IOP Publishing J. Neural Eng. 18 (2021) 066014 https://doi.org/10.1088/1741-2552/ac2bb6

Journal of Neural Engineering

RECEIVED
23 December 2020

REVISED
17 August 2021

ACCEPTED FOR PUBLICATION
30 September 2021

PUBLISHED
16 November 2021

PAPER

Source space reduction for eLORETA

A Faes*, A de Borman and M M Van Hulle

KU Leuven—University of Leuven, Department of Neurosciences, Laboratory for Neuro- & Psychophysiology, B-3000 Leuven, Belgium
* Author to whom any correspondence should be addressed.

E-mail: axel.faes@kuleuven.be

Keywords: source localization, sparsity constraint, eLORETA

Supplementary material for this article is available online

Abstract

Objective. We introduce Sparse exact low resolution electromagnetic tomography (eLORETA), a novel method for estimating a nonparametric solution to the source localization problem. Its goal is to generate a sparser solution compared to other source localization methods including eLORETA while benefitting from the latter's superior source localization accuracy. *Approach.* Sparse eLORETA starts by reducing the source space of the Lead Field Matrix using structured sparse Bayesian learning from which a Reduced Lead Field Matrix is constructed, which is used as input to eLORETA. *Main results.* With Sparse eLORETA, source sparsity can be traded against signal fidelity; the proposed optimum is shown to yield a much sparser solution than eLORETA's with only a slight loss in signal fidelity. *Significance.* When pursuing a data-driven approach, for cases where it is difficult to choose specific regions of interest, or when subsequently a connectivity analysis is performed, source space reduction could prove beneficial.

1. Introduction

It is commonly accepted that perception, motor behavior, language, and cognition emerge from coordinated information flows between functionally specialized regions. As these activities and flows are dynamic, imaging techniques such as MEG (magnetoencephalography) and (scalp-based) EEG (electroencephalography) are preferred over spatially more accurate yet temporally less detailed ones such as fMRI (functional magnetic resonance imaging) and PET (positron emission tomography) [1, 2].

When localizing active brain regions, MEG and EEG signals need to be projected back into the brain, a procedure called source localization [3, 4]. It requires solving an inverse problem (from scalp electrode- or MEG sensor space to 3D brain space) that is bound by the recordings not providing sufficient information to guarantee a unique and stable solution [5]. The recordings reflect mixtures of local neural generators of which the activity is spatially smeared out due to volume conduction, causing localization errors or even a failure of the localization procedure.

In the case of EEG, source localization starts with a forward equation expressing the scalp electric potential differences as a function of the current densities produced by the neural generators, an easy to solve problem using linear algebra. The difficulty arises when the scalp potential differences are known and we need to solve the inverse solution for unknown amplitudes and orientations of the current densities [6]. Because there are many more unknowns than equations, additional assumptions are needed, leading to several algorithmic approaches. Two mainstream source localization techniques have been proposed: the equivalent current dipole (parametric) and the distributed source models (nonparametric) [6]. The latter differs from the former in that it does not require the prior specification of the number of neural sources. The former solves the inverse solution to obtain the location and orientation of these sources. The latter assumes the entire brain volume to consist of fixed source locations and solves the inverse solution to estimate their amplitudes; as these sources are by no means assumed to be independent, they cannot be regarded as different neural generators. There is a tendency to rely on distributed source modeling when there is no prior information about the expected active brain regions, in the opposite case about the equivalent current dipoles. Within the realm of the nonparametric approaches, several source localization methods have been proposed.

Linearly constrained minimum variance (LCMV) [7, 8] is a spatial filtering method that accepts activity originating from a specific location, while attenuating that from other locations, thus a way of beamforming. The result is an estimate of the activity generated by the target brain locations. LCMV assumes that the sources are not temporally correlated. The inverse filter is based on minimizing the source variance at a given location, subject to a 'unit-gain constraint'. The mapping between multiple sources and scalp-EEG is obtained by combining corresponding LCMV filters.

Exact low resolution electromagnetic tomography (eLORETA) [9] is a linear inverse method characterized by a spatially smooth current density. It belongs to the LORETA family with zero localization error in the case of one active dipole and no noise. Several studies in realistic settings (multiple sources, noise, actual recordings) reported that eLORETA outperforms other linear methods in localization accuracy [10].

Brain connectivity variable resolution tomographic analysis (BC-VARETA) is a recent inverse method developed by Gonzales *et al* [11] that has already been adopted in a few studies [12–14]. The method estimates the inverse solution and its precision matrix, which represents the connectivity parameters, by using the frequency domain representation of the stationary time series. It belongs to the category of parametric source localization methods. As BC-VARETA is computationally expensive, to reduce source space dimensionality, a prior screening of the sources is performed using structured sparse Bayesian learning (SSBL), which has a parameter that controls to what extent (in %) source space is reduced, called the sparsity constraint (κ). If it is set to 5%, then source space is reduced to 5% of its original size [15]. The extent of dimensionality reduction is controlled by an upper bound on the number of selected sources so that κ is chosen depending on the context. One should note that BC-VARETA localizes sources but does not reconstruct their time series. However, it has several interesting features, one of which is the much sparser solution it produces than mainstream source localization techniques. When applied to the Cuban Brain Mapping Project, BC-VARETA's results have been shown to be in alignment with the physiology of human resting state EEG in the spectral domain [11].

In this paper, we propose a novel method for source localization from multiple stationary time series. The method aims at a sparser solution compared to other source localization methodologies while maintaining localization accuracy. We first introduce our source localization method called Sparse eLORETA. Then, in section 3, we apply the proposed method to one simulated- and two real-world EEG datasets, the Cuban Human Brain Project, and the Berlin Brain Connectivity Benchmark and Data Analysis Challenge Dataset, and report the results. Finally, we summarize and discuss our results and draw our main conclusion.

Algorithm 1. Source space reduction for eLORETA.

$[miu] \leftarrow ssbl(Svv, HM, seg, dip\_map)$
$[indms] \leftarrow smoothing(miu, \kappa * dip, vert, fac, ind)$
$ReducedHM \leftarrow HM(:, indms)$
$P \leftarrow eLORETA(ReducedHM, \gamma)$

2. Materials and methods

The goal of the method we propose is to generate a sparser solution compared to other source localization methods including eLORETA while benefitting from the latter's superior source localization accuracy. When pursuing a data-driven approach, for cases where it is difficult to choose *a priori* regions of interest (ROIs), or when subsequently a connectivity analysis is performed, source space reduction could prove beneficial. The advantage of BC-VARETA in this respect is that it generates a sparse, yet accurate source localization. This motivated us to utilize several components from BC-VARETA to sparsify source localization. Conceptually, we use the source localization map obtained from BC-VARETA as a mask for eLORETA. In what follows we briefly discuss BC-VARETA. Implementation details, as well as of the other used methods, are listed in the supplementary materials section (available online at stacks.iop.org/JNE/18/066014/mmedia).

First, we apply SSBL to reduce the source space of the Lead Field Matrix. SSBL extracts the possibly active generators and gives us the indices of dipoles that are potentially active. These indices are used to reconstruct a new, partial Lead Field Matrix. Finally, the resulting Lead Field Matrix is used within eLORETA to generate the inverse solution.

In algorithm 1, we list the pseudocode behind Sparse eLORETA. First, we apply SSBL to reduce the source space of the Lead Field Matrix. SSBL requires the Covariance Matrix of the input data (*Svv*), the head model (*HM*), the number of epochs (*seg*) and a dictionary mapping each dipole onto itself (i.e. $1 \rightarrow 1$). SSBL extracts the possibly active generators as a list of indices (*miu*). In the next step, a smoothing effect is applied to the results from SSBL. Here we use the κ parameter to indicate to what percentage of all dipoles (*dip* indicates the number of dipoles) smoothing should prune dipoles. When taken too small, relevant dipoles could be pruned, when taken too large irrelevant dipoles could remain. Note that the smoothing function also requires the location of all dipoles (*vertices*), the triangulation (*fac*), and the range of all dipole indices (*ind*). The smoothing step returns the indices of the remaining, potentially active dipoles, and this concludes source space reduction. These indices are used to construct the reduced head

model that is input to eLORETA, which generates the inverse solution P. eLORETA takes only one other parameter, γ, which is the regulariation parameter.

2.1. Simulated EEG dataset

A simulation study provides us with a ground truth against which our results can be verified and compared with those of other methods. We adopt the simulation framework of Anzolin *et al* which consists of several steps: the generation of brain signals with a predetermined connectivity pattern as ground truth, noise generation, forward modeling, inverse modeling, connectivity estimation, and performance evaluation [16]. However, we propose a different approach for the first step, as we are not interested in connectivity estimation.

Pseudo-EEG data were generated by employing the toolbox described in the work from Haufe *et al* [17]. We simulated three time series of 500 samples. A three-dimensional multivariate autoregressive model of order 2 was used to simulate pseudo-EEG data. This process has previously been used to evaluate non-stationary, directed interactions in multivariate neural data [18]. Hereby, x_1, x_2 and x_3 represent the electrical activity of three dipoles within the brain. Each one is an active source contributing to the pseudo-EEG measured on the scalp assuming an EEG cap with 108 electrodes.

$$\begin{cases} x_1(n) = 0.5x_1(n-1) - 0.7x_1(n-2) \\ \qquad + 0.25x_2(n-1) + w_1(n) \\ x_2(n) = 0.7x_2(n-1) - 0.5x_2(n-2) \\ \qquad + 0.2x_1(n-1) + 0.25x_3(n-1) + w_2(n) \\ x_3(n) = 0.8x_3(n-1) + w_3(n). \end{cases} \tag{1}$$

In order to probe the robustness of source localization, we considered not only brain sources but also 500 noise sources as background activity. We considered multiple types of noise: (incoherent) pink noise (i.e. standard pink noise), coherent autoregressive noise [19], and coherent pink noise [20], and gauged their effect on source localization accuracy.

BC-VARETA requires information about the cortex mesh. We have used the New York Head Model available in the Berlin Brain Connectivity Benchmark and Data Analysis Challenge with mesh information in terms of vertex coordinates and triangles or faces [21]. We introduced 2004 electric equivalent dipoles, homogeneously distributed and located using the ICBM152v2009 stereotaxic registration model. The 2004 dipole positions were obtained by subsampling the 75 000 MNI coordinates available in the New York Head Model [22].

For each simulation, a different triple dipole configuration was created by having two dipoles at fixed positions and the third dipole at a randomly chosen one. The fixed dipoles were positioned in four configurations using the same stereotactical coordinates as in [16]: two superficial dipoles (distance from the origin $>$ 6.5 cm), two deep ones (distance from the origin $<$ 6 cm), or two close together (relative distance $<$ 5 cm) or far apart (relative distance $>$ 8 cm) (table 1). We further assumed that the dipoles are oriented perpendicularly to the cortical surface. The 500 noise sources were randomly distributed within the brain mesh excluding the three active dipole locations.

Table 1. Fixed dipole coordinates within the subsampled New York head model, for each configuration. The first table lists the positions of the first fixed dipole, the second table those of the second fixed dipole.

Close superficial	(33.92, 59.40, −3.27)
Far superficial	(25.01, −93.32, 7.55)
Close deep	(−2.78, −2.56, 4.47)
Far deep	(58.19, −32.33, 33.18)
Close superficial	(40.87, 59.98, 5.69)
Far superficial	(40.87, 59.98, 5.69)
Close deep	(−5.88, −34.51, 34.22)
Far deep	(−30.73, 2.01, −6.17)

Once we activate the three brain sources and the 500 noise sources, we can project their activities to the 108 EEG electrodes. Once projected, the scalp activity of the brain source $x^{active}(t)$ and the scalp activity of the noise sources $x^{noise}(t)$ are summed. This is done with a coefficient λ, set to 0.5. $\|x^{active}(t)\|_F$ and $\|x^{noise}(t)\|_F$ are the Frobenius norms of the multivariate time series $x^{active}(t)$ and $x^{noise}(t)$, respectively, i.e. the square root of the squared activity summed spatially and temporally. The Frobenius norm is used to scale the brain source and noise activities:

$$x^{brain}(t) = \lambda \times \frac{x^{active}(t)}{\|x^{active}(t)\|_F} + (1-\lambda) \times \frac{x^{noise}(t)}{\|x^{noise}(t)\|_F}. \tag{2}$$

The signal definition, seen in equation (2), implies that deep active dipoles can have the same strength as shallow ones. This could make the discovery of deep active dipoles more difficult but, on the other hand, as the signal to noise ratio (SNR) is computed not per source, but for all three dipoles simultaneously, this imbalance is mitigated. Finally, in order to simulate the measurement noise, spatially and temporally uncorrelated signals, Gaussian-distributed uncorrelated white noise x^{m_noise} is added to $x^{brain}(t)$ with a predefined Mean Squared Amplitude ratio of 0.9. The overall pseudo-EEG data is defined by the following equation:

$$x^{measured}(t) = 0.9 \times \frac{x^{brain}(t)}{\|x^{brain}(t)\|_F} + 0.1 \times \frac{x^{m\_noise}(t)}{\|x^{m\_noise}(t)\|_F}. \tag{3}$$

2.2. Real-world EEG datasets

Two EEG datasets are considered as real-world cases. The first set was recorded from a 32 year old healthy

male under resting state (eyes closed) condition using 128 channels of a MEDICID 5 system operating at 200 Hz sampling rate. It is part of the Cuban Brain Mapping Project [23], created in 2005 with the aim to obtain brain atlases of the Cuban population.

The second dataset consists of 118 channel EEG activity recorded from five healthy subjects [24] during a visually-cued multi-class motor imagery task (imagined left hand, right hand and right foot movements) at 100 Hz sampling rate, available under Dataset IVa of the BCI Competition III (www.bbci.de/competition/iii/desc_IVa.html). We consider the case of imagined right hand movement of subject ay. Prior to source localization, we apply a fourth order zero-phase Butterworth filter to extract alpha band activity (8–12 Hz) and cut out epochs from 0.5 s after the start cue until 0.5 s before stop cue to avoid contamination by the visual cue and early stopping, respectively. We determine within each epoch the degree of event-related desynchronization (ERD) or -synchronization (ERS) with respect to baseline activity. Baseline activity epochs are taken from the rest periods between trials starting 0.3 s after the onset of a rest period until 0.3 s before it ends. In total, there are 18 epochs (18 epochs for baseline activity and 18 epochs for imagined movement activity). These epochs are then concatenated.

Similar to performed movement, we expect imagined movement activity in motor areas including the primary motor cortex (M1) and supplementary motor area (SMA), and ventral and dorsal parts of the premotor cortex, although activation of M1 during imagined movement is weaker compared to that during performed movement [25]. Activity is expected in areas that relate to action planning, such as the dorsolateral prefrontal cortex, ventrolateral prefrontal cortex (VLPFC) (which anatomically corresponds to the inferior frontal cortex, IFC), and posterior parietal cortex (PPC) including the angular gyrus [26]. However, given the rather simple motor task, we expect less involvement of the medial prefrontal cortex (MPFC), such as SMA. As far as we know, no source localization results have been published for this dataset.

For both datasets, we have used the Head Model of the Cuban Brain Mapping Project [23]. It contains mesh information in terms of vertex coordinates and triangles or faces. We introduced 6003 homogeneously distributed electric equivalent dipoles.

3. Results

3.1. Simulated EEG dataset results

The accuracy with which the active source signals are estimated with LCMV, eLORETA and Sparse eLORETA, for different source configurations and noise types, are summarized in tables 2 and 3. The tables list the average correlation coefficients between the reconstructed dipoles and ground truth signals (obtained from equation (1)), averaged over the 1000 randomly chosen positions of the third source with the other two sources kept fixed.

Table 2. Average correlation coefficients of the Far Superficial dipole configuration given three noise models. PN = (incoherent) pink noise, CPN = coherent pink noise, CAN = autoregressive pink noise.

	LCMV	eLORETA	Sparse eLORETA
PN	0.9867	0.9537	0.9234
CPN	0.9119	0.9329	0.9241
CAN	0.9807	0.9479	0.9020

Table 3. Average correlation coefficients in the presence of (uncoherent) pink noise for different dipole configurations. FS = far superficial, CS = close superficial, FD = far deep, CD = close deep.

	LCMV	eLORETA	Sparse eLORETA
FS	0.9867	0.9537	0.9234
CS	0.9707	0.9750	0.9023
FD	0.9204	0.8833	0.7721
CD	0.8383	0.6297	0.5991

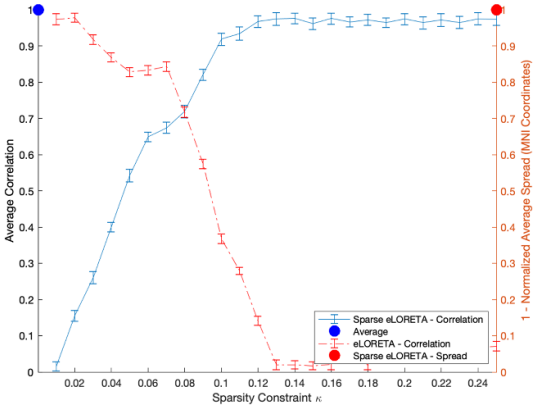

Figure 1. Average correlation (blue line, left scale) and complement of the average normalized spatial spread (red line, right scale), normalized with respect to eLORETA's outcome (red ball), plotted as a function of the sparsity constraint κ. The blue ball denotes eLORETA's average correlation. The Pareto-optimal solution is at the intersection ($\kappa \approx 0.08$).

A first observation is that LCMV yields the best correlation coefficients, which is counter-intuitive since in real-world cases we expect eLORETA to outperform LCMV. eLORETA does not provide an exact localization due to the presence of noise models in the simulation. There is a slight decrease in correlation obtained with Sparse eLORETA compared to eLORETA. This is a side effect of our source space reduction procedure: we found that, by increasing the sparsity constraint κ, the correlation coefficient increases and from $\kappa = 0.12$ onwards coincides with eLORETA's (blue line in figure 1). However, κ also affects source localization accuracy as reported next.

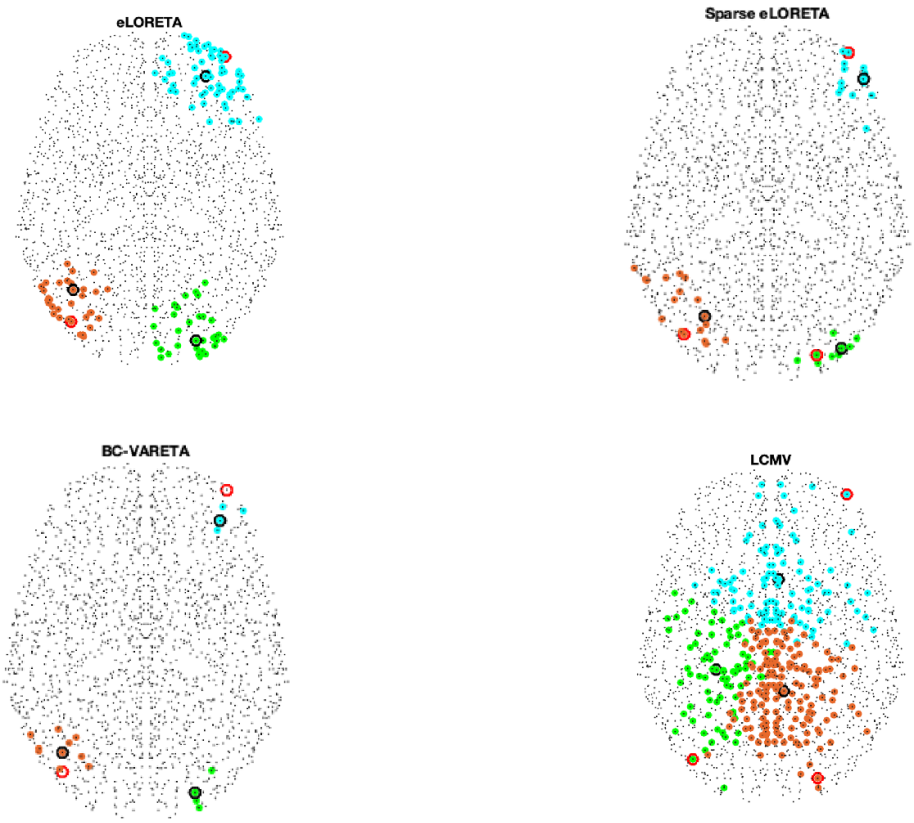

Figure 2. Distribution of reconstructed dipoles (brain top views, nose pointing upwards) of the simulation study, (uncoherent) pink noise case. Shown are the results obtained using eLORETA, Sparse eLORETA, BC-VARETA and LCMV given the Far Superficial configuration. with the three medoid clusters marked in orange, green and cyan. Medoids centers are indicated with black filled circles, the ground truth dipoles with red filled circles. In the bottom right cluster (green for eLORETA), the medoid center overlaps with the ground truth dipole.

Figure 2 shows the source distributions obtained with eLORETA, BC-VARETA and Sparse eLORETA for (uncoherent) pink noise in the case of the Far Superficial dipole configuration. We observe that, despite LCMV's higher signal correlation (table 3), the reconstructed dipoles are dispersed without showing any evidence of clustering around the true ones. With eLORETA, the found dipoles are clustered around the true ones but still less than with BC-VARETA. Compared to eLORETA, we clearly see the effect of source space reduction: the reconstructed dipoles are grouped and close to the true ones. Results are analogous for other configurations given the quantitative results listed in table 3.

We also assessed the computational effort (i.e. elapsed time) required by the source localization methods (table 4). We observe that BC-VARETA takes much more time compared to the other methods. Although the SSBL step is time-consuming, Sparse eLORETA requires only slightly more time than eLORETA, a beneficial side effect of reducing source space.

Table 4. Average elapsed time (in seconds) of LCMV, eLORETA, Sparse eLORETA and BC-VARETA, determined using Matlab's tic toc function, averaged over 1000 runs.

	Timing
LCMV	0.0286
eLORETA	0.3567
Sparse eLORETA	1.0459
BC-VARETA	68.0538

3.2. Clustering reconstructed dipoles

The results reveal major differences in the spatial distributions of reconstructed sources returned by the different methods. In order to quantify these distributions, we applied medoid clustering [27] to partition the reconstructed sources into clusters, extract their centers ('medoids') and determine their 'spatial spread' defined as the average distance of the cluster members to their corresponding medoids. Note that medoid clustering assumes the number of clusters is known. We have applied the method to cluster the eLORETA, Sparse eLORETA, BC-VARETA and LCMV reconstructed source distributions. Figure 2 shows the result of the Far Superficial configuration in the presence of (uncoherent) pink noise. The Supplementary Materials section shows the implementation details for the mediod clustering.

Table 5 lists, for all dipole configurations considered, the average spatial spread, averaged over the three clusters. The first point to note is the widespread nature of the LCMV dipole solutions.

Table 5. Average spatial spread in the presence of (uncoherent) pink noise for different dipole configurations. FS = far superficial, CS = close superficial, FD = far deep, CD = close deep.

	LCMV	eLORETA	Sparse eLORETA	BC-VA RETA
FS	1196.7	199.47	113.16	113.89
CS	1165.6	335.71	57.61	4.41
FD	4473.1	9239.1	175.34	4.48
CD	3001.4	907.3	242	65.4

Table 6. Average distance between medoids and true sources in the presence of (uncoherent) pink noise for different dipole configurations. FS = far superficial, CS = close superficial, FD = far deep, CD = close deep.

	LCMV	eLORETA	Sparse eLORETA	BC-VARETA
FS	68.71	11.88	11.64	11.43
CS	60.14	15.23	12.76	8.61
FD	19.24	41.46	37.07	35.28
CD	45.95	42.93	42.93	45.39

We also observe that Sparse eLORETA consistently yields smaller cluster sizes compared to eLORETA yet slightly larger than BC-VARETA's. Table 6 shows the average distance between the cluster elements and the true sources. We observe that the performances of eLORETA, Sparse eLORETA and BC-VARETA are very similar. In line with the (average) spatial spread, LCMV tends to generate larger distances between medoids and true sources.

When plotting the complement of the spatial spread against the sparsity constraint κ (red line in figure 1), we see that the larger κ, the smaller the complement of the spatial spread, or the larger the spatial spread, until about $\kappa = 0.12$ after which it levels off. When comparing this with the average correlation (blue line), we observe that $\kappa = 0.12$ is also the point where it levels off and, in addition, where there is no significant difference between eLORETA (blue ball) and sparse eLORETA. This is likely due to the fact that, for higher κ's, the sparsity constraint only prunes irrelevant dipoles. Hence, there seems to be a tradeoff between signal correlation and spatial spread or, in other words, between temporal and spatial accuracy. When normalizing the two metrics, thus, time series correlation and normalized complement of the average spread, they can be plotted on the same [0,1] scale. We take κ at the intersection as in this way we cannot improve on one metric without worsening on the other (Pareto optimum).

3.3. Real-world EEG dataset results

Figure 3 shows 300 source localizations (aka bootstrapping) for the first real world case (Cuban EEG dataset) as obtained with eLORETA, Sparse eLORETA, BC-VARETA and LCMV. We observe that the sources reconstructed with LCMV are more spread out. The most active sources (darker color) in Sparse eLORETA are in areas where BC-VARETA and Sparse eLORETA also locate their sources. Compared to eLORETA, we clearly observe the effect of source space reduction in Sparse eLORETA. There are no notable differences between Sparse eLORETA and BC-VARETA's results. For κ we took 0.08, as in the simulation setup, also given that it does not seem to prune sources from relevant regions while maintaining the desired sparsity.

Figure 4 shows 300 source localizations for the second real world set (subject ay, imagined right hand movement) obtained with eLORETA, Sparse eLORETA, BC-VARETA. We observe that BC-VARETA does not perform well on this dataset as only a few sources are localized. LCMV also has issues as practically all sources are labeled as ERS and distributed uniformly without showing any identifiable clusters (not shown). eLORETA, and Sparse eLORETA in particular, are able to reconstruct sources in the motor cortex. For the latter two methods, we observe alpha band ERS activity (i.e. deactivation) in DLFPC (bilaterally), angular gyrus (ANG, contralaterally), and superior PPC (contralateral in Sparse eLORETA, bilateral in eLORETA), and ERD activity in M1 (contralateral in Sparse eLORETA, bilateral in eLORETA albeit less ipsilateral) and VLPFC (bilaterally), but no activity in MPFC (SMA) (see figure 5 for a labeling of the ERD/ERS active brain regions for Sparse eLORETA). Time-locked deactivations during goal-directed behavior in these regions have been observed in several studies and are considered part of a default mode network [28]. We observe that eLORETA exhibits close-to-zero ERD/ERS activities near the temporal pole and ERS activities along the inferior temporal gyrus bilaterally but all these have been pruned by SSBL in Sparse eLORETA.

4. Discussion

Localizing active brain regions from non-invasive recordings is a challenging problem as there is simply not enough information to arrive at a unique, stable solution without relying on additional assumptions. This has led to a gamut of localization methods at least in the case of scalp EEG, most importantly LCMV and several members of the LORETA family, such as eLORETA and sLORETA. Some methods in addition provide estimates of the functional connectivity between reconstructed sources as in BC-VARETA [11], where sources and connectivities are computed simultaneously. A more common approach is to first anatomically define well-separated ROIs and then to estimate their functional connectivity [29] with recent studies pushing for adaptive cortical meshes to improve sensitivity and distinguishability of the ROIs [30]. Others have focused on demixing the recorded signals, instead of localizing them, after which their

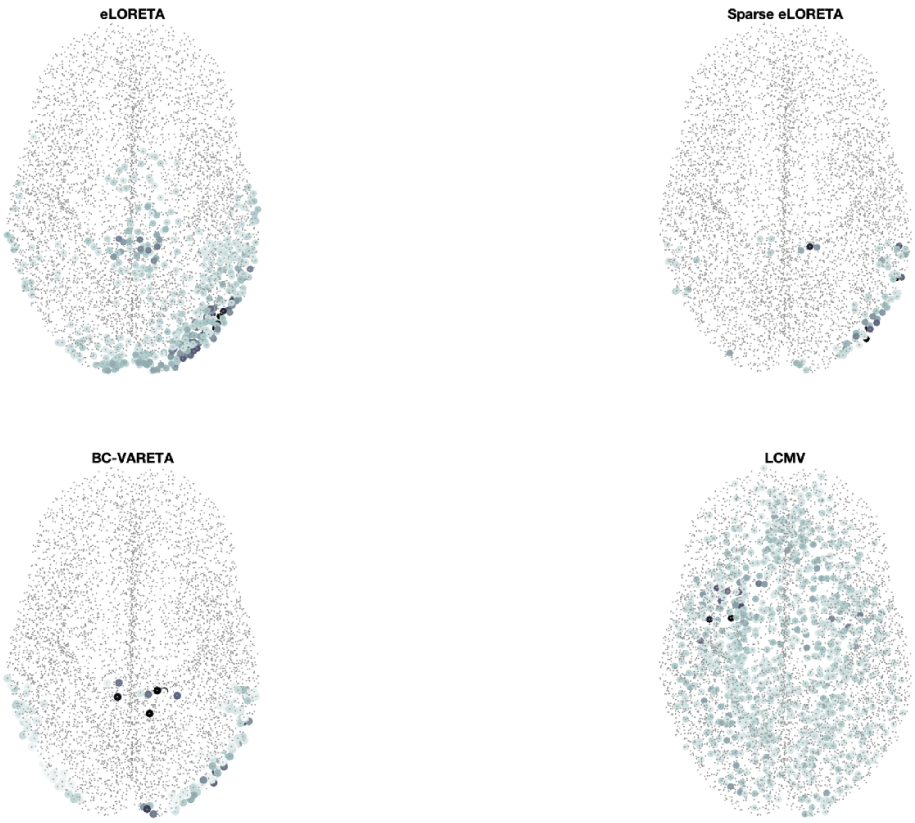

Figure 3. Distribution of reconstructed dipoles (brain top views, nose pointing upwards) of the Cuban EEG dataset using eLORETA, Sparse eLORETA, BC-VARETA and LCMV. The grey nodes in the background are mesh nodes. The filed circles are active dipoles with a darker color indicating more activity.

connectivity is estimated, e.g. using the method proposed in [16].

We proposed Sparse eLORETA for source localization. It relies on SSBL to reduce source space dimensionality. Combined with eLORETA, it yields a much sparser solution, while retaining the advantages of the latter (i.e. a low localization error and few ghost sources). The much sparser solution is expected to facilitate subsequent connectivity analysis.

Similar to other simulation studies, we modulated the depth of the sources, their reciprocal distance, and the SNR (brain noise level). We adopted the framework of Anzolin *et al* to develop a realistic simulation of brain activity to assess Sparse eLORETA's performance for multiple configurations of the three active sources [16]. As expected, superficial sources that are far apart were much easier to localize, compared to deep sources close together, as seen in table 3. Compared to LCMV and eLORETA, Sparse eLORETA performed better, gaining more sparsity while sacrificing only minimally on accuracy, showing the benefit of SSBL. We mainly looked at how accurately these state-of-the-art techniques could locate three active sources. We had instances where the sources were not accurately located but their signals correlated well with the true ones. Moreover, we could show that, despite the SSBL step, the computational effort of our method is similar to eLORETA's since our solution is more sparse. However, BC-VARETA requires significantly more effort compared to eLORETA and our method. LCMV is the fastest to compute, although the results were disappointing.

To quantify source localization accuracy of our simulation results, we performed a medoid clustering analysis on the reconstructed source distributions. We extracted two metrics from this analysis, the distance of the medoids to the true sources and the spatial spread around the former, and showed that our method performed better than eLORETA. However, our results also showed that BC-VARETA still holds an edge over the methodology we propose as seen in table 5. Since BC-VARETA localizes sources but does not reconstruct their time series, a comparison with BC-VARETA cannot account for the fidelity with which these time series are reconstructed. In that sense, our method is a compromise between source sparsity and signal fidelity, two aspects of the source reconstruction problem. Medoid clustering could also be used to reduce source space similar to Wang *et al* [31], albeit using different clustering approaches. We believe that source screening, which we proposed for eLORETA, and (data-driven) clustering, or a combination thereof, have their merits in generating sparse solutions.

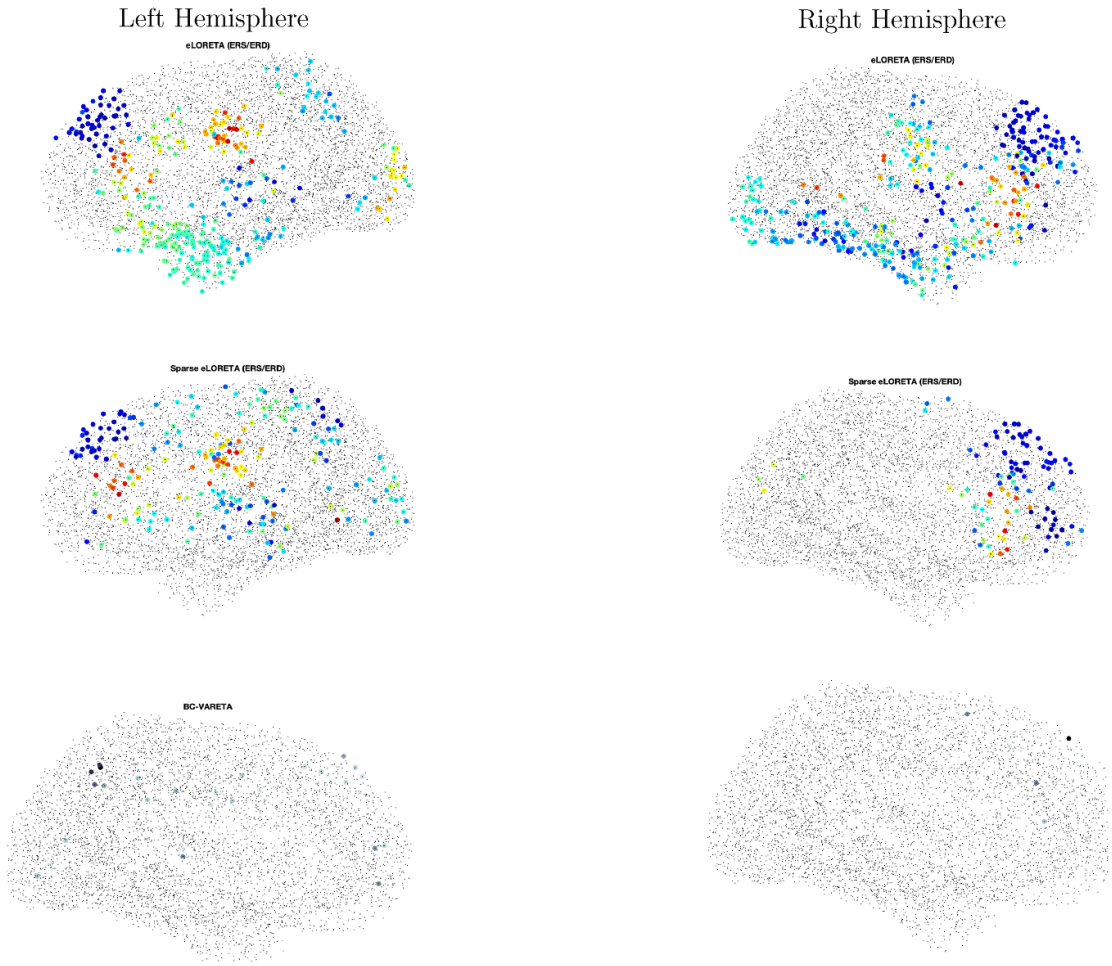

Figure 4. ERD/ERS post-cue imagined right hand movement activation of 300 sources (aka bootstrapping) obtained with eLORETA, Sparse eLORETA and BC-VARETA, mapped on left and right lateral brain views (arranged column-wise). The grey nodes in the background are mesh nodes. Blue dots correspond to ERS, yellow/red to ERD and green to neither. For BC-VARETA dipoles are grey dots as no ERS/ERS assessment is possible.

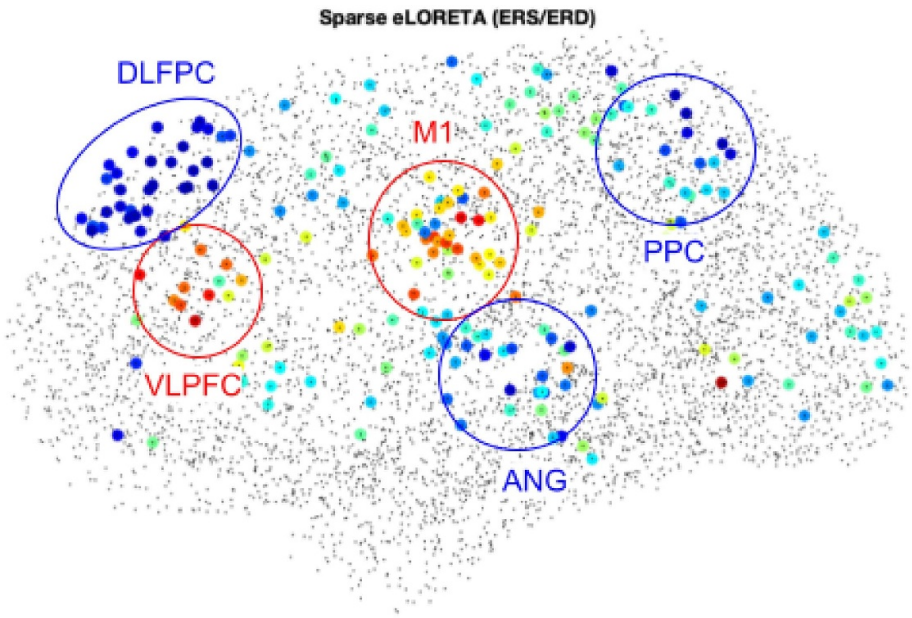

Figure 5. Contralateral view of the Sparse eLORETA solution with clusters labeled by the corresponding brain regions.

In order to show the applicability in practice, we considered a resting state and a cued imagined movement EEG dataset. For the first case, except for LCMV, we observe that the source estimates of eLORETA, Sparse eLORETA and BC-VARETA are located in similar regions but with a smaller spread for the latter two, as expected (figure 3). However, while in our simulation study BC-VARETA exhibited a smaller spread, this was no longer the case with the real-world dataset.

As to the results of the second case (figure 4), we observe that, compared to Sparse eLORETA, eLORETA has more ipsilateral clusters of sources but at the same time more spurious ones. Again, BC-VARETA performs much worse by consistently locating sources in medial central and temporal cortices (with almost zero spread).

Finally, for cases where ROIs are not predefined, such as with the real world datasets we have also considered, our method could be used to perform source space reduction prior to applying source localization. In the future, it would be interesting to expand on our work by integrating Sparse eLORETA with a method for selecting ROIs. In this way, we could arrive at a completely data-driven method. Since source localization results returned by Sparse eLORETA are sparse, it would be interesting to see how clustering algorithms based on pairwise connectivities could assist in charting ROIs.

5. Conclusion

We show that Sparse eLORETA offers the possibility for eLORETA to trade temporal for spatial accuracy.

Funding

A F was supported by the Belgian Fund for Scientific Research—Flanders (FWO 1157019N). A d B participated as part of an internship. M M V H is supported by research grants received from the European Union's Horizon 2020 research and innovation programme onder Grant Agreement No. 857 375, the Financing Program (PFV/10/008) and the special research fund of the KU Leuven (C24/18/098), the Belgian Fund for Scientific Research—Flanders (G088314N, G0A0914N, G0A4118N), the Interuniversity Attraction Poles Programme—Belgian Science Policy (IUAP P7/11), and the Hercules Foundation (AKUL 043).

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

ORCID iDs

A Faes https://orcid.org/0000-0002-1637-255X
M M Van Hulle https://orcid.org/0000-0003-1060-7044

References

[1] Niedermeyer E and Lopes da Silva F H 2005 *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields* (Baltimore, MD: Williams & Wilkins)
[2] Cohen M X 2014 *Analyzing Neural Time Series Data: Theory and Practice* (Cambridge, MA: MIT Press)
[3] Michel C M and Brunet D 2019 EEG source imaging: a practical review of the analysis steps *Front. Neurol.* **10** 325
[4] Schoffelen J-M and Gross J 2009 Source connectivity analysis with MEG and EEG *Human Brain Mapp.* **30** 1857–65
[5] Roberto Domingo P-M 1999 Review of methods for solving the EEG inverse problem *Int. J. Bioelectromagn.* **1** 75–86
[6] Grech R, Cassar T, Muscat J, Camilleri K P, Fabri S G, Zervakis M, Xanthopoulos P, Sakkalis V and Vanrumste B 2008 Review on solving the inverse problem in EEG source analysis *J. Neuroeng. Rehabil.* **5** 25
[7] Van Veen B D, Van Drongelen W, Yuchtman M and Suzuki A 1997 Localization of brain electrical activity via linearly constrained minimum variance spatial filtering *IEEE Trans. Biomed. Eng.* **44** 867–80
[8] Whittingstall K, Stroink G, Gates L, Connolly J F and Finley A 2003 Effects of dipole position, orientation and noise on the accuracy of EEG source localization *Biomed. Eng. Online* **2** 1–5
[9] Pascual-Marqui R D 2007 Discrete, 3D distributed, linear imaging methods of electric neuronal activity. Part 1: exact, zero error localization (arXiv:0710.3341)
[10] Pascual-Marqui R D *et al* 2011 Assessing interactions in the brain with exact low-resolution electromagnetic tomography *Phil. Trans. R. Soc.* A **369** 3768–84
[11] Gonzalez-Moreira E, Paz-Linares D, Martinez-Montes E and Valdes-Sosa P A 2018 Third generation MEEG source connectivity analysis toolbox (BC-VARETA 1.0) and validation benchmark (arXiv:1810.11212)
[12] Paz-Linares D, Gonzalez-Moreira E and Duru A 2019 Measuring and suppressing m/EEG connectivity leakage: the hidden Gaussian graphical state-model (HIGGS) *Organization for Human Brain Mapping* **2019**
[13] Bringas Vega M L *et al* 2019 An age-adjusted EEG source classifier accurately detects school-aged barbadian children that had protein energy malnutrition in the first year of life *Front. Neurosci.* **13** 1222
[14] Rykaczewski K, Nikadon J, Duch Włodzisław and Piotrowski T 2021 supFunSim: spatial filtering toolbox for EEG *Neuroinformatics* **19** 107–25
[15] Paz-Linares D, Vega-Hernandez M, Rojas-Lopez P A, Valdes-Hernandez P A, Martínez-Montes E and Valdes-Sosa P A 2017 Spatio temporal EEG source imaging with the hierarchical Bayesian elastic net and elitist lasso models *Front. Neurosci.* **11** 635
[16] Anzolin A, Presti P, Frederik V D S, Astolfi L, Haufe S and Marinazzo D 2019 Quantifying the effect of demixing approaches on directed connectivity estimated between reconstructed EEG sources *Brain Topography* **32** 655–74
[17] Haufe S and Ewald A 2019 A simulation framework for benchmarking EEG-based brain connectivity estimation methodologies *Brain Topography* **32** 625–42
[18] Omidvarnia A, Mesbah M, O'Toole J M, Colditz P and Boashash B 2011 Analysis of the time-varying cortical neural connectivity in the newborn EEG: a time-frequency approach *Int. Workshop on Systems, Signal Processing and their Applications, WOSSPA* (IEEE) pp 179–82

[19] Haufe S, Nikulin V V, Müller K-R and Nolte G 2013 A critical assessment of connectivity measures for EEG data: a simulation study *Neuroimage* **64** 120–33
[20] Habets Eel A P, Cohen I and Gannot S 2008 Generating nonstationary multisensor signals under a spatial coherence constraint *J. Acoust. Soc. Am.* **124** 2911–17
[21] Haufe S and Ewald A 2016 Berlin brain connectivity benchmark and data analysis challenge (Available at: http://bbci.de/supplementary/EEGconnectivity/BBCB.html)
[22] Huang Y, Parra L C and Haufe S 2016 The new york head-a precise standardized volume conductor model for EEG source localization and TES targeting *NeuroImage* **140** 150–62
[23] Hernandez-Gonzalez G *et al* 2011 Multimodal quantitative neuroimaging databases and methods: the cuban human brain mapping project *Clin. EEG Neurosci.* **42** 149–59
[24] Dornhege G, Blankertz B, Curio G and Muller K-R 2004 Boosting bit rates in noninvasive EEG single-trial classifications by feature combination and multiclass paradigms *IEEE Trans. Biomed. Eng.* **51** 993–1002
[25] Munzert Jorn, Lorey B and Zentgraf K 2009 Cognitive motor processes: the role of motor imagery in the study of motor representations *Brain Res. Rev.* **60** 306–26
[26] Deiber M-P, Ibanez V, Honda M, Sadato N, Raman R and Hallett M 1998 Cerebral processes related to visuomotor imagery and generation of simple finger movements studied with positron emission tomography *Neuroimage* **7** 73–85
[27] Kaufman L and Rousseeuw P J 2009 *Finding Groups in Data: An Introduction to Cluster Analysis* vol 344 (New York: Wiley)
[28] Mantini D and Vanduffel W 2013 Emerging roles of the brain's default network *Neuroscientist* **19** 76–87
[29] Fahimi Hnazaee M, Khachatryan E, Chehrazad S, Kotarcic A, De Letter M and Van Hulle M M 2020 Overlapping connectivity patterns during semantic processing of abstract and concrete words revealed with multivariate Granger Causality analysis *Sci. Rep.* **10** 1–19
[30] Farahibozorg S-R, Henson R N and Hauk O 2018 Adaptive cortical parcellations for source reconstructed EEG/MEG connectomes *NeuroImage* **169** 23–45
[31] Wang S H, Lobier M, Siebenhühner F, Puoliväli T, Palva S and Palva J M 2018 Hyperedge bundling: a practical solution to spurious interactions in MEG/EEG source connectivity analyses *NeuroImage* **173** 610–22