Caulking the “Leakage Effect” on MEEG Sources Connectivity Analysis

Deirel Paz-Linares[1,2#], Eduardo Gonzalez-Moreira[1,3#], Eduardo Martinez-Montes[2], Pedro A. Valdes-Hernandez[5], Jorge Bosch-Bayard[4], Maria Luisa Bringas-Vega[1,2] and Pedro A. Valdés-Sosa[1,2*]

*contributed equally as first authors*

deirel.paz@neuroinformatics-collaboratory.org

eduardo.g.m@neuroinformatics-collaboratory.org

***** *corresponding author*

pedro.valdes@neuroinformatics-collaboratory.org

Affiliations:

(1) *The Clinical Hospital of Chengdu Brain Science Institute, MOE Key Lab for Neuroinformation, University of Electronic Science and Technology of China, Chengdu, China*
(2) *Cuban Neuroscience Center, La Habana, Cuba*
(3) *Centro de Investigaciones de la Informática, Universidad Central “Marta Abreu” de las Villas*
(4) *Departamento de Neurobiología Conductual y Cognitiva, Instituto de Neurobiología, Universidad Nacional Autónoma de México. Boulevard Juriquilla 3001, Querétaro, 76230, México*
(5) *Department of Biomedical Engineering, Florida International University, Miami, FL, United States*

Abstract

Its currently acknowledged that a simplified estimation of connectivity at the Magnetoencephalography and Electroencephalography (MEEG) sensor space is fraught with several problems due to the volume conduction distortion. It would seem however, that a solution to this issue is carrying out connectivity analysis at the generators space. Meanwhile, a major criticism falls on that most of the source connectivity methods are distressed by the “Leakage Effect”, i.e. linear mixing of the reconstructed sources. We address this problem by means of a novel method, that allows us to “caulk” the “Leakage Effect” in MEEG sources activity and connectivity: BC-VARETA. This is done by leveraging two methodological aspects. First: Joint estimation of source activity and connectivity as a frequency domain linear dynamical system identification approach. Second: Incorporating priors into the sources graphical model of the connectivity estimator. Our claims are supported by a large simulation framework, that uses realistic head models, diverse sources setup, biological/instrumentation noisy signals, and Inverse Crime evaluation. Also, a fair quantitative analysis is performed based on measures of the ‘Leakage Effect’, in which state of the art inverse solvers were tested.

Theory

For the MEEG techniques the discrete measurements Forward Model, in the Frequency Domain representation of the recorded signals, is expressed by the general equation:

$$\boldsymbol{V} = \boldsymbol{K}\boldsymbol{J} + \boldsymbol{e} \qquad [2.1.1]$$

The MEEG measurements $\boldsymbol{V}$ and signal noise $\boldsymbol{e}$, are defined as Random vectors on the p-size Scalp Sensors space $\mathbf{Z}$ meanwhile the sources activity random vector $\boldsymbol{J}$ is defined on the q-size discretized Gray Matter space $\mathbf{\Omega}$. The random vectors of MEEG measurements $\boldsymbol{V}$, MEEG signal noise $\boldsymbol{e}$ and sources activity $\boldsymbol{J}$ represent 1st-dimension (spatial) fibers of 3-dimensional Tensors $\boldsymbol{\mathcal{V}}$, $\boldsymbol{\mathcal{E}}$ and $\boldsymbol{\mathcal{J}}$, i.e. $\boldsymbol{V} = \boldsymbol{\mathcal{V}}_{:,\not{f},m}$, $\boldsymbol{e} = \boldsymbol{\mathcal{E}}_{:,\not{f},m}$ and $\boldsymbol{J} = \boldsymbol{\mathcal{J}}_{:,\not{f},m}$. The Tensors dimensions represent correspondingly the Scalp Sensors $\mathbf{Z}$ (Gray Matter $\mathbf{\Omega}$), i.e. the 1st-dimension, Frequency Domain $\boldsymbol{\mathcal{F}}$, i.e. the 2nd-dimension, and Samples $\boldsymbol{\mathcal{M}}$, i.e. the 3rd-dimension. The $p \times q$-size design matrix $\boldsymbol{K}$ (transformation of spaces $\mathbf{\Omega} \rightarrow \mathbf{Z}$) builds on a discretization of the Lead Field from a head conductivity model.

Construing a Model of MEEG source localization and connectivity, upon the equation [2.1.1], can be tackled in general by the Bayesian formalism, which involves categorizing as random variables the MEEG measurements Tensor $\boldsymbol{\mathcal{V}}$ and source activity Tensor $\boldsymbol{\mathcal{J}}$ spatial fibers i.e. (***Data***) $\boldsymbol{V}$ and (***Parameters***) $\boldsymbol{J}$. The model builds on a Parametric representation Tensor spatial fibers of signal noise $\boldsymbol{e}$ and sources activity $\boldsymbol{J}$ Probability Density Functions (***pdf***). It is commonly given by embedded Gaussian Graphical Models (GGM), i.e. hierarchically conditioned Multivariate Gaussians N (in either the Real or Circularly Symmetric Complex case), of the ***Data*** Likelihood and ***Parameters*** Prior. The parametrization within these distributions introduces an additional category of random variables denominated (***Hyperparameters***) Ξ. Below we summarize the two levels GGM of the ***Data*** and ***Parameters***, along with the specification of ***Hyperparameters*** (parametrization) structure and Priors (defined as Gibbs ***pdf***s). See its schematic representation by More-Penrose diagrams in *Figure 1*.

$$\boldsymbol{V}|\boldsymbol{J},\Xi \sim N_p(\boldsymbol{V}|\boldsymbol{K}\boldsymbol{J},\sigma_e^2\boldsymbol{R}) \qquad [2.1.2]$$

$$\boldsymbol{J}|\Xi \sim N_q\left(\boldsymbol{J}\middle|0,\mathbf{\Theta}_{\boldsymbol{JJ}}^{-1}\right) \qquad [2.1.3]$$

$$p\left(\mathbf{\Theta}_{\boldsymbol{JJ}}\right) \propto e^{-\alpha P(\mathbf{\Theta}_{\boldsymbol{JJ}},\mathbf{\Lambda})} \qquad [2.1.4]$$

$$p(\sigma_e^2) \propto e^{-\beta\sigma_e^{-2}} \qquad [2.1.5]$$

Above, $\boldsymbol{V} = \boldsymbol{\mathcal{V}}_{:,\not{f},m}$, $\boldsymbol{J} = \boldsymbol{\mathcal{J}}_{:,\not{f},m}$ are the Tensor spatial fibers, $\boldsymbol{R} = \boldsymbol{\mathcal{R}}_{:,:,\not{f}}$ and $\mathbf{\Theta}_{\boldsymbol{JJ}} = \mathbf{\Theta}_{:,:,\not{f}}$ the Tensor spatial slices and $\Xi = \{\mathbf{\Theta},\sigma_e^2\}$.represents the Model Parametrization. The ***Data*** GGM conditional Covariance matrix, i.e. the product $\sigma_e^2\boldsymbol{R}$ in formula [2.1.2], is assumed to be composed by a scalar random variable σ_e^2, representing the unknown nuisance Variance, and a known matrix $\boldsymbol{R}$ (in the Cartesian space product $\mathbf{Z} \times \mathbf{Z}$) of the noise Covariance structure. The noise Covariance structure represents the 1st and 2nd spatial dimensions slice of a 3-dimensional Tensor $\boldsymbol{\mathcal{R}}$, encoding information about the sensors correlated activity for all frequency components $\boldsymbol{\mathcal{F}}$ (3rd-dimension of the Tensor). These correlations are given either by shorting currents between adjacent electrodes' due to the scalp conductivity or common inputs from instrumentation/environmental noisy sources. The noise Precision (Variance) Exponential (Jeffry Improper) Gibbs ***pdf*** set up on, see formula [2.1.5], aims to bypass the nuisance level that could be assimilated into the ***Parameters***. This is possible due to the monotonically increasing values of the noise Variance probability density assigned by the Jeffry Improper Prior, which allows for encoding the information about the noise inferior threshold into the parameter β.

The Precision matrix $\mathbf{\Theta}_{\boldsymbol{JJ}}$ (in the Cartesian space product $\mathbf{\Omega} \times \mathbf{\Omega}$) of the ***Parameters*** GGM, represents the 1st and 2nd spatial dimensions slice of a 3-dimensional Tensor $\mathbf{\Theta}$ of source connectivity for all frequency components $\boldsymbol{\mathcal{F}}$ (3rd-dimension of the Tensor), see formula [2.1.3]. The general penalization function P at the argument of the Gibbs Prior in formula [2.1.4], imposes certain Structured Sparsity pattern on the connectivity. The Structured Sparsity can be encoded given information from the Gray Matter anatomical segmentation, by penalizing the groups of variables corresponding to the Gray Matter areas Intra/Interconnections. The Matrix $\mathbf{\Lambda}$ (in the Cartesian space product $\mathbf{\Omega} \times \mathbf{\Omega}$) within the General Penalization function, represents a probability mask of the anatomically plausible connections. The probability mask in case of the dense short-range connections, e.g. Intra-Cortical Connections, is defined as a deterministic spatially invariant empirical Kernel of the connections strength decay with distance. For the long-range connections, e.g. Inter-Cortical connection, it is given by probabilistic maps of the White Matter tracks connectivity strength from Diffusion Tensor Imaging (DTI). The global influence in the ***Parameters*** GGM of the connectivity Structured Sparsity penalization $P(\mathbf{\Theta_{JJ}}, \mathbf{\Lambda})$ is controlled by the Regularization Parameter α, which can be fitted to the ***Data*** by means of some statistical criteria of goodness.

In general, the construction of the Model corresponds to the ubiquitous Bayesian representation of Linear State Space Models (LSSM), in both Time (Real) and Frequency (Complex) domain. The LSSM ***Data*** (Observation Equation) and ***Parameters*** (Autoregressive State Equation) are modeled by Multivariate Gaussian ***pdf*** 's, whereas the Connectivity (Autoregression Coefficients Matrix) is represented by the Parameters' Precision Matrix.

Results

Simulation substrate: We evaluate the proposed estimators of Sources Localization and Connectivity on simulated EEG data. The simulation substrate was set up on a Cortical Manifold space $\mathbf{\Omega}$, defined as 15K points Surface of the Gray Matter, with coordinates on the MNI Brain template (http://www.bic.mni.mcgill.ca). The Scalp Sensors space $\mathbf{Z}$ was built on 343 electrodes, within 10-5 EEG Sensors system. The Lead Fields, for both Simulations $\boldsymbol{K}_{sim}$ and for Reconstruction $\boldsymbol{K}$, were computed by BEM integration method accounting for a model of 5 head compartments (gray matter, cerebrospinal fluid, inner skull, outer skull, scalp). To avoid the Inverse Crime two individual subject Head Models were extracted from the corresponding T1 MRI images. The Electrophysiological Noise was defined by a composition of Sensors Noise and 500 Noisy Cortical Sources (approximating a 3% of the 15K points of the Cortical Manifold space $\mathbf{\Omega}$), projected to the Scalp Sensors space $\mathbf{Z}$. For practical computational limitations the Reconstruction was compute on a 6K points of Cortical Manifold space $\mathbf{\Omega}$, obtained as a reduction from the original 15K points Surface of the Gray Matter.

Comparisons: For comparison purpose we selected the Standardized version of LORETA (sLORETA), which constitutes the most stablished and robust (under a wide range of conditions), among the of family Source Activity estimators described in ***Table 1***. The sLORETA ***SD-PSF*** and ***EMD-GSF*** were computed from the SEC at the Fist Level of Inference after convergence, according to the theory in *Subsection 2.2.1*. The ***SD-CPSF*** and ***EMD-CGSF*** was taken from the Inverse of the Source Covariance matrix at the Second Level of Inference after convergence.

Source set up: Four Sources were placed in 500 random configurations where only three of them were connected. 400 EEG Data trials were created, at the Scalp Sensors space, by projecting Samples from a Gaussian Random Generator given for each configuration. The Gaussian Random Generator Covariance structure was defined as the Inverse of the Unitary Precision matrix (Connectivity), given in formula [2.4.3]. Also, 5dB Level Noise were added to the Data of each projected Sample. The simulation was set with the following aims. First: Evaluating the Type I Leakage given the composition the distortive effects of Volume Conduction and multiple Sources superposed at the Scalp projection. Second: Evaluating the Type II Leakage given to Volume Conduction distortion in a scenario where multiple connections are present. See the results in *Figure 2.*

Figure1: More-Penrose diagram of the MEEG Sources Activity and Connectivity Bayesian Model and its Priors. The Model Variables and Prior Knowledge are represented with Gray Circles and Squares respectively.

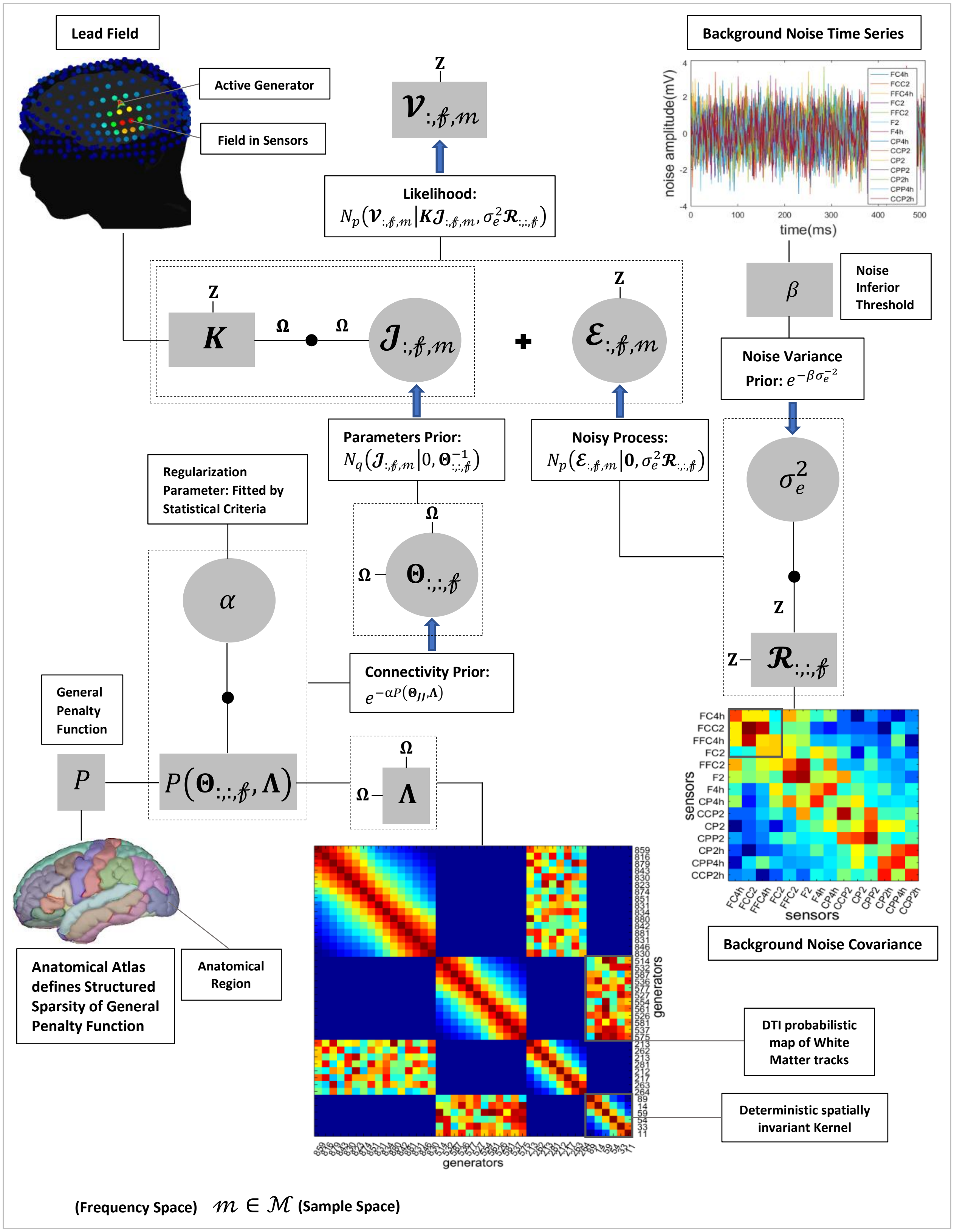

Figure 2: A typical simulation trial to evaluate the Volume Conduction effect at the scalp and Source analysis performance. From left to right (activations are shown at the upper part and connectivity at the bottom part): 1- The simulated point sources configuration and connectivity. 2- Scalp projection of the activity and connectivity. 3- Activity and connectivity estimated with eLORETA method. 4- Activity and connectivity estimated with the BC-VARETA method. Both the bidimensional and tridimensional thermal colormaps are shown within an interval between 0 and maximum value 1.