

An extendable simulation framework for benchmarking EEG-based brain connectivity estimation methodologies*

Stefan Haufe¹

Abstract—Due to its high temporal resolution, electroencephalography (EEG) is a promising research tool for studying functional and effective brain interaction. Yet, it is rather uncommon for researchers to validate their connectivity estimation methodologies prior to applying them to real data, even though problems have been pointed out regarding the validity of some of the predominant approaches. We here provide an extendable simulation framework that enables researchers to test their analysis pipelines on customizable realistically simulated EEG data. We define three simple criteria to measure source localization, connectivity detection and directionality estimation performance. All data and code needed to generate pseudo-EEG data and to benchmark a method’s estimation performance are provided.

I. INTRODUCTION

The reconstruction of the activation time courses and locations of the neuronal populations (here referred to as *sources*) contributing to an EEG (electroencephalography) measurement, is a challenging inverse problem. To date, a variety of methods serving that purpose exist, accounting for the fact that different experimental settings may require different source characterizations. Inverse source reconstruction algorithms have been validated in a considerable body of literature using simulated and real data.

A related problem is the EEG-based estimation of functional or effective brain connectivity. Here, not the source activity itself is of interest, but the interaction between brain sites, which may be estimated from an inverse solution in a subsequent step. Brain connectivity analyses are more challenging than inverse source reconstructions alone, since they rely both on the correct estimation of the source activity and the correct subsequent inference of the interaction structure from the estimated sources.

The property of being inherently two-step procedures makes systematic benchmarking of connectivity analyses difficult. At the same time, it is also difficult if not impossible to devise an EEG experiment in which the interactions between all sources contributing to the measurements (that is, the entire brain) are known. It is therefore rather common to publish results of EEG-based brain connectivity analyses without or with only limited empirical validation of the employed methodology. Common limitations include the disregard or insufficient modeling of the source mixing caused

*This work was supported by a Marie Curie International Outgoing Fellowship (grant No. PIOF-GA-2013-625991) within the 7th European Community Framework Programme.

¹Stefan Haufe is with the Laboratory for Intelligent Imaging and Neural Computing, Columbia University, New York, NY 10027, USA, and with the Machine Learning Dept., Technische Universität Berlin, 10587 Berlin, Germany. stefan.haufe@tu-berlin.de

by volume conduction in the head, the disregard of correlated noise sources, and generally the overestimation of the signal-to-noise ratio (SNR). This is problematic since, it is especially the mixing of signal and noise sources present in any real EEG dataset that causes established connectivity measures such as Granger Causality to suffer from wrongly detected connections as well as missed connections [11], [6], [5], [14].

Considering the steadily growing body of literature on EEG-based brain connectivity analysis, the field would benefit from a standardized benchmark. In this work, we present an extendable simulation framework that can be used to generate realistic EEG data from underlying interacting brain sources, as well as to validate the performance of inverse source reconstruction and brain connectivity analyses on that data.

II. A SOURCE CONNECTIVITY BENCHMARK

In order to provide a standardized environment capturing the most important aspects of real EEG recordings, our simulation framework is characterized by

- The use of a high-resolution average anatomy template aligned to the Montreal Neurological Institute (MNI) coordinate system.
- The use of a realistic volume conductor model.
- The use of the extended 10/20 electrode montage.
- The presence of interacting sources exerting time-delayed influence on another.
- Interactions being confined to a narrow frequency band.
- Realistic source locations being confined to the cortical manifold and emitting electrical currents perpendicular to the local surface.
- Generally, varying locations, spatial extents and depths of the sources.
- The presence of independent background brain processes with pink noise spectra.
- The presence of white sensor noise.
- Realistic SNR ranges.

At the same time, the simulation setting is designed to be as simple as possible in order to ensure a transparent and unambiguous performance evaluation. That is, while the framework can be extended in many possible ways, we here consider only

- The presence of only two interacting sources.
- Linear interaction.
- Uni-directional information flow.
- Spatially non-overlapping sources.

Since mislocalization or misestimation of source activity will negatively affect subsequent connectivity analysis, the problem of source connectivity estimation cannot be decoupled

from the problem of inverse source reconstruction. Consequently, we here provide means to evaluate methodologies both with respect to source reconstruction and connectivity estimation accuracy. Precisely, all evaluations are carried out only on the coarse level of eight regions of interest (ROIs) that are identical to the octants of the brain, which allows us to introduce very simple performance metrics.

All data and code required to generate pseudo-EEG datasets will be made publicly available at <http://bbci.de/supplementary/EEGconnectivity/>. This includes lead fields, source and electrode coordinates, as well as Matlab scripts for generating simulated data, evaluating performance metrics, and conducting source reconstructions and connectivity estimations using baseline methods from the literature. Researchers benchmarking their methods have the opportunity to adjust and extend this simulation framework according to their needs.

A. Forward model of EEG data

Inverse source reconstruction requires a forward model of how neural activity in the brain maps to the sensors. In its discretized form, this model reads

$$\mathbf{x}(t) = \mathbf{L}\mathbf{j}(t) + \boldsymbol{\varepsilon}(t). \quad (1)$$

Here, the time-dependent $3R$ -dimensional vector $\mathbf{j}(t)$ represents the directed intracellular neuronal currents at R distinct locations on the cortical surface. The $M \times 3R$ lead field matrix \mathbf{L} describes the relationship between intracellular currents and the observable scalp potentials at M sensors. Finally, $\boldsymbol{\varepsilon}(t)$ is a M -dimensional noise vector. Inverse source reconstruction is concerned with the estimation of the source primary currents $\mathbf{j}(t)$ given the measurements $\mathbf{x}(t)$ in a given head model \mathbf{L} , while functional or effective brain connectivity estimation is concerned with the estimation of the information flow between brain sites from either $\mathbf{j}(t)$ or (less commonly) $\mathbf{x}(t)$.

B. Head model

Lead fields were computed in the ICBM152b brain, a standard head geometry obtained by nonlinearly averaging the anatomical magnetic resonance images of 152 adults [3]. Surfaces of the brain, skull and skin shells were extracted using the Brainstorm software [13]. Within this 3-shell geometry, the EEG forward problem was solved using the boundary element method (BEM) [4]. Note here that the final version of this benchmark will make use of a highly-detailed finite element model (FEM) of the same anatomy [16]. The forward model was evaluated at 2504 nodes of a mesh of the cortical surface, and for 64 EEG electrodes to yield the lead field. Only for plotting purposes, a higher-resolution cortical mesh was created to which source distributions defined on the lower-dimensional mesh could be mapped. Cortical surfaces were extracted using the BrainVISA Morphologist toolbox. Electrodes were placed according to the extended international 10-20 system [12].

C. Regions of interest

Eight regions of interest (ROIs) being identical to the octants of the brain are defined. Two ensure that octants cover brain areas of roughly similar size, octant boundaries were determined based on cutting the cortical mesh into two halves each containing an equal number of nodes. The cutting planes obtained this way are defined by the equations $x = 0$ mm (separating left and right hemispheres), $y = -18.7$ mm (separating anterior and posterior hemispheres) and $z = 12.8$ mm (separating superior and inferior hemispheres). All coordinates given are in MNI space. The combination of the three hyperplanes defines eight octants as shown in Figure 1.

Octant	Definition ($[x, y, z]$ in mm)	Color
RAI	$x \geq 0$ & $y \geq -18.7$ & $z < 12.8$	Gray
RAS	$x \geq 0$ & $y \geq -18.7$ & $z \geq 12.8$	Pink
RPI	$x \geq 0$ & $y < -18.7$ & $z < 12.8$	Brown
RPS	$x \geq 0$ & $y < -18.7$ & $z \geq 12.8$	Yellow
LAI	$x < 0$ & $y \geq -18.7$ & $z < 12.8$	Orange
LAS	$x < 0$ & $y \geq -18.7$ & $z \geq 12.8$	Green
LPI	$x < 0$ & $y < -18.7$ & $z < 12.8$	Blue
LPS	$x < 0$ & $y < -18.7$ & $z \geq 12.8$	Red

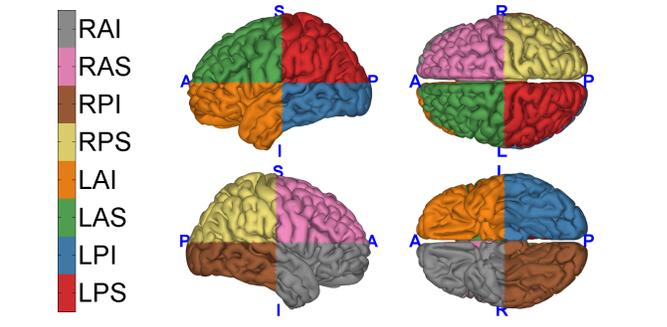


Fig. 1. Division of the brain into eight octants.

D. Spatial structure of the sources

In each simulated dataset, two distinguished sources are modeled. These two sources are constrained to lie in different randomly sampled brain octants. Within each octant, a random node of the cortical mesh is picked as the center of the source activity. The center nodes are required to be at least 10 mm away from the octant boundaries. Note that the randomized sampling of source locations leads to a considerable variation of source depth with sources in inferior regions being deeper than sources in corresponding superior regions, and sources in posterior regions being deeper than corresponding sources in anterior regions (see Figure 2). Here, depth is defined as the mean Euclidean distance of the center node from all scalp electrodes.

The spatial distribution of the source current amplitudes is modeled by a Gaussian function, where the geodesic distance between nodes of the cortical mesh is used as the distance metric. The spatial standard deviation of the amplitude distributions is sampled uniformly between 10 mm and 40 mm. The amplitude at nodes located outside the seed octant is

set to zero, such that no ‘leakage’ of activity across octant borders occurs, and the true connectivity between octants can be defined unambiguously. The amplitude distributions are divided by their ℓ_2 -norm for each source separately. The orientation of the neuronal current at each node is defined as the normal vector w. r. t. the mesh surface at that node. Scalp topographies for each source are computed by multiplying the 3D current distribution (the product of amplitude and orientation) with the lead field, that is, by summing up the contributions from all nodes of the source octant. Figure 2A depicts the source amplitude distributions, as well as the resulting scalp potentials, for two representative sources.

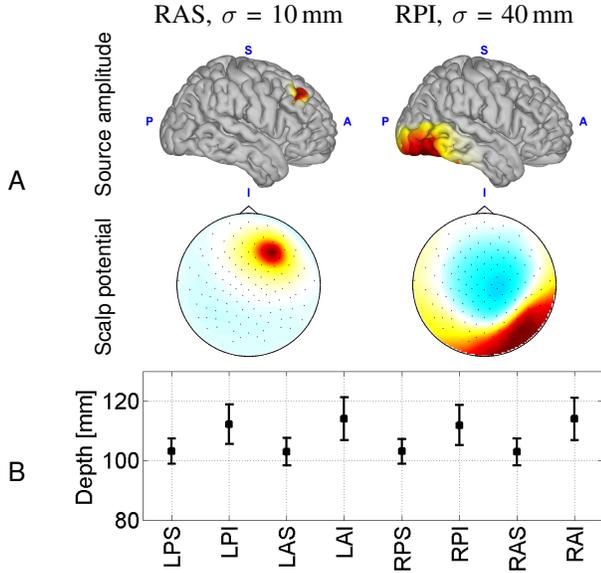


Fig. 2. A: examples of brain sources. Left: source with small spatial extent (spatial standard deviation along cortical manifold $\sigma = 10$ mm) in the right anterior superior octant of the brain. Right: source with large ($\sigma = 40$ mm) spatial extent in the right posterior inferior octant of the brain. Upper panel: source amplitude distribution. Note that sources do not extend into neighboring octants. Lower panel: resulting theoretical EEG field potential assuming currents oriented perpendicular to the cortical manifold. B: depth distribution (mean and standard deviation) of the cortical surface points belonging to each of the brain octants.

E. Source dynamics

The time courses of the two distinguished sources are modeled using bivariate linear autoregressive (AR) models of the form

$$\begin{bmatrix} j_1(t) \\ j_2(t) \end{bmatrix} = \sum_{p=1}^P \begin{bmatrix} a_{11}(p) & a_{12}(p) \\ a_{21}(p) & a_{22}(p) \end{bmatrix} \begin{bmatrix} j_1(t-p) \\ j_2(t-p) \end{bmatrix} + \begin{bmatrix} \epsilon_1(t) \\ \epsilon_2(t) \end{bmatrix},$$

where the $a_{ij}(p)$, $i, j \in \{1, 2\}$, $p \in \{1, \dots, P\}$ are linear AR coefficients, and $\epsilon_i(t)$, $i \in \{1, 2\}$ are uncorrelated standard normal distributed noise variables (innovations). Importantly, the offdiagonal entries $a_{12}(p)$ and $a_{21}(p)$ describe time-delayed linear influences of one source on another. A sampling rate of 100 Hz, and an AR model order of $P = 5$ is used.

Two variants of the linear system $\mathbf{j}(t) = [j_1(t), j_2(t)]^\top$ are constructed. For the first variant, \mathbf{j}^{int} , $a_{12}(p)$, $p \in \{1, \dots, P\}$ is set to zero for all lags p , modeling a unidirectional time-delayed influence of $j_1^{\text{int}}(t)$ on $j_2^{\text{int}}(t)$. For the second

variant, $\mathbf{j}^{\text{nonint}}$, all offdiagonal coefficients $a_{12}(p)$ and $a_{21}(p)$, $p \in \{1, \dots, P\}$ are set to zero, leaving the two time series $j_1^{\text{nonint}}(t)$ and $j_2^{\text{nonint}}(t)$ completely independent. The AR coefficients are sampled from the univariate standard normal distribution. Only stable AR system, for which the combined spectral power of the two sources in the alpha band (8–13 Hz), normalized by the width of the alpha band, is higher than the overall normalized power, are selected. Sources are bandpass-filtered in the alpha band using a third-order Butterworth filter. The resulting time series thus represent non-/interacting alpha-band oscillations being either independent or characterized by a clearly defined sender-receiver relationship.

F. Generation of pseudo-EEG data

Pseudo-EEG data containing simulated underlying brain interaction are created as follows. A total of $T = 18000$ samples of the source time series $\mathbf{j}^{\text{int}}(t)$ are generated, corresponding to a 3 minute recording. The source time courses are then mapped to two patches of the cortical surface and projected to the EEG sensors through multiplication with the lead field, giving rise to the signal contribution $\mathbf{s}^{\text{int}}(t)$ of the EEG. In addition, 500 brain noise time series obeying $1/f$ -shaped (pink noise) power spectra and random phases are generated. These noise sources are placed at 500 locations randomly sampled from the entire cortical surface, and are also mapped to EEG sensor space. This procedure yields the brain noise contribution $\mathbf{n}^{\text{brain}}(t)$ of the EEG. In addition, spatially and temporally uncorrelated sensor noise $\mathbf{n}^{\text{sensor}}(t)$ is sampled from a univariate standard normal distribution. The overall noise contribution is defined as

$$\mathbf{n}(t) = 0.9 \frac{\mathbf{n}^{\text{brain}}(t)}{\|\mathbf{n}^{\text{brain}}(t)\|_F} + 0.1 \frac{\mathbf{n}^{\text{sensor}}(t)}{\|\mathbf{n}^{\text{sensor}}(t)\|_F}, \quad (2)$$

where $\|\mathbf{n}(t)\|_F$ is the Frobenius norm. The overall pseudo-EEG measurement is generated according to

$$\mathbf{x}^{\text{int}}(t) = \alpha \frac{\mathbf{s}^{\text{int}}(t)}{\|\mathbf{s}^{\text{int}}(t)\|_F} + (1 - \alpha) \frac{\mathbf{n}(t)}{\|\tilde{\mathbf{n}}(t)\|_F}, \quad (3)$$

where $\tilde{\mathbf{n}}(t)$ is the alpha-band filtered version of $\mathbf{n}(t)$. The signal-to-noise parameter α is drawn uniformly from the interval $[0.1, 0.9]$. Lastly, a highpass filter at 0.5 Hz (third order Butterworth) is applied to $\mathbf{x}^{\text{int}}(t)$.

In the same way as $\mathbf{x}^{\text{int}}(t)$, a second pseudo-EEG measurement $\mathbf{x}^{\text{nonint}}(t)$ containing no form of brain interaction at all is generated. The purpose of this pseudo-measurement is to benchmark methods in their ability to detect the presence or absence of time-delayed brain interaction at all.

G. Task

For a given dataset comprising $\mathbf{x}^{\text{int}}(t)$, $\mathbf{x}^{\text{nonint}}(t)$ (renamed as $\mathbf{x}^1(t)$ and $\mathbf{x}^2(t)$ so that no information about the presence of underlying connectivity is revealed), the following three questions are asked.

- 1) Localization: In which two brain octants are the alpha-band signal sources are located?
- 2) Connectivity: Which one of the two datasets $\mathbf{x}^1(t)$ and $\mathbf{x}^2(t)$ contains actual simulated brain interaction?
- 3) Direction: In the dataset estimated to contain interacting sources, which one of the two octants estimated to contain the sources is sending information, and which one is receiving information?

H. Performance measures

The correctness of the answers to the above-mentioned questions is evaluated in a straightforward way using the following three performance measures. It is suggested that researchers analyze $K \geq 100$ datasets and report mean and standard errors of these three measures.

LOC: This measure compares the true octants containing the alpha-band sources with the estimated ones. Each octant estimated correctly adds a score of $1/2$. The measure therefore takes one of the values 0, $1/2$ and 1, while the expected value under random guessing is $\approx 1/4$. Neither the order in which octants are given, nor their assumed interaction is taken into account.

CONN: This measure evaluates whether the correct one of the anonymized measurements $\mathbf{x}^1(t)$ and $\mathbf{x}^2(t)$ has been estimated to contain actual time-delayed brain interaction. Correct estimates lead to a score of +1, whereas incorrect estimates lead to a score of -1. Importantly, researchers can refuse to make a decision, which leads to a score of 0. Thus, CONN may take one of the values -1, 0 and +1, where the expected value under random guessing is 0.

DIR: This measure evaluates the correct assessment of interaction directionality by comparing the estimated connectivity between estimated sources with the true connectivity between simulated sources. The evaluation is split into three parts corresponding to the assessment of interactions between left and right, anterior and posterior, as well as superior and inferior hemispheres. Due to this split, DIR can provide a positive measure of connectivity estimation performance even in the case of moderate source mislocalization.

For a given pair of true and estimated source octants, as well as the true and estimated directionality between those, DIR evaluates for each of the three spatial directions separately whether the estimated flow is compatible with the true flow. For each correct estimate, a score of $+1/3$ is given, whereas for each incorrect estimate $-1/3$ is given. Refusal to decide on direction leads to a score of 0. If the direction is estimated, but CONN = -1, then the DIR score is also -1. The scores attainable by DIR are therefore -1, $-1/3$, 0, $1/3$, and 1, while the expected value under random guessing is $\approx -1/2$.

I. Strategies

It is generally up to the researcher using the benchmark to decide in which way and order they infer the location and connectivity of the two sources. One way would be to apply inverse source reconstruction [7], [8] first, define the source octants based on power maps, and then analyze connectivity between those octants in a second step. Another valid way would be to analyze the full connectivity graph after source localization, and then choose source octants based on maximal connectivity. Approaches using blind source separation techniques or avoiding source representations entirely are also in principle valid as long as they lead to source octant and connectivity estimates. An example estimation pipeline serving as a baseline will be included in the full-length version of this paper.

J. Extensions

The framework can be freely modified or extended.

Non-linear interaction: Currently, only linear dynamics are considered through the use of linear AR models. For researchers proposing non-linear or non-parametric approaches to connectivity estimation [15], [9], it will be useful to create a variant of the benchmark simulating sources with particular non-linearities.

Bi-directional interaction: The current framework only considers unidirectional information flow, while studying the more realistic bidirectional case is also worthwhile [14].

More than two interacting sources and brain network analyses: Graph-theoretical analyses of networks derived from, e. g., Granger-causal analyses are becoming increasingly popular [2], [10], [1]; it would therefore be useful to extend the simulation setting to permit testing of such approaches.

III. CONCLUSIONS

We present a simulation framework enabling researchers working in the field of EEG-based brain connectivity to validate their approaches.

ACKNOWLEDGMENT

We thank Pedro Valdes-Sosa for suggesting the creation of a benchmark, and Guido Nolte for providing Matlab code.

REFERENCES

- [1] P. Barttfeld et al. Functional connectivity and temporal variability of brain connections in adults with attention deficit/hyperactivity disorder and bipolar disorder. *Neuropsychobiology*, 69(2):65–75, 2014.
- [2] F. De Vico Fallani et al. Cortical functional connectivity networks in normal and spinal cord injured patients: Evaluation by graph analysis. *Hum Brain Mapp*, 28(12):1334–1346, 2007.
- [3] V. Fonov et al. Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage*, 54(1):313–327, 2011.
- [4] A. Gramfort et al. OpenMEEG: opensource software for quasistatic bioelectromagnetics. *Biomed Eng Online*, 9:45, 2010.
- [5] S. Haufe et al. A critical assessment of connectivity measures for EEG data: a simulation study. *NeuroImage*, 64:120–133, 2012.
- [6] S. Haufe et al. Alleviating the influence of weak data asymmetries on Granger-causal analyses. In *Latent Variable Analysis and Signal Separation*, volume 7191 of *Lecture Notes in Computer Science*, 25–33.
- [7] S. Haufe et al. Estimating vector fields using sparse basis field expansions. In *Advances in Neural Information Processing Systems 21*, 617–624.
- [8] S. Haufe et al. Large-scale EEG/MEG Source Localization with Spatial Flexibility. *Neuroimage*, 54:851–859, 2011.
- [9] D. Marinazzo et al. Nonlinear connectivity by Granger causality. *Neuroimage*, 58(2):330–338, 2011.
- [10] T. Mullen et al. Modeling cortical source dynamics and interactions during seizure. *Conf Proc IEEE Eng Med Biol Soc*, 2011:1411–1414, 2011.
- [11] G. Nolte et al. Robustly estimating the flow direction of information in complex physical systems. *Phys Rev Lett*, 100:234101, 2008.
- [12] R. Oostenveld and P. Praamstra. The five percent electrode system for high-resolution EEG and ERP measurements. *Clin Neurophysiol*, 112:713–719, 2001.
- [13] F. Tadel et al. Brainstorm: a user-friendly application for MEG/EEG analysis. *Comput Intell Neurosci*, 2011:879716, 2011.
- [14] M. Vinck et al. How to detect the Granger-causal flow direction in the presence of additive noise? *Neuroimage*, 108:301–318, 2015.
- [15] M. Wibral et al. Transfer entropy in magnetoencephalographic data: quantifying information flow in cortical and cerebellar networks. *Prog. Biophys. Mol. Biol.*, 105(1-2):80–97, 2011.
- [16] S. Haufe et al. A highly detailed FEM volume conductor model based on the ICBM152 average head template for EEG source imaging and TCS targeting. *Conf Proc IEEE Eng Med Biol Soc*, 2015. In Press.