HYPERPARAMETERS ESTIMATION FOR THE BAYESIAN LOCALIZATION OF THE EEG SOURCES WITH TV PRIORS

Antonio López[1], *Jesús M. Cortés*[2,3], *Domingo López-Oller*[4], *Rafael Molina*[2] *and Aggelos K. Katsaggelos*[5]

[1]Departamento de Lenguajes y Sistemas Informáticos, Universidad de Granada, 18071 Granada, Spain, alopez@ugr.es
[2]Departamento de Ciencias de la Computación e I.A., Universidad de Granada, 18071 Granada, Spain, jcortes,rms@decsai.ugr.es
[3]Ikerbasque: The Basque Foundation for Science. Biocruces Health Research Institute. 48903. Barakaldo, Bizkaia. Pais Vasco, Spain
[4]Departamento de Teoría de la Señal, Telemática y Comunicaciones. Universidad de Granada, 18071 Granada, Spain, domingolopez@ugr.es
[5]Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, Illinois 60208-3118, aggk@eecs.northwestern.edu

ABSTRACT

In this work we propose a new Bayesian method for the non-invasive localization of EEG sources. For this problem, most of the existing methods assume that the sources are distributed throughout the brain volume according to smooth 3D patterns. However, this assumption might fail in pathological conditions, such as in an epileptic brain, where it can occur that the neurophysiological generators are localized in a narrow region, highly compacted, what originates abrupt profiles of electrical activity. This new method incorporates a Total Variation (TV) prior which has been used before in image processing for edge detection and applies variational methods to approximate the probability distributions to estimate the unknown parameters and the sources. The procedure is tested and validated on synthetic EEG data.

Index Terms— EEG Source Localization, Bayesian Inference, TV Prior, Variational Methods, Hyperparameters Estimation

1. INTRODUCTION

The problem of the EEG source localization has a long tradition in computational neuroimaging [1, 2, 3]. It consists of providing optimal solutions to the localization of the EEG sources within the skull based on EEG measurements on the scalp (observations). Thus, it gives a non-invasive localization of the electrical sources. Despite its relevance, this problem is intrinsically ill-posed as infinite solutions are consistent with the same scalp observations. To overcome this problem, many methods have been proposed (for reviews see [1, 2, 3] and references therein) and many of these methods have used Bayesian estimation. However, previous prior models assumed an L_2-norm for the gradients of dipole intensity, such as the widely used method LORETA (Low Resolution Electromagnetic Tomography) [4], Minimum Norm Estimation (MNE) or Restricted Maximum Likelihood (ReML) [5].

For situations in which sources are following abrupt patterns of electrical activity, as for instance, it occurs in focalized epilepsy (in which it is possible even to have a strong localization in a single dipole), L_2-norm algorithms will underestimate the localization as the strong activity of a single-dipole is thermalized by its neighborhood.

We need methods for the localization of the focal epileptic activity, resulting in accurate determination of the seizure onset for helping presurgical evaluation. In this paper we present a new method applicable to these situations and compare its performance with the well-known LORETA algorithm [4], which is the control for our method.

The paper is organized as follows. In section 2 we briefly present the modeling. In section 3 we introduce the Bayesian formulation of the problem. In section 4 we perform the Bayesian inference analysis and, as a result, we give a detailed algorithm for EEG source localization. In section 5 we apply such algorithm to synthetic data and validate its efficiency. Finally, conclusions and discussion are given in Section 6.

2. MODEL

We will assume that there are n_e scalp electrodes on the scalp (observations at sensors) measuring the electrical activity originated by $n_d \gg n_e$ dipoles (sources), equally distributed throughout the brain volume (a 3D grid within a skull).

A standard formulation for the relationship between ob-

Work supported by MICINN, Ministerio de Ciencia e Innovacion, Ref. TIN2010-15137.

servations and sources is given by the equation

$$\mathbf{v} = \mathcal{L}\mathbf{j} + \epsilon, \tag{1}$$

where $\mathbf{v}$ is a vector of size $n_e \times 1$ representing the scalp electrical potentials measured by sensors, $\mathbf{j}$ is an $n_d \times 1$ vector accounting for the source magnitude and ϵ is modeling additive noise. The proportionality matrix between observations and sources is given by the Lead Field matrix $\mathcal{L}$, which has dimensions of $n_e \times n_d$. To calculate it one needs the positions of both sensors and sources and knowledge on the specific head model (for details on our implementation see results).

3. BAYESIAN FORMULATION

The Bayesian formulation of the EEG source localization problem requires the definition of the joint distribution $\mathrm{p}(\alpha, \beta, \mathbf{j}, \mathbf{v})$ of the observation $\mathbf{v}$, the unknown original source $\mathbf{j}$ and the hyperparameters β and α. To model this joint distribution we utilize the hierarchical Bayesian paradigm in which the estimation is performed in two stages. In the first stage, we form the distributions $\mathrm{p}(\mathbf{v}|\mathbf{j}, \beta)$ and $p(\mathbf{j}|\alpha)$, and in the second stage the hyperpriors for the hyperparameters β and α are defined. For that purpose, the joint posterior distribution is written as

$$\mathrm{p}(\alpha, \beta, \mathbf{j}, \mathbf{v}) = \mathrm{p}(\alpha)\mathrm{p}(\beta)\mathrm{p}(\mathbf{j} \mid \alpha)\mathrm{p}(\mathbf{v} \mid \mathbf{j}, \beta). \tag{2}$$

In our case, noise realizations at the different sensors are considered independent from each other. For each sensor, the noise model in Eq. (1) is assumed to be Gaussian with mean zero and variance equal to β^{-1}. Thus, the distribution of the observations $\mathbf{v}$ given $\mathbf{j}$ and β is given by

$$\mathrm{p}(\mathbf{v}|\mathbf{j}, \beta) \propto \beta^{\frac{n_e}{2}} \exp\left[-\frac{\beta}{2} \parallel \mathbf{v} - \mathcal{L}\mathbf{j} \parallel^2\right]. \tag{3}$$

For modeling the sources we use a TV prior given by

$$\mathrm{p}(\mathbf{j}|\alpha) \propto \alpha^{\frac{n_d}{2}} \exp\left[-\alpha TV(\mathbf{j})\right], \tag{4}$$

which has been computed from a partition function with an empirical validation [6]. The hyperparameter α (the *scale* parameter) is weighting the effect that the prior distribution has in the source localization versus the contribution coming purely from the observations; the higher the α, the bigger the prior contribution.

For a three-dimensional grid of dipole locations, we write

$$TV(\mathbf{j}) = \sum_{i}^{n_d} \sqrt{(\Delta_i^x \mathbf{j})^2 + (\Delta_i^y \mathbf{j})^2 + (\Delta_i^z \mathbf{j})^2}, \tag{5}$$

where $\Delta_i^x \mathbf{j}$, $\Delta_i^y \mathbf{j}$ and $\Delta_i^z \mathbf{j}$ are the first order differences at dipole i with respect to space directions x, y, z, respectively. More specifically, they are defined as $\Delta_i^x \mathbf{j} \equiv \mathbf{j}_i - \mathbf{j}_{x(i)}$, $\Delta_i^y \mathbf{j} \equiv \mathbf{j}_i - \mathbf{j}_{y(i)}$ and $\Delta_i^z \mathbf{j} \equiv \mathbf{j}_i - \mathbf{j}_{z(i)}$, where $x(i)$, $y(i)$ and $z(i)$ denote the nearest neighbors of i, to the right, front and below, respectively (see Fig. 1).

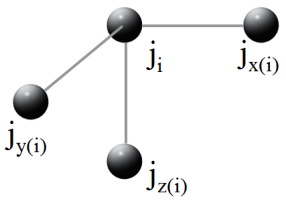

Fig. 1. **Sketch illustrating the three-dimensional directions for the first order differences for a dipole i.**

At the second stage of the hierarchical Bayesian paradigm, to model the vector of hyperparameters $\omega \in (\alpha, \beta)$ we use gamma hyperpriors defined by

$$\mathrm{p}(\omega) = \Gamma(\omega|a_\omega^o, b_\omega^o) \propto \omega^{a_\omega^o - 1} \exp\left[-b_\omega^o \omega\right], \tag{6}$$

where $a_\omega^o > 0$ and $b_\omega^o > 0$ are, respectively, the scale and shape parameters of the gamma distribution, which are assumed to be known.

4. BAYESIAN INFERENCE

To perform inference we need to calculate the posterior distribution $\mathrm{p}(\alpha, \beta, \mathbf{j}|\mathbf{v})$. Since this posterior can not be found in a closed form, we will use variational methods to approximate the posterior to the distribution $\mathrm{q}(\alpha, \beta, \mathbf{j}) = \mathrm{q}(\alpha, \beta)\mathrm{q}(\mathbf{j})$. In particular, the variational criterion we use is to find $\mathrm{q}(\alpha, \beta, \mathbf{j})$ after minimization of the Kullback-Leibler (KL) divergence, i.e.

$$\begin{aligned} &C_{KL}(\mathrm{q}(\alpha, \beta, \mathbf{j}) \parallel \mathrm{p}(\alpha, \beta, \mathbf{j} \mid \mathbf{v})) = \\ &= \int_{\alpha,\beta,\mathbf{j}} \mathrm{q}(\alpha, \beta, \mathbf{j}) \log\left(\frac{\mathrm{q}(\alpha, \beta, \mathbf{j})}{\mathrm{p}(\alpha, \beta, \mathbf{j} \mid \mathbf{v})}\right) d\alpha d\beta d\mathbf{j} \\ &= \int_{\alpha,\beta,\mathbf{j}} \mathrm{q}(\alpha, \beta, \mathbf{j}) \log\left(\frac{\mathrm{q}(\alpha, \beta, \mathbf{j})}{\mathrm{p}(\alpha, \beta, \mathbf{j}, \mathbf{v})}\right) d\alpha d\beta d\mathbf{j} \\ &\quad + \log \mathrm{p}(\mathbf{v}), \end{aligned} \tag{7}$$

which is always non negative and equal to zero if and only if, $\mathrm{q}(\alpha, \beta, \mathbf{j}) = \mathrm{p}(\alpha, \beta, \mathbf{j} \mid \mathbf{v})$.

The form of the TV prior (5) makes the evaluation of the KL distance difficult, but following [7] we will make use of an approximation which allows to obtain a quadratic lower bound for the prior model, ie.

$$p(\alpha|\mathbf{j}) \geq cM(\alpha, \mathbf{j}, \mathbf{u}), \tag{8}$$

where c is a constant, $\mathbf{u} \in (R^+)^{n_d}$ is a vector which depends on the spatial first-order differences of the sources $\mathbf{j}$ under the

distribution $\mathrm{q}(\mathbf{j})$ and $M(\alpha, \mathbf{j})$ is given by

$$M(\alpha, \mathbf{j}, \mathbf{u}) = \alpha^{n_d/2} \times \exp\left[-\frac{\alpha}{2}\sum_{i}^{n_d}\frac{(\Delta_i^x \mathbf{j})^2 + (\Delta_i^y \mathbf{j})^2 + (\Delta_i^z \mathbf{j})^2 + u_i}{\sqrt{u_i}}\right]. \quad (9)$$

Consequently, Eq. (8) allows for a lower bound for the joint probability distribution, i.e.

$$\begin{aligned} p(\alpha, \beta, \mathbf{j}, \mathbf{v}) &\geq c\mathrm{p}(\alpha)\mathrm{p}(\beta)M(\alpha, \mathbf{j}, \mathbf{u})\mathrm{p}(\mathbf{v} \mid \mathbf{j}, \beta) \\ &\equiv F(\alpha, \beta, \mathbf{j}, \mathbf{u}, \mathbf{v}), \end{aligned} \quad (10)$$

which by substituting it in Eq. (7), it provides an upper bound for the KL distance

$$\begin{aligned} &C_{KL}(\mathrm{q}(\alpha, \beta, \mathbf{j}) \parallel \mathrm{p}(\alpha, \beta, \mathbf{j} \mid \mathbf{v})) = \\ &\leq \min_{\mathbf{u}} \int_{\alpha,\beta,\mathbf{j}} \mathrm{q}(\alpha, \beta, \mathbf{j}) \log\left(\frac{\mathrm{q}(\alpha, \beta, \mathbf{j})}{F(\alpha, \beta, \mathbf{j}, \mathbf{u}, \mathbf{v})}\right) d\alpha d\beta d\mathbf{j}. \end{aligned} \quad (11)$$

The minimization of Eq. (11) leads to the following iterative procedure for finding $q(\alpha, \beta, \mathbf{j})$ and $\mathbf{u}$:

Algorithm

1. Give the initial estimates for $\mathbf{u}$ *and* $q(\alpha, \beta)$*, named respectively* $\mathbf{u^1}$ *and* q^1.

2. Do the following until algorithm convergence:

2.1. Find the solution of

$$\begin{aligned} q^k(\mathbf{j}) &= \arg\min_{q(\mathbf{j})} \Big(\int_{\mathbf{j}} \int_{\alpha} \int_{\beta} q^k(\alpha, \beta) q(\mathbf{j}) \\ &\quad \times \log\left(\frac{q^k(\alpha, \beta) q(\mathbf{j})}{F(\alpha, \beta, \mathbf{j}, \mathbf{u}^k, \mathbf{v})}\right) d\alpha d\beta d\mathbf{j}\Big), \end{aligned} \quad (12)$$

which is given by

$$q^k(\mathbf{j}) \propto \exp\{E_{q^k(\alpha,\beta)}\left[\ln F(\alpha, \beta, \mathbf{j}, \mathbf{u}^k)\right]\}. \quad (13)$$

At the k *iteration the estimation of the distribution of* $\mathbf{j}$ *is Gaussian with mean given by*

$$E_{q^k(\mathbf{j})}[\mathbf{j}] = cov_{q^k(\mathbf{j})}[\mathbf{j}] E_{q^k(\beta)}[\beta] \mathcal{L}^T \mathbf{v}, \quad (14)$$

and covariance

$$cov_{q^k(\mathbf{j})}[\mathbf{j}] = \left(E_{q^k(\beta)}[\beta]\mathcal{L}^T\mathcal{L} + E_{q^k(\alpha)}[\alpha] D^T Z(\mathbf{u}^k) D\right)^{-1} \quad (15)$$

2.2. Find the solution of

$$\begin{aligned} \mathbf{u}^{k+1} &= \arg\min_{\mathbf{u}} \Big(\int_{\mathbf{j}} \int_{\alpha} \int_{\beta} q^k(\alpha, \beta) q^k(\mathbf{j}) \\ &\quad \times \log\left(\frac{q^k(\alpha, \beta) q^k(\mathbf{j})}{F(\alpha, \beta, \mathbf{j}, \mathbf{u}, \mathbf{v})}\right) d\alpha d\beta d\mathbf{j}\Big), \end{aligned} \quad (16)$$

which is given by

$$u_i^{k+1} = \left(\Delta_i^x E_{q^k(\mathbf{j})}[\mathbf{j}]\right)^2 + \left(\Delta_i^y E_{q^k(\mathbf{j})}[j]\right)^2 + \left(\Delta_i^z E_{q^k(\mathbf{j})}[\mathbf{j}]\right)^2 + \left[cov_{q^k(\mathbf{j})}[\mathbf{j}] D^T D\right]_{ii} \text{ with } i = 1, \ldots, n_d. \quad (17)$$

2.3. Find the solution of

$$\begin{aligned} q^{k+1}(\alpha, \beta) &= \arg\min_{q(\alpha,\beta)} \Big(\int_{\mathbf{j}} \int_{\alpha} \int_{\beta} q(\alpha, \beta) q^k(\mathbf{j}) \\ &\quad \times \log\left(\frac{q(\alpha, \beta) q^k(\mathbf{j})}{F(\alpha, \beta, \mathbf{j}, \mathbf{u}^{k+1}, \mathbf{v})}\right) d\alpha d\beta d\mathbf{j}\Big), \end{aligned} \quad (18)$$

which is given by

$$\begin{aligned} q^{k+1}(\alpha, \beta) &= q^{k+1}(\alpha) q^{k+1}(\beta) \\ &\propto \exp\{E_{q^k(\mathbf{j}}\left[\ln F(\alpha, \beta, \mathbf{j}, \mathbf{u}^{k+1})\right]\}. \end{aligned} \quad (19)$$

Therefore, $q^{k+1}(\alpha)$ *and* $q^{k+1}(\beta)$ *are Gamma distributions. Their expectations are the estimates of the parameters* α *and* β*; the inverse expectations are given by:*

$$\left(E_{q^{k+1}(\alpha)}[\alpha]\right)^{-1} = \gamma_\alpha \frac{1}{\bar{\alpha}^o} + (1 - \gamma_\alpha) \frac{2\sum_i^{n_d} \sqrt{u_i^{k+1}}}{n_d}, \quad (20)$$

and

$$\left(E_{q^{k+1}(\beta)}[\beta]\right)^{-1} = \gamma_\beta \frac{1}{\bar{\beta}^o} + (1 - \gamma_\beta) \frac{E_{q^k(\mathbf{j})}\left[\| \mathbf{v} - \mathcal{L}\mathbf{j} \|^2\right]}{n_e}. \quad (21)$$

Several remarks to the algorithm:

1. In Eq. (13), $E_q[f(X)]$ is denoting the expected value of $f(X)$ over the distribution q.
2. In Eq. (15) we have defined $D \equiv \left[(D^x)^T (D^y)^T (D^z)^T\right]^T$ a 3×1 block matrix in which the blocks D^x, D^y, and D^z denote convolution matrices $n_d \times n_d$, such that when applied to $\mathbf{j}$ returns the first order differences with respect to directions x, y, z.
3. The matrix $Z(\mathbf{u}^k)$ in Eq. (15) is a 3×3 block diagonal matrix defined as $Z(\mathbf{u}^k) \equiv diag\left[B(\mathbf{u}^k), B(\mathbf{u}^k), B(\mathbf{u}^k)\right]$ where $B(\mathbf{u}^k)$ is a $n_d \times n_d$ diagonal matrix of the form
$$B(\mathbf{u}^k) = diag\left(\frac{1}{\sqrt{u_i^k}}\right) i = 1, \ldots, n_d.$$
4. In Eqs. (20) and (21) we have defined the following constants: $\bar{\alpha}^o \equiv a_\alpha^o / b_\alpha^o$, $\bar{\beta}^o \equiv a_\beta^o / b_\beta^o$, $\gamma_\alpha \equiv \frac{a_\alpha^o}{a_\alpha^o + n_d/2}$, $\gamma_\beta \equiv \frac{a_\beta^o}{a_\beta^o + n_e/2}$.
5. The expression $E_{q^k(\mathbf{j})}\left[\| \mathbf{v} - \mathcal{L}\mathbf{j} \|\right]^2$ in Eq. (21) is obtained by
$$\begin{aligned} E_{q^k(\mathbf{j})}\left[\| \mathbf{v} - \mathcal{L}\mathbf{j} \|\right]^2 &= \| \mathbf{v} - \mathcal{L} E_{q^k(\mathbf{j})}[\mathbf{j}] \|^2 + \\ &\quad trace\left[cov_{q^k(\mathbf{j})}[\mathbf{j}] \mathcal{L}^T \mathcal{L}\right]. \end{aligned}$$

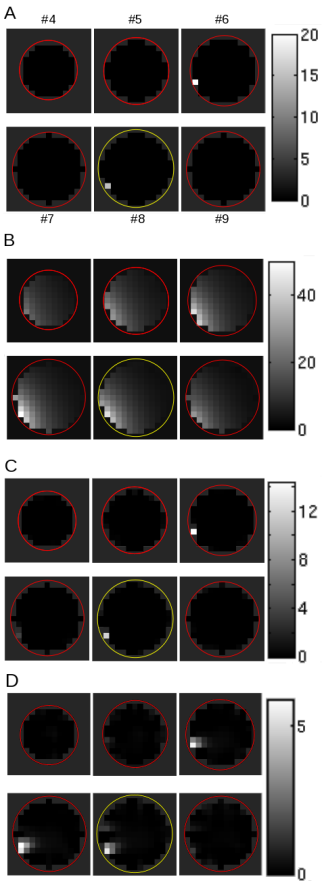

Fig. 2. **Visualizarion of the algorithm performance for the source localization**. A: Original sources. B: Sources at first iteration. C: Final localization performed by our algorithm. D: Loreta localization. A-D: Depicted 6 different slices, enumerated from 4 to 9 in panel A. Red circles are the cross-sections of the sphere containing the sources. The circle of the central slice, number 8, is colored in yellow. Noise here is 20 dB.

6. To compute the expectations given by Eqs. (20) and (21) we have made use of the Gamma distributions given by

$$q^{k+1}(\alpha) \propto \alpha^{n_d/2+a_\alpha^o-1} \times \exp\left[-\alpha\left(\sum_i^{n_d}\sqrt{u_i^{k+1}}+b_\alpha^o\right)\right],$$

$$q^{k+1}(\beta) \propto \beta^{n_e/2+a_\beta^o-1} \times \exp\left[-\beta\left(\frac{E_{q^k(\mathbf{j})}\,\|\,\mathbf{v}-\mathcal{L}\mathbf{j}\,\|^2}{2}+b_\beta^o\right)\right].$$

7. Both γ_α and γ_β take values in [0,1), so they can be interpreted as normalized confidence parameters. When they are zero, the initial values of the hyperparameters are not meaningful at all, and at the other limit, when they approximate to 1, there is no estimation for the hyperparameters as they always coincide with the initial values in the iterative process.
8. As the sensitivity strength for the electrodes decreases proportionally $1/r^2$ [8], with r the source-sensor distance, the situation of deep sources requires the incorporation of new *weighting* corrections, but in this paper they are not needed as we only have considered superficial sources.

5. EXPERIMENTAL RESULTS

In this section, we apply our algorithm for source localization to EEG synthetic data and compare its performance with the classical LORETA [4].

The sources are included in a discrete cubic-grid consisting of 15 transversal slices each one containing 15×15 points and with an inter-slice distance of $d=0.133$. We use a head model of three-concentric spheres [9], in which all sources are within a sphere with radius 0.815, thus containing $n_d=949$ different sources. There are $n_e=128$ EEG electrodes (sensors) at a distance of radius 1 and with distribution given by the Geodesic Sensor Net.

Note that for clinical applications the calculation of the $\mathcal{L}$ matrix assumes a more realistic geometry (based on the MRI volume) but the application of our algorithm is still valid as $\mathcal{L}$ is considered an input parameter; so given another $\mathcal{L}$, the same algorithm can be applied.

The initial condition for the source activity is given by $\mathbf{j}^1=\mathcal{L}^T\mathbf{v}$ (cf. Eq. (1)) and from here we will consider for the rest of the variables that their initial condition only depend on $\mathbf{j}^1$. Thus, we take, $u_i^1=\left(\Delta_i^x\mathbf{j}^1\right)^2+\left(\Delta_i^y\mathbf{j}^1\right)^2+\left(\Delta_i^z\mathbf{j}^1\right)^2$ (cf. Eq. (17)); $\bar{\alpha}^o=n_d\left(2\sum_i\sqrt{u_i^1}\right)^{-1}$ (cf. Eq. (20) for $\gamma_\alpha=0$); $\alpha^1=\bar{\alpha}^o$; $\bar{\beta}^o=n_e\,\|\,\mathbf{v}-\mathcal{L}\mathbf{j}^1\,\|^{-2}$ (cf. Eq. (21) for $\gamma_\beta=0$); $\beta^1=\bar{\beta}^o$. And other initial conditions were considered but our results did not change substancially.

Without loss of generality, we use $\gamma_\alpha=0.5$ and $\gamma_\beta=0.5$, so it exists a balance between the hyperparameters estimation achieved in each iteration and their initial values.

We have simulated three different levels of degradation noise at 40, 30, and 20 dB.

Fig. 2 illustrates the performance of our algorithm. We are depicting the sources at only transversal slices, from the total number of 15, the represented slices have numbers from 4 to 9, Fig. 2A. The circle of the central slice, number 8, is colored in yellow for all figures. Two point sources at the superficial sphere are originally created (cf. Fig. 2A). Note that the two sources, although close one to each other, are separated by one slice what makes their localization to be difficult. At the first iteration, sources are plotted in Fig. 2B. After applying our algorithm, the source estimate is illustrated in Fig.

2C, which is in agreement with Fig. 2A. For comparison purposes, in Fig. 2D we also plot the localization performed by the algorithm LORETA [4] for the value of the regularization parameter chosen with the best MSE. Not only the LORETA localization is biased but the sources strength is strongly underestimated (about one third of the original strength). Thus, our method performs better than LORETA.

After algorithm convergence, the final estimation by our algorithm for 30 dB (not illustrated in Fig. 2) gave a value of MSE= 32, obtained for $\hat{\alpha} = 11$ and $\hat{\beta} = 38200$. For 20 dB, the performance was MSE= 113 (shown in Fig. 2C and obtained for $\hat{\alpha} = 0.6$ and $\hat{\beta} = 53$), whilst the best performance achieved by LORETA was MSE= 513 (Fig. 2D).

For a quantitative evaluation of the hyperparameters estimates, we have computed the MSE curve for different values of the ratio α/β. The results are represented in Fig. 3A (20 dB) and Fig. 3B (30 dB). The MSE value corresponding to the estimated solution $(\hat{\alpha}/\hat{\beta})$ is marked with an arrow, thus showing that the estimation performed by our algorithm is near the minimum of the MSE curves.

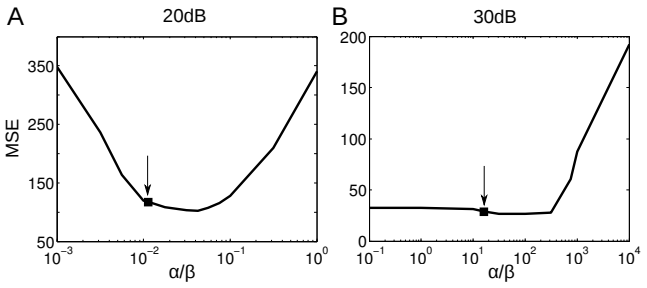

Fig. 3. **Quantitative validation of hyperparameters estimation**. The MSE between the original localization and the final estimation was computed for different values of the ration α/β. A: 20 dB noise. B: 30 dB. A-B: Arrows are indicating the MSE for the estimated solution $(\hat{\alpha}/\hat{\beta})$.

The situation for 40dB noise is not shown in any of the figures, but in this case our Algorithm performs with a negligible Mean Squared Error (MSE= 3.6).

6. CONCLUSIONS

We have presented a new algorithm for EEG source localization based on Bayesian inference with a Total-Variation (TV) prior. Such priors can be very useful for the localization of abrupt sources, as they can occur during seizures. Importantly, our algorithm does not require hand-tuning of the hyperparameters as they have been estimated as well by using the hierarchical Bayesian paradigm and variational approximation. We have applied our algorithm to synthetic data and quantitatively validated the accuracy of the hyperparameters estimation. Interesting to remark, the simulated scenario consisted in two point sources (one slice separation from each other). The better source localization employed by our method compared to the classical LORETA might suggest that our method might work well for detection/localization of multiple (sparse) sources. Future work will include the comparison of our method to other existing methods workable as well for sparse sources, eg. [10].

7. REFERENCES

[1] S. Baillet, J. C. Mosher, and R. M. Leahy, "Electromagnetic brain mapping," *IEEE Signal Processing Magazine*, vol. 18, pp. 14–30, 2001.

[2] C. M. Michel, M. M. Murray, G. Lantz, S. Gonzalez, L. Spinelli, and R. Grave de Peralta, "EEG source imaging," *Journal of Clinical Neurophysiology* , vol. 115, pp. 2195–2222, 2004.

[3] J. M. Schoffelen and J. Gross, "Source connectivity analysis with MEG and EEG," *Human Brain Mapping*, vol. 2009, pp. 1857–1865, 2009.

[4] R. D. Pascual-Marqui, C. M. Michel, and D. Lehmann, "Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain," *International Journal of Psychophysiology*, vol. 18, pp. 49–65, 1994.

[5] C. Phillips, J. Mattout, M. D. Rugg, P. Maquet, and K. J. Friston, "An empirical Bayesian solution to the source reconstruction problem in EEG," *NeuroImage*, vol. 24, pp. 997–1011, 2005.

[6] J. M. Bioucas-Dias, M. A. Figueiredo, and J. P. Oliveira, "Adaptive total variation image deconvolution: A majorization-minimization approach," Presented at the *EUSIPCO*, Florence, Italy, Sep. 2006.

[7] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Parameter estimation in TV image restoration using variational distribution approximation," *IEEE Transactions on Image Processing*, vol. 17, pp. 326–339, 2008.

[8] F. H. Lin, T. Witzel, S. P. Ahlfors, S. M. Stufflebeam, J. W. Belliveau, and M. S. Hamalainen, "Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates," *NeuroImage*, vol. 31, pp. 160–171, 2006.

[9] J. P. Ary, S. A. Klein, and D. H. Feder, "Location of sources of evoked scalp potentials: corrections for skull and scalp thickness," *IEEE Transactions on Biomedical Engineering*, vol. 28, pp. 447–452, 1981.

[10] I. Gorodnitsky, B. D. Rao "Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm," *IEEE Transactions on Signal Processing*, vol. 45, pp. 600–616, 1997.