the image sparsity. Since this information is not always easy to obtain, this paper proposes to estimate these hyperparameters from the data using a hierarchical Bayesian approach. This approach requires to define priors for the hyperparameters. As in [9], inverse gamma (denoted as $IG(\lambda|\alpha, \beta)$) and uniform (denoted as $\mathcal{U}_{[0,1]}$) priors have been chosen for the hyperparameters λ and ω .

As regards the linear operator variance σ_h^2 , a Jeffrey's prior has been used to guarantee its positivity while keeping this prior non-informative:

$$f(\sigma_h^2) \propto \frac{1}{\sigma_h^2} \mathbf{1}_{\mathbb{R}^+}(\sigma_h^2). \quad (7)$$

C. Bayesian inference

We adopt here a maximum a posteriori (MAP) strategy in order to estimate the model parameter vector $\boldsymbol{\theta}$ based on the posterior of the unknown parameters and hyperparameters. More precisely, if we denote by $\boldsymbol{\Phi} = \{\lambda, \omega, \sigma_h^2\}$ the model hyperparameters, the joint posterior distribution of $\{\boldsymbol{\theta}, \boldsymbol{\Phi}\}$ can be expressed as

$$f(\boldsymbol{\theta}, \boldsymbol{\Phi}|\mathbf{y}, \alpha, \beta) \propto f(\mathbf{y}|\boldsymbol{\theta})f(\boldsymbol{\theta}|\boldsymbol{\Phi})f(\boldsymbol{\Phi}|\alpha, \beta). \quad (8)$$

The posterior Eq. (8) is not sufficiently simple to derive a closed form expression for the MAP estimator of $\boldsymbol{\theta}$. As a consequence, we propose here to use a Gibbs sampler [7] that iteratively generates samples according to the conditional distributions $f(\mathbf{x}|\mathbf{y}, \omega, \lambda, \mathcal{H}, \sigma_n^2)$, $f(\sigma_n^2|\mathbf{y}, \mathbf{x}, \mathcal{H})$, $f(\mathcal{H}|\sigma_h^2, \mathbf{y}, \mathbf{x})$, $f(\lambda|\mathbf{x}, \alpha, \beta)$, $f(\omega|\mathbf{x})$ and $f(\sigma_h^2|\mathcal{H})$. This sampler is summarized in Algorithm 1.

The conditional distributions $f(\mathbf{x}|\mathbf{y}, \omega, \lambda, \mathcal{H}, \sigma_n^2)$, $f(\sigma_n^2|\mathbf{y}, \mathbf{x}, \mathcal{H})$, $f(\lambda|\mathbf{x}, \alpha, \beta)$ and $f(\omega|\mathbf{x})$ have been derived in [9] where the uncertainties about \mathcal{H} were not considered. As regards $f(\sigma_h^2|\mathcal{H})$, straightforward calculations show that this posterior is the following inverse gamma distribution

$$\sigma_h^2|\mathcal{H} \sim \mathcal{IG}\left(\sigma_h^2|PM/2, \|\mathcal{H} - \overline{\mathcal{H}}\|_2^2\right). \quad (9)$$

Algorithm 1 Gibbs sampler.

Initialize with some $\mathbf{x}^{(0)}$ and $\mathcal{H}^{(0)}$.

repeat

- ① Sample σ_n^2 according to $f(\sigma_n^2|\mathbf{y}, \mathbf{x}, \mathcal{H})$.
- ② Sample λ according to $f(\lambda|\mathbf{x}, \alpha, \beta)$.
- ③ Sample ω according to $f(\omega|\mathbf{x})$.
- ④ Sample \mathbf{x} according to $f(\mathbf{x}|\mathbf{y}, \omega, \lambda, \mathcal{H}, \sigma_n^2)$.
- ⑤ Sample σ_h^2 according to $f(\sigma_h^2|\mathcal{H})$.
- ⑥ Sample \mathcal{H} according to $f(\mathcal{H}|\sigma_n^2, \mathbf{y}, \mathbf{x})$.

until convergence

Concerning $f(\mathcal{H}|\sigma_n^2, \mathbf{y}, \mathbf{x})$, after decomposing \mathcal{H} on the orthonormal basis $B = \{U_1, \dots, U_{MP}\}$ such that $\mathcal{H} = \tilde{\mathcal{H}}_{-l} + h_l U_l$ where $\tilde{\mathcal{H}}_{-l}$ is nothing but \mathcal{H} whose l^{th} element has been set to 0, straightforward calculations yield the following conditional distribution for h_l :

$$h_l|\mathbf{x}, \mathbf{y}, \sigma_h^2, \sigma_n^2 \sim \mathcal{N}(\mu_l, \sigma_l^2) \quad (10)$$

where $\sigma_l^2 = \frac{\sigma_n^2 \sigma_h^2}{\sigma_n^2 + \sigma_h^2 \|U_l \mathbf{x}\|_2^2}$, $\mu_l = \sigma_l^2 (\mathbf{e}_l^T U_l \mathbf{x} + \bar{h}_l)$ and $\mathbf{e}_l = \mathbf{y} - \tilde{\mathcal{H}}_{-l} \mathbf{x}$.

After convergence, the proposed algorithm ends up with sampled sequences that are used to compute a minimum mean square error (MMSE) or MAP estimator of the unknown parameter vector, allowing us to compute the estimated signal $\hat{\mathbf{x}}$ and operator $\hat{\mathcal{H}}$, in addition to the estimated noise variance $\hat{\sigma}_n^2$ and the estimated hyperparameters $\hat{\lambda}$, $\hat{\omega}$ and $\hat{\sigma}_h^2$.

III. APPLICATION TO EEG SIGNAL RECOVERY

In this experiment, we consider an EEG reconstruction problem where the observed signals correspond to the activity measured by each electrode during the acquisition time. We used the MNE software¹, to simulate an EEG dataset using 20 electrodes with 35 sources (35 voxels on the brain surface), where only 3 of them have been chosen to be active. The simulation involved 21 time points, which means that we have to recover a 35×21 image \mathbf{x} (see Fig. 1[Ground truth]) from an observation \mathbf{y} of size 20×21 , where each line of \mathbf{x} represents the activity at a given voxel (source) of the brain surface. The linear operator \mathcal{H} is supposed to represent the brain model geometry. The simulation has been performed by taking a Gaussian acquisition noise of variance $\sigma_n^2 = 0.16$. To simulate uncertainties in the observation operator, a zero-mean Gaussian noise of variance $\sigma_h^2 = 0.0064$ has been added to the reference lead field matrix $\bar{\mathcal{H}}$.

Fig. 1 displays the ground truth and reconstructed signals using a standard weighted least squares method (WLS), the proposed method and the Bernoulli-Laplace regularization ($\ell_0 + \ell_1$) of [9]. Visual inspection of reconstructed images clearly show the reconstruction errors caused by the linear operator imprecisions for WLS and the $\ell_0 + \ell_1$

regularization. These errors are clearly reduced by the proposed method by recovering a better sparsity support and accurate signal. Table I confirms this observation by reporting quantitative evaluation of the sparsity level for the reconstructed signals. In fact, when evaluating the ℓ_0 pseudo-norm of reconstructed signals ($\|\cdot\|_0$), the proposed method gives the closest sparsity level ($\|\hat{\mathbf{x}}\|_0 = 31$) to the ground truth ($\|\mathbf{x}_{\text{ref}}\|_0 = 51$). As regards quantitative evaluation of reconstruction quality, Table I also reports signal to noise ratios (SNRs)². Reported values clearly show the ability of the proposed method (SNR = 13.8 dB) to recover more accurate signals than the $\ell_0 + \ell_1$ regularization (SNR = 1.34 dB), which is due to the adopted model that accounts for imprecisions in the observation operator. This SNR gain confirms the importance of accounting for such errors.

TABLE I
SNR AND SPARSITY LEVEL OF RECONSTRUCTED SIGNALS.

| | Ground truth | WLS | $\ell_0 + \ell_1$ | Proposed method |
|---------------|--------------|------|-------------------|-----------------|
| SNR (dB) | - | 0.41 | 1.34 | 13.80 |
| $\ \cdot\ _0$ | 51 | 595 | 165 | 31 |

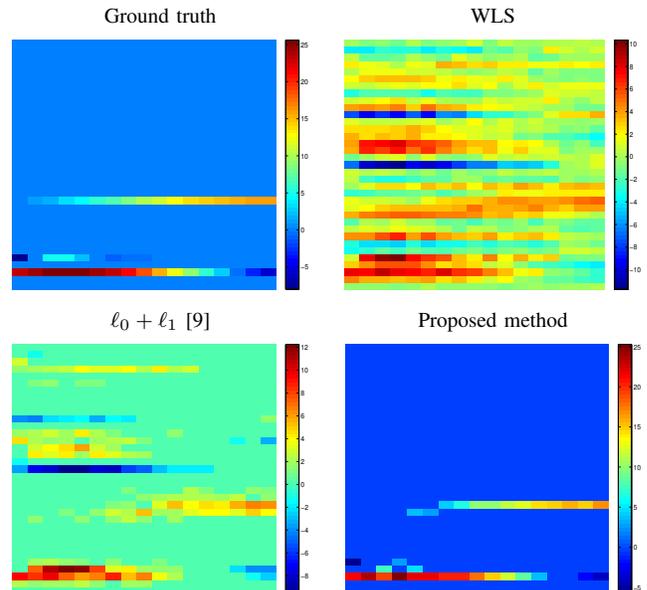


Fig. 1. Original and restored EEG signals using WLS, $\ell_0 + \ell_1$ regularization and the proposed method.

As regards the linear operator restoration, Fig. 2 shows the true observation operator used for simulation, as well as the noisy and restored ones. This figure also reports SNR values which show that our method allows getting accurate estimates of the observation operator. This accurate estimation explains the obtained SNR gain in the estimation of the signal $\hat{\mathbf{x}}$.

When looking at the model hyperparameters, the sampled chains allowed us to derive the following estimates:

¹<http://www.martinos.org/mne/>

² $\text{SNR} = 20 \log_{10} \frac{\|\mathbf{x}_{\text{ref}}\|_2}{\|\mathbf{x}_{\text{ref}} - \hat{\mathbf{x}}\|_2}$

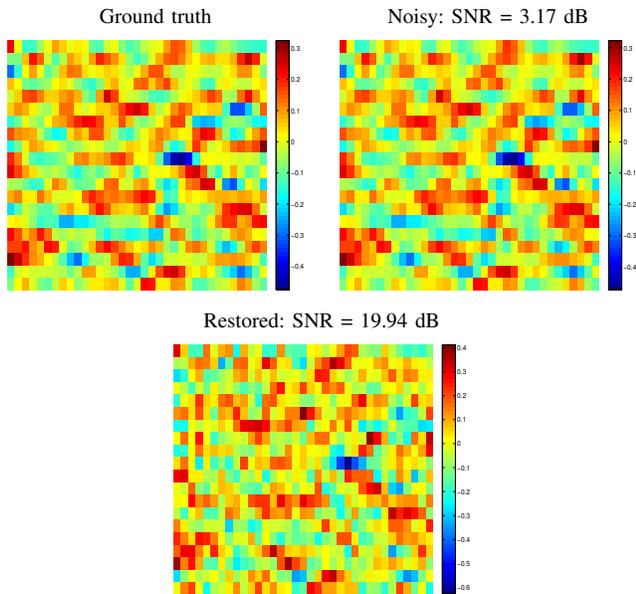


Fig. 2. Ground truth, noisy and restored observation operator using the proposed method. SNR values are also provided.

$\hat{\sigma}_n^2 = 0.13$ and $\hat{\sigma}_h^2 = 0.0065$, which are good estimates of the reference values σ_n^2 and σ_h^2 . Concerning the Bernoulli parameter, the estimated value $\hat{\omega} = 0.1265$ indicates indeed a high sparsity level of the signal \hat{x} .

As regards computational cost, 1000 iterations are required for the proposed method including 500 iterations corresponding to the burn-in period (have been discarded from the estimation). With a Matlab implementation on a 64-bit 2.00GHz i7-3667U architecture, the proposed method takes 176 seconds, while the $\ell_0 + \ell_1$ regularization takes only 165 seconds. The proposed method presents therefore a good compromise in terms of performance/cost ratio. Note that this computational cost can be highly decreased by resorting to C implementation for example.

IV. CONCLUSIONS

In this contribution, we proposed a novel method for sparse signal/image regularization which, in addition to standard signal imperfections, models imprecisions about the observation operator. The proposed method was developed in a flexible Bayesian framework where the regularization parameters were automatically estimated from the data. In application to EEG/MEG signal restoration, promising results show the potential of the proposed method in handling such complicated restoration problems. Future work will investigate the application of this method on real EEG/MEG signals where experimental conditions could be more severe and not fully controlled.

REFERENCES

- [1] L. Ke and J. Zhang, "Dynamic-parallel MR image reconstruction based on adaptive coil sensitivity estimation," in *Int. Symp. on Biomed. Imag.*, Paris, France, May 14-17 2008, pp. 1015 – 1018.
- [2] D. Weinstein, L. Zhukov, and C. Johnson, "Lead-field bases for electroencephalography source imaging," *Annals of Biomedical Engineering*, vol. 28, no. 9, pp. 1059–1065, Sep. 2000.

- [3] W. Ou, M. Hamalainen, and P. Golland, "A distributed spatio-temporal EEG/MEG inverse solver," *NeuroImage*, vol. 44, no. 3, pp. 932–946, 2009.
- [4] K. Friston, L. Harrison, J. Daunizeau, S. Kiebel, C. Phillips, N. Trujillo-Barreto, R. Henson, G. Flandin, and J. Mattout, "Multiple sparse priors for the M/EEG inverse problem," *NeuroImage*, vol. 39, no. 3, pp. 1104–1120, 2008.
- [5] A. Gramfort, D. Strohmeier, J. Hauelsen, M.S. Hmlinen, and M. Kowalski, "Time-frequency mixed-norm estimates: Sparse M/EEG imaging with non-stationary source activations," *NeuroImage*, vol. 70, pp. 410–422, 2013.
- [6] L. Bouchchir and B. Boashash, "Wavelet denoising based on the MAP estimation using the BKF prior with application to images and EEG signals," *IEEE Trans. on Signal Process.*, vol. 61, no. 8, pp. 1880–1894, Apr. 2013.
- [7] C. Robert and G. Casella, *Monte Carlo statistical methods*, Springer, New York, 2004.
- [8] L. Chaari, S. Mériaux, J.-Ch. Pesquet, and P. Ciuciu, "Spatio-temporal wavelet regularization for parallel MRI reconstruction: application to functional MRI," 2013, arxiv:1201.0022.
- [9] L. Chaari, J.-Y. Tournet, and H. Batatia, "Sparse Bayesian regularization using Bernoulli-Laplacian priors," in *Proc. European Signal Processing Conference (EUSIPCO)*, Marrakech, Morocco, September, 9-13 2013.
- [10] N. Dobigeon, A. O. Hero, and J.-Y. Tournet, "Hierarchical Bayesian sparse image reconstruction with application to MRFM," *IEEE Trans. Image Process.*, vol. 18, no. 9, pp. 2059–2070, Sept. 2009.