Statistics and Its Interface Volume 0 (202x) 1–13

Depth-Invariant Beamforming for Functional Connectivity with MEG data *

Jian Zhang

The conventional beamformers that reconstruct the cerebral origin of brain activity measured outside the head via electro- and magnetoencephalography (EEG/MEG) suffer from depth bias and smearing of nearby sources. Here, to meet these methodological challenges, we propose a depth-invariant and forward beamformer for magnetoencephalography (MEG) data. Based on the new proposal, we further develop a two-step approach for inferring functional connectivity in the brain. The proposed methodology is invariant with respect to source depths in the brain. It nulls smearing of nearby sources and allows for time-varying source orientations. We illustrate the new approach with MEG data derived from a face-perception experiment, revealing patterns of functional connectivity for face perception. We identify a set of brain regions where their responses and connectivity are significantly varying when stimuli alter between faces and scrambled faces. By simulation studies, we show that the proposed forward beamformer can outperform the forward methods based on conventional beamformers in terms of localization bias.

AMS 2000 subject classifications: Primary 62H35.
Keywords and phrases: MEG neuroimaging, depth-invariant beamforming, functional network, source localization, reconstruction.

1. INTRODUCTION

As cognitive functions arise from dynamic communication between brain regions, there is an increasing interest in the detection of the underlying functional networks constituted by these regions (Ishai, 2008; Brooks et al., 2011; Friston, 2011). Functional connectivity, defined as temporal dependencies among neurophysiological events originating from different brain regions, can be inferred from dynamic signals which are measured by use of non-invasive devices such as functional magnetic resonance imaging (fMRI) and magnetoencephalography (MEG). fMRI measures brain activity by detecting changes associated with blood flow whereas MEG measures magnetic signals outside the brain, which originate from neural firing. Note that the accuracy of estimated functional network strongly depends on the temporal resolution of the measured events (e.g., Drakesmith et. al, 2013). It is widely acknowledged that the haemodynamic measurement-based fMRI can resolve detailed spatial patterns of activity, but has notoriously poor temporal resolution. As the magnitude of the measured fields is proportional to that of the neural currents, using MEG, one is able to perform a real-time tracking of brain activities with an millisecond temporal resolution, the same time scale as that of neural interactions. However, MEG has a relatively low spatial precision. Therefore, MEG is a useful tool to measure functional brain connectivity when one is able to improve its spatial resolution.

The electrical currents derived from the neural firing are often modeled by neural dipoles which are described by their locations, orientations and amplitudes (called time-courses). See Zhang and Su (2015) and Yao at al. (2018) for an introduction. In light of this, a distributed model with thousands of current dipoles is commonly proposed to infer the neuronal generators of MEG data. Dipole locations (i.e., voxels) are created by discretizing the dipole space in the brain, where the number of these grid points can be much larger than the number of available sensors. Therefore, searching for a small subset of dipole sources among a large number of candidates poses a challenge to modern statistics.

In this paper we focus on the study of cognitive functions related to face perception, a most developed visual skill in human, by using MEG measurements. Face-perception is many ways a microcosm of object recognition; and the solution to face-perception will give insights into the general problem of object recognition (Tsao and Livingstone, 2008). The cognitive development of face perception is linked to developing social interaction skills and language. For example, infants prefer to imitate facial expressions at a very early age. Adults perceive the gender, expression, age and mood by looking at faces. Processing information gleaned from faces requires the integration of activity across a network of cortical regions. This suggests that face perception is mediated by a specialized neural network system (Ishai, 2008). However, there have been few studies in which researchers attempt to elucidate the features of such a system. With fMRI data, Zhen et al. (2013) revealed a functional network of various regions which showed differential responses to facial stimuli over non-facial stimuli. The findings from MEG data are also very limited although they

* We thank Professor Richard Henson, MRC Cognition and Brain Sciences Unit, Cambridge, for sharing the MEG neuroimaging data with us.

have a better temporal resolution. For example, Henson et al. (2011) designed a series of face-perception MEG experiments, identifying only a region called fusiform face area (FFA) which responded differently to face over scrambled face. But they missed other regions which were found in fMRI studies. It is unclear whether there are networks underlying the observed MEG data, similar to the fMRI setting. To fill the gap, Zhang et al. (2014) and Zhang and Liu (2015) made some efforts by a covariance-thresholding technique, revealing some more but weakly active regions. Unfortunately, source smearing often jeopardized these analyses in general (Sekihara et al., 2010). Here, we propose a novel forward beamforming to tackle the problem of source cross-talk and thus to improve the data analysis of the above face-perception experiment. By cross-talk we mean that signals generated at one dipole location can leak into estimated activity at spatially separate locations.

The choice of an appropriate beamforming method heavily influences the reconstructed brain activity, as well as the subsequent connectivity analysis as the above inverse problem is ill-posed. Distributed approaches such as minimum-norm estimate (MNE) and adaptive beamformers are commonly employed to address the source reconstruction problem (Hämäläinen et al., 1993; Veen et al., 1997; Robinson and Vrba, 1998; Sekihara et al., 2010; Zhang et al., 2014; Zhang and Liu, 2015; Yao et al., 2020). In these methods, one scans through dipole locations inside the brain by use of spatial filters and assess their contributions to the observed magnetic fields in terms of neuronal activity index. Conventional beamforming approaches such as synthetic aperture magnetometry (SAM) assume that the dipole orientations keep fixed in time and henceforth can be obtained through an extra optimization step. The constant orientation assumption can be relaxed by designing separate beamformers for each individual principal dipole orientation which are called linearly constrained minimum variance (LCMV) beamformers. Both SAM and LCMV have the advantage over traditional multiple dipole fittings that they do not require knowing the number of dipoles in the model in advance. Furthermore, the MNE has low spatial resolution for focal point-like neuronal sources, while SAM and LCMV solutions can model the point sources down to the size of the mesh grid (Huang et al., 2004). The SAM and LCMV also provide a convenient framework to suppress source cross-talk by imposing certain constrains in optimization. So, we opt for SAM and LCMV as building blocks in our study. The SAM and LCMV procedures can be implemented in two steps: In the first step, at each dipole point, one searches for the optimal projections of the sensor data to the three principal directions by minimizing the trace of the sample covariance of the projected data subject to a unit matrix gain constraint. Through the minimization, we reduce the contributions of white noises and source cross-talks to powers and obtain an optimal power estimator at the grid point. Similarly, using the baseline data, we obtain an optimal baseline power estimator. In the second step, we calculate a neuronal activity index (NAI) by contrasting the above two power matrices at each grid point. Plotting these index values against the grid points, we can create an NAI map over a given time window. The map presents the distribution of dipolar currents which underlies the recorded sensor data.

Appropriately summarizing neuronal activity is fundamentally important in the above implementation. Veen et al. (1997) proposed an output signal-to-noise-ratio based NAI (i.e., the ratio between the traces of the signal and noise power matrices) for this purpose. NAI takes the value of 1 when the above two power matrices are equal. The SAM can be implemented in a similar way. The SAM and LCMV beamformers have been shown to have good performance in various scenarios (e.g., Sekihara et al., 2010). Despite this, there are a few important issues remained to address. Firstly, these beamformers can fail to detect dipole sources where the signal and noise power matrices are not equal but NAI takes the value of 1. This is because the NAI is not a bijective operator as it is based on the ratio between the traces (or the the maximum relative eigenvalue) of the signal and noise matrices. Secondly, these beamformers can be biased as they are not dipole depth-invariant. If one principal orientation generates a weaker lead field than the other two principal orientations do, then both the traces of the signal power matrix and the noise power matrix are dominated by the above unreliable weak dipole orientation. This will increase the false discovery rate of beamforming (Huang et al., 2004). Finally, SAM beamformer relies on a potentially invalid assumption that dipole orientations is time-invariant.

Here, to address these issues, we first propose a novel vector-beamformer based on the Bregman-divergence index (Davis et al., 2007). We show that the proposed beamformer is dipole-depth invariant and that it allows for time-varying dipole orientations. Based on this new technique, we are able to implement a forward beamforming by nulling the previous identified NAI peaks in each iteration. To define a termination rule, we divide dipole candidates into two groups, one with high NAI values and the other with low NAI values, by minimizing the within-group variability. Using the low NAI values, we decide whether the NAI map has been whitened and when forward beamforming should be terminated. Previous findings have suggested that temporal correlation of neuronal time-courses may be an indication of communication and information flow between cortical neurons (Chan et. al, 2015). In light of this, we then develop a method for constructing a functional network of contrast between stimuli based on correlations of dipole time-courses. In particular, we apply the proposed method on an MEG dataset derived from a face-perception experiment, revealing a novel functional network of contrast for face stimuli against scrambled face stimuli. The network suggests that the selected regions coordinate each other in response to the change of face stimuli in a hierarchical way: Face-perception at the top depends on the source clusters in the middle layer

and each cluster depends on neuronal regions at the bottom layer. This finding improves our understanding of the neuronal mechanism underpinning face perception. Finally, to assess the new procedure, we conduct simulation studies to compare the Bregman-divergence based procedure to the LCMV and SAM based procedures, suggesting that the former is most promising.

The remaining paper is organized as follows. The details of the proposed methodology and its theory are provided in Section 2. The applications of the proposed methods to the face-perception dataset and synthetic data are presented in Section 3. Conclusions are made in Section 4.

2. METHODOLOGY

To propose a new approach, we begin with a brief derivation of the classical LCMV beamformer followed by a discussion on its weakness. Then, we give the details of the new approach.

2.1 Dipole-depth dependance of LCMV

Suppose that we observe a vector of time series $\mathbf{Y}(t) = (Y_1(t), ..., Y_n(t))^T \in \mathbf{R}^n$, $t = t_j, 1 \le j \le J$ from n sensors. We also have the corresponding baseline data for the sensors. We consider a list of candidate source locations in the brain, $r_k, 1 \le k \le p$, each with an $n \times 3$ lead field matrix H_k (i.e., the unit output derived from Maxwell's equations). Let $\mathbf{m}_k(t)$ be an unknown 3×1 moment (time-course) at time t and location r_k with covariance matrix Σ_k (called power matrix), $\varepsilon(t)$ the white noises at the MEG sensors with covariance matrix $\sigma_0^2 I_n$, and I_n the $n \times n$ identity matrix. Then, the distributed model can be described by the following equations

$$\mathbf{Y}(t) = \sum_{k=1}^{p} H_k \mathbf{m}_k(t) + \varepsilon(t). \quad (1)$$

See Sekihara et al. (2010) for the details. The classic LCMV beamforming can be defined in two steps as follows.

Step 1. For each location r_k in the source space, we search for an optimal projection-matrix $\hat{W}_k$ and the corresponding power matrix $\hat{\Sigma}_k$ by minimizing the trace of the sample covariance of the projected data $W^T \mathbf{Y}(t_j)$, $1 \le j \le J$ with respect to the $n \times 3$ projection matrix W, subject to $W^T H_k = I_3$. By use of the Lagrange multiplier, we solve the above optimization problem, obtaining:

$$\begin{aligned} \hat{W}_k &= \hat{C}^{-1} H_k \left[H_k^T \hat{C}^{-1} H_k \right]^{-1}, \\ \hat{\Sigma}_k &= \hat{W}_k^T \hat{C} \hat{W}_k = \left[H_k^T \hat{C}^{-1} H_k \right]^{-1}, \end{aligned}$$

where $\hat{C}$ is a sensor covariance estimator. Similarly, based on the baseline data, we obtain the noise power matrix estimator $\hat{\sigma}_0^2 \left[H_k^T H_k \right]^{-1}$.

Step 2. For location r_k, calculate the neuronal activity index

$$\mathrm{NAI}_{1k} = \mathrm{NAI}_{1k}(\hat{C}, \hat{\sigma}_0^2, H_k) = tr(\hat{\Sigma}_k)/(\hat{\sigma}_0^2 tr([H_k^T H_k]^{-1})),$$

which is a contrast between the traces of the signal power matrix and the noise power matrix. We then plot the index values against the grid points, creating a neuronal activity map over a given time window. Finally, one estimates the true sources by thresholding the map.

However, the sensitivity of MEG sensors is not uniform to sources across different cortical regions and different orientations (Hillebrand and Barnes, 2002). In fact, it follows generally from Maxwell's equations that the lead fields of MEG sensors have a maximum at the border of the source space, closest to the sensors (Heller and van Hulsteyn, 1992). The classic LCMV solution is biased to those locations and orientations, preferring superficial source locations. This observation can be validated by the following numerical example. We constructed a spherical volume conductor with 10cm radius from the origin and with $n = 91$ sensors, created by using the software FieldTrip (Oostenveld et al., 2011). We then discretized the inside brain space into a 3D-grid of resolution 1 cm. This yielded a grid with $p = 2222$ grid points $r_k, k = 1, ..., p$ inside the brain. We calculated the magnetic fields emanating from unit inputs of electric dipole neuronal activities at these locations respectively, obtaining an $n \times 3$ lead field matrix for each grid point. Denote the lead field matrix at the k-th location by $H_k = (H_{k1}, H_{k2}, H_{k3})$, where H_{kj} is an $n \times 1$ column vector. For principal orientation $j = 1, 2, 3$, we calculated $(||r_k||, ||H_{kj}||), k = 1, ..., p$ and made a scatter plot in Figure 1. The plots show that in general $||H_{kj}||$ is relatively smaller when location r_k is close to the origin than when r_k is away from the origin. We also made pairwise plots between principal orientations. These plots demonstrate that $||H_{kj}||$ are varying across different orientations. There were a number of grid points with weak dipole orientations.

Then, the depth of a source at r_k will be inversely proportional to the total of squared unit outputs $||H_{kj}||$ which the sensors receive from that source. To remove the depth effects, we normalize the lead field vectors, obtaining the normalized lead field vector

$$\tilde{H}_{kj} = H_{kj}/||H_{kj}||.$$

Under the normalized lead field, the sensitivities of the dipoles inside the brain to the sensors are spatially homogenous. Furthermore, the power matrix estimators admit the following decompositions:

$$\begin{aligned} \hat{\sigma}_0^2 \left[H_k^T H_k \right]^{-1} &= \hat{\sigma}_0^2 D_k^{-1} \left[\tilde{H}_k^T \tilde{H}_k \right]^{-1} D_k^{-1}, \\ \left[H_k^T \hat{C}^{-1} H_k \right]^{-1} &= D_k^{-1} \left[\tilde{H}_k^T \hat{C}^{-1} \tilde{H}_k \right]^{-1} D_k^{-1} \end{aligned}$$

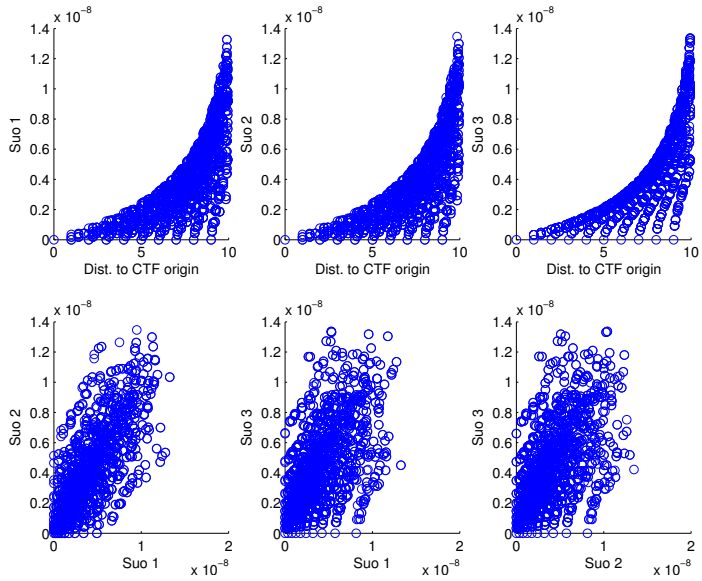

Figure 1. The first row displays the scatter plots of the total squared unit outputs at 2222 grid points along the principal orientation j against the corresponding Euclidean distances from dipole locations to the CTF origin. The plots from the left to the right are corresponding to $j = 1, 2, 3$ (i.e., x, y, z axes). In general, the closer to the origin, the smaller the unit output will be. The second row shows the pairwise scatter plots of the total squared unit outputs at 2222 grid points for the principal orientations j, $j = 1, 2, 3$. These plots demonstrate that the conditional number c_{jk} can be close to 0. Color will not be used in print.

with $D_k = \text{diag}(||H_{1k}||, ||H_{2k}||, ||H_{3k}||)$. For any $1 \le j \le 3$, define the conditional number $c_{jk} = \max_{i \neq j}\{||H_{jk}||/||H_{ik}||\}$ which is roughly equal to the maximum ratios of the jth orientation to other orientations. The scatter plots in the 2nd row of Figure 1 indicate that for some ks, the conditional number c_{jk} is close to zero because r_k is close to the origin. For these ks, the unit output H_{jk} along the principal dipole orientation j at r_k are much smaller than those along other principal orientations. In these cases, the sensors will not be sensitive to signals generated along these orientations.

The following proposition indicates that the index NAI_{1k} is not scale-invariant with respect to the lead field matrix H_k and that the index can be dominated by a weak dipole orientation.

Proposition 2.1. *There exists a diagonal matrix $S = diag(s_1, s_2, s_3)$ such that*

$$NAI_{1k}(\hat{C}, \hat{\sigma}_0^2, SH_k) \neq NAI_{1k}(\hat{C}, \hat{\sigma}_0^2, H_k).$$

Moreover, for any $1 \le j \le 3$, if $[\tilde{H}_k^T \tilde{H}_k]^{-1}$ and $[\tilde{H}_k^T \hat{C}^{-1} \tilde{H}_k]^{-1}$ are fixed but the condition number is close to zero, then NAI_{1k} is approximately determined by the jth principal orientation in the form

$$NAI_{1k} \approx \frac{u_{jj}}{v_{jj}},$$

where u_{jj} and v_{jj} are the jth diagonal entries of matrices $[\tilde{H}_k^T \hat{C}^{-1} \tilde{H}_k]^{-1}$ and $[\tilde{H}_k^T \tilde{H}_k]^{-1}$ respectively.

Note that a property similar to Proposition 2.1 holds for the SAM.

2.2 Bregman-divergence-based beamformer

The above proposition implies that the classical LCMV can have the bias to deep dipoles in the brain. To reduce the bias, we weight the estimated power matrix $\hat{\Sigma}_k = \hat{W}_k^T \hat{C} \hat{W}_k$ by the baseline power matrix estimator $\hat{\sigma}_0^2 \hat{W}_k^T \hat{W}_k$, obtaining a normalized power matrix estimator at r_k:

$$\begin{aligned} R_k &= \hat{W}_k^T \hat{C} \hat{W}_k (\hat{\sigma}_0^2 \hat{W}_k^T \hat{W}_k)^{-1} \\ &= \hat{\sigma}_0^{-2} \left[H_k^T \hat{C}^{-2} H_k\right]^{-1} \left[H_k^T \hat{C}^{-1} H_k\right]. \end{aligned}$$

Then, we define the normalized power at r_k by using the Bregman matrix divergence

$$\text{NAI}_{bk} = \text{NAI}_{bk}(\hat{C}, \hat{\sigma}_0^2, H_k) = tr(R_k) - \log(\det(R_k)) - 3.$$

Note that two power matrix estimators $[H_k^T \hat{C}^{-1} H_k]$ and $[H_k^T \hat{C}^{-2} H_k]$ are non-negative. So, there exist 3×3 matrix U and diagonal matrix $\Lambda = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$ satisfying

$$[H_k^T \hat{C}^{-1} H_k] U = [H_k^T \hat{C}^{-2} H_k] U \Lambda, \quad U^T [H_k^T \hat{C}^{-2} H_k] U = I_3,$$

where $\lambda_1 \ge \lambda_2 \ge \lambda_3 \ge 0$ are the relative eigenvalues of $[H_k^T \hat{C}^{-1} H_k]$ to $[H_k^T \hat{C}^{-2} H_k]$. Therefore, we have

$$\text{NAI}_{bk} = \sum_{j=1}^{3} \left(\hat{\sigma}_0^{-2} \lambda_j - \log(\hat{\sigma}_0^{-2} \lambda_j) - 1\right).$$

Let k_0 be the index value at which NAI_{bk} attains the maximum. We denote by r_0 the corresponding location estimate and by H_{k_0} the lead field matrix at r_0. We estimated the time-course at location r_0 along the orientation determined by the eigenvector u_{01} associated with the maximum eigenvalue λ_{01} of $\hat{W}_{k_0}^T \hat{C} \hat{W}_{k_0}$ relative to $\hat{\sigma}_0^2 \hat{W}_{k_0}^T \hat{W}_{k_0}$, which is equal to $u_{01}^T \hat{W}_{k_0}^T \mathbf{Y}$.

The next proposition says that the new index is scale-invariant.

Proposition 2.2. *For any scale transformation $\tilde{S}_k = H_k S$ on H_k with $S = diag(s_1, s_2, s_3)$, we have*

$$NAI_{bk}(\hat{C}, \hat{\sigma}_0^2, H_k) = NAI_{bk}(\hat{C}, \hat{\sigma}_0^2, \tilde{S}_k).$$

In particular, letting $S = D_k^{-1}$ defined in Section 2.1 and $\tilde{H}_k = H_k D_k^{-1}$, we have

$$NAI_{bk}(\hat{C}, \hat{\sigma}_0^2, H_k) = NAI_{bk}(\hat{C}, \hat{\sigma}_0^2, \tilde{H}_k).$$

2.3 Forward beamforming

We define an iterative Bregman forward beamformer called BBFB below:

- Step 0: Run the Bregman beamformer on the sensor data to initialize the procedure, finding r_0 and k_0 defined in the previous subsection.
- Step 1: Define and run an updated beamformer on the data by nulling the contribution of the signal at r_0 as follows. For each $k \neq k_0$, we consider the optimization problem

$$\min_W tr(W^T \hat{C} W), \text{ subject to } W^T H_k = 1, W^T H_{k_0} = 0,$$

which gives the optimal weight

$$\begin{aligned} W_k &= \hat{C}^{-1}(H_k, H_{k_0})((H_k, H_{k_0})^T \hat{C}^{-1}(H_k, H_{k_0}))^{-1} \\ &\quad \times (I_3, 0_3)^T \end{aligned}$$

and signal power matrix estimate

$$W_k^T \hat{C} W_k = (I_3, 0_3)((H_k, H_{k_0})^T \hat{C}^{-1}(H_k, H_{k_0}))^{-1}(I_3, 0_3)^T.$$

Similarly, if the baseline noise covariance matrix is assumed of the form $\sigma_0^2 I_n$, then the noise power matrix estimate at the projection direction W_k can be estimated by

$$\begin{aligned} \hat{\sigma}_0^2 W_k^T W_k &= (I_3, 0_3)((H_k, H_{k_0})^T \hat{C}^{-1}(H_k, H_{k_0}))^{-1} \\ &\quad \times (H_k, H_{k_0})^T \hat{C}^{-2}(H_k, H_{k_0}) \\ &\quad \times ((H_k, H_{k_0})^T \hat{C}^{-1}(H_k, H_{k_0}))^{-1}(I_3, 0_3)^T, \end{aligned}$$

where $\hat{\sigma}_0^2$ is a noise level estimate derived from the baseline noise data. The Bregman index can be defined by

$$\begin{aligned} \mathrm{NAI}_{bk|k_0} &= tr\left(\hat{\sigma}_0^{-2} W_k^T \hat{C} W_k (W_k^T W_k)^{-1}\right) \\ &\quad - \log\det\left(\hat{\sigma}_0^{-2} W_k^T \hat{C} W_k (W_k^T W_k)^{-1}\right) - 3. \end{aligned}$$

Let k_1 be the value of k at which $\mathrm{NAI}_{bk|k_0}$ attains the maximum and let r_1 be the corresponding location. There exist relative eigenvector matrix $U_1 = (u_{11}, u_{12}, u_{13})$ and relative eigenvalue diagonal matrix $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$, $\lambda_1 \geq \lambda_2 \geq \lambda_3$, satisfying

$$U_1^T W_k^T \hat{C} W_k = \Lambda U_1^T W_k^T W_k, \quad U_1^T W_k^T W_k U_1 = I_3.$$

We project the data $\mathbf{Y}$ along the orientation determined by the eigenvector corresponding to the maximum relative eigenvalue at location r_1, obtaining the associated time-course $\hat{W}_{k_1}\mathbf{Y}$.
- Step 2: Similarly, given location estimates $r_k, 1 \leq k \leq K-1$ found in the previous steps, define the $\mathrm{NAI}_{bk|k_s, 0 \leq s \leq K-1}$. We calculate the location estimate r_K and the associated time-course $\hat{W}_K\mathbf{Y}$.
- Stopping rule: Iteratively run Step 2 until the following stopping criterion is meet as follows. Let $\phi_K = \{k_s, 0 \leq s \leq K\}$. For $0 \leq K \leq n/3$, we first order $\mathrm{NAI}_{bk|\phi_K}, 1 \leq k \leq g$ decreasingly as $\mathrm{NAI}_{b(k)|\phi_K}, 1 \leq k \leq g$. Let k^* be the value of $k = k^*$ at which $\mathrm{NAI}_{bk|\phi_K}$ attains the maximum. For any v, we split NAI's into two parts and calculate the sample variances $\mathrm{var}_{1:v}$ and $\mathrm{var}_{(v+1):g}$ respectively. Choose the smallest v, denoted by v_K, at which NAI's have a best split in the sense that $\mathrm{var}_{1:v} + \mathrm{var}_{(v+1):g}$ attains the minimum. Let μ be the mean of $\mathrm{NAI}_{b(k)|\phi_K}, v_K + 1 \leq k \leq g$. Taking into account of multiple testing adjustment, we stop the further beamforming if

$$\mathrm{NAI}_{bk^*|\phi_K} < \mu + c \times \sqrt{\mathrm{var}_{(v_K+1):g}},$$

where c is the quantitle of the standard normal distribution at $0.05/g$.

Similarly, we define the SAM-based and the LCMV-based forward beamforming procedures called SAMFB and LCMVFB.

2.4 Constructing a functional network of contrast between two stimuli

We are interested in the problem of the functional connectivity change when stimuli varies. Functional connectivity is commonly inferred on the basis of correlations between source time-courses (Friston, 2011). Following this idea, we construct an undirected contrast network of the inferred sources with edges determined by their correlation contrast statistics. The details are as follows.

Suppose that we run n_f trials under face stimuli and n_s trials under scrambled face stimuli. We first calculated individual trial covariances, followed by an average over trials in order to estimate the sensor covariance matrices for face and scramble face respectively (Zhang and Su, 2015). For each pair of the estimated sources and for each trial, we projected the sensor data in the trial along the optimal directions obtained in the procedure, estimating the corresponding source-time series. The Fisher's z-transformation of the Pearson correlation coefficient between these two series in the pre-stimulus periods and in the post-stimulus periods were calculated respectively. This gave pre-stimulus z-values and post-stimulus z-values for n_f face trials and for n_s scramble-face trials, denoted by z_{ijk}. Here, $i = 1, 2$ stand for pre-stimulus and post-stimulus respectively, $j = 1, 2$ for face and scramble face respectively, and k for the trial index. Using these values, we calculated an observed t-value

$$t_{\mathrm{obs}} = \frac{(\bar{z}_{21} - \bar{z}_{11}) - (\bar{z}_{22} - \bar{z}_{12})}{\sqrt{(\mathrm{var}(z_{21}) + \mathrm{var}(z_{11}))/n_f + (\mathrm{var}(z_{22}) + \mathrm{var}(z_{12}))/n_s}}.$$

We then conducted a large number of say 10^5 random permutations on the trial labels. For each permutation, we calculated the corresponding permuted t-value. We counted the

proportion of times that the absolute permuted t-values exceeded the observed t-value t_{obs}, obtaining estimated P-values for each pair of the estimated sources. We carried out Hochberg's step-up corrections for multiple testing of significance of these P-values (Hochberg, 1988). To construct a correlation network for these selected sources, we took these estimated sources as nodes and set a link between two nodes if their P-value of correlation contrasts was less than or equal to some pre-selected level, say 5% or 1%.

3. NUMERICAL RESULTS

In this section, we present the results of data analyses on real and synthetic data by use of the proposed methodology.

3.1 Face-perception data

Henson et al. (2011) conducted a series of face-perception experiments, where 96 face trials and 50 scrambled face trials were performed on a healthy young adult subject. Each trial started with a central fixation cross (presented for a random duration of 400 to 600 ms), followed by a face or scrambled face (presented for a random duration of 800 to 1000 ms), and followed by a central circle for 1700 ms. The subject used either his/her left or right index finger to report whether he/she thought the stimulus was symmetrical or asymmetrical vertically through its center. The data were collected with a Neuromag VectorView system, containing a magnetometer and two orthogonal, planar gradiometers located at each of 102 positions within a hemispherical array situated in a light, magnetically shielded room. The sampling rate was 1100Hz. See https://www.fil.ion.ucl.ac.uk/spm/data/mmfaces/ for the details. Based on the measurements of 102 planar gradiometers, we employed the proposed method to localize brain regions where neuronal activities increase for the face stimuli relative to the scrambled face stimuli, and to reveal connectivity changes between the detected regions.

For this purpose, we first normalized the subject's MRI scan to a MRI template by using the software FieldTrip, on which a grid CTF system of 1 cm resolution was created with 1487 points. We then applied the neuroimaging software SPM8 to read and preprocess the recorded data, and to epoch the data generated from the face stimulus trials and the scrambled face stimulus trials respectively. This gives rise to a 102×771 data matrix for each trial: the first 220 columns for 200ms pre-stimuli and the later 551 columns for the stimuli. Following Zhang and Su (2015), we adopted the following strategy for estimating sensor covariance matrix. For each trial, we calculated the sample covariance matrix and noise covariance matrix by using the stimulus data and the pre-stimulus data respectively. Under some stationary assumption, we then averaged these sample covariance matrices over face trials and over scrambled face trials respectively, obtaining covariance estimators $\hat{C}_f$ and $\hat{C}_{f0}$ for the face dataset and $\hat{C}_s$ and $\hat{C}_{s0}$ for the scrambled face dataset. Here, the implicit assumption of stationarity on the source time-courses is made when we averaged these sample covariance matrices. We estimated the baseline noise levels by $\hat{\sigma}^2_{f0}$ and $\hat{\sigma}^2_{s0}$, the minimum diagonal elements in $\hat{C}_{f0}$ and $\hat{C}_{s0}$ respectively. Finally, we applied the proposed forward beamforming procedure on $\hat{C}_f$ and $\hat{C}_s$. In each forward step, we calculated log-contrasts of face over scrambled face at every grid points. We interpolated and overlaid these contrasts on the structural MRI of the subject, obtaining a contrast map of neuronal activity. We took the peak location as an source estimate. After applying the stopping rule, we acquired 31 dipole sources as listed in Table 1. The nulled NAI brain source maps of these sources are displayed in Figures 3 to 5. The corresponding source time-courses under normal face stimuli paired with those under scrambled face stimuli are presented in Figure 6.

[Put Figure 3∼5 here.]

[Put Figure 6 here.]

To recover the functional regions they belonged to, we performed a K-means clustering with a silhouette analysis on their physical locations, resulting in following clusters:

- Cluster 1 consisted of sources $v_{17}, v_{18}, v_{20}, v_{24}, v_{25}, v_{29}, v_{30}, v_{31}$.
- Cluster 2 consisted of sources $v_{11}, v_{14}, v_{15}, v_{16}, v_{19}, v_{21}, v_{22}, v_{26}, v_{27}, v_{28}$.
- Cluster 3 consisted of sources $v_1, v_3, v_7, v_8, v_9, v_{10}, v_{23}$.
- Cluster 4 consisted of sources $v_2, v_4, v_5, v_6, v_{12}, v_{13}$.

These clusters were located around the centers $(-3.4, 2.4, 7.5)$ cm, $(-0.8, -4, 2.2)$ cm, $(5.3, -3.1, 6.9)$ cm and $(4.5, 4.8, 5.8)$ cm respectively. Note that clusters 1, 2, 3, 4 are physically close to (or include) the superior temporal sulcus (STS) in parietal lobe, to the fusiform face area (FFA) and the occipital face area (OFA) in occipital lobe, to the precuneus (PCu) and the amygdala (AMG) in frontal lobe respectively. FFA and OFA analyze invariant aspects of faces that underlies recognition of individuals, whereas STS, PCu and AMG are responsible for retrieval of personal knowledge (i.e., semantic) and analysis of facial expression (e.g., Zhen et al., 2013; Kanwisher et al., 1997; Davies-Thompson and Andrews, 2011). This supported the hypothesis for the face-perception that the presence of faces was likely detected, characterized and categorized by clusters 1 and 2. The semantic information and facial expression were then retrieved by clusters 1, 3 and 4.

[Put Table 1 here.]

We further hypothesized that the information flow among these source clusters was mediated by a brain network. To validate this assumption, taking the dipole sources as nodes, we constructed a functional network which showed an increasing response to face stimuli over scrambled face stimuli. We assigned an edge to two nodes if there was a correlation change between them for face over scrambled face at the 5% significant level. Note that multiple testing was involved in these edge assignments. So we performed Hochberg's adjustment on the related p-values to remove false assignments.

There were still many assignments which survived from the adjustment using the threshold level of 5%. There were 10 edges between clusters 1 and 2, 1 edge between clusters 1 and 3, 2 edges for each of cluster pairs (1,4), (2,3), (2,4) and (3,4). We also made Hochberg's adjustment on these p-values by using a stringent threshold level, say 1%, obtaining the edge links for the following six pairs of nodes (v_{31}, v_{15}), (v_{17}, v_{21}), (v_{27}, v_{30}), (v_2, v_{25}), (v_6, v_{18}) and (v_{23}, v_{24}). Among them, the first three showed the links between clusters 1 and 2, followed by two links between clusters 1 and 4 and one link between clusters 1 and 3. These results suggested that the above four clusters worked dependently with a varying functional connectivity in response to the change of stimuli. In particular, clusters 1 and 2 were more active than the other clusters as there were more connectivity changes between them when stimuli were switched from scrambled faces to normal faces. This implied that face processing network dynamically adjusted its weight in connectivity among key face-selective regions such as STS, FFA and OFA in order to adapt to varying stimuli.

[Put Figure 7 here.]

3.2 Synthetic data

We begin with defining some notations. For any estimator $\hat{r}$ of an source location r, we define the localization bias by $|\hat{r}-r|$, the L_1 distance between $\hat{r}$ and r. We define the L_1 bias between two location sets B_1 and B_2 by $D(B_1, B_2) = \max_{b_1 \in B_1} \min_{b_2 \in B_2} |b_1 - b_2|$. Let m_r be the underlying $3 \times n$ source-course matrix at the location r and $\hat{\mathbf{m}}_{\hat{r}}$ be its estimator, where the three rows stand for source-courses in the x, y and z orientations respectively. We calculate the cross-correlation coefficients between the rows in $\mathbf{m}_r$ and the rows in $\hat{\mathbf{m}}_{\hat{r}}$, forming $\mathrm{Corr}(\hat{\mathbf{m}}_{\hat{r}}, \mathbf{m}_r)$, a 3×3 correlation coefficient matrix. The association index between $\hat{\mathbf{m}}_{\hat{r}}$ and $\mathbf{m}_r$ is defined as the Frobenius norm of the above correlation matrix, namely $||\mathrm{Corr}(\hat{\mathbf{m}}_{\hat{r}}, \mathbf{m}_r)||_F$. The larger the index, the stronger the association will be. For an experiment with multiple trials, an association index is define by averaging these trial-based indices. Note that the above association index attains the maximum value of 3 under the constraint that $\hat{\mathbf{m}}_{\hat{r}} = \mathbf{m}_r$ and the true source-courses in the x, y and z orientations are equal; it attains the maximum value of $\sqrt{3}$ under the constraint that the true source-courses in the x, y and z orientations are orthogonal to each other.

To evaluate the performances of BBFB, SAMFB and LCMVFB, we first created a 102-sensor MEG system (CTF/VSM) by use of the same head model as in the above face-perception data. We constructed 1487 regular 3-D grid points of resolution 1 cm within the head. These candidate source positions were aligned with the axes of the head coordinate system. A lead field matrix H (with dimension 102×4461) between the sensors and the grid points was then calculated by using the open software FieldTrip. The sampling rate is 1100 Hz. We considered two scenarios with time-invariant and time-variant orientations respectively.

Scenario 1 (Sources with time-invariant orientations). We assumed that there were non-zero neuronal sources $\theta_{fk}(t), t = 0, 1, ..., 771$ (for face stimuli) and $\theta_{sk}(t), t = 0, 1, ..., 771$ (for scramble face stimuli) at the locations $r_k, 1 \le k \le 31$ derived from the previous real data analysis, where

$$\begin{aligned} \theta_{fk}(t) &= \begin{cases} 0 & 0 \le t \le 221 \\ \sqrt{102}\eta_f \beta_{fk}(t) & 221 < t \le 771 \end{cases} \\ \theta_{sk}(t) &= \begin{cases} 0 & 0 \le t \le 221 \\ \sqrt{3}\eta_s \beta_{sk}(t) & 221 < t \le 771. \end{cases} \end{aligned}$$

were time-courses along time-invariant orientations $\eta_f = \left(\frac{10}{\sqrt{102}}, \frac{1}{\sqrt{102}}, \frac{1}{\sqrt{102}}\right)^T$, $\eta_s = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})^T$. The former orientation was dominated by x axis while the later one was weighted equally in x, y, z axes. Here, $(\beta_{fk}(t), \beta_{sk}(t)), 1 \le k \le 31$ were pairs of estimated time-courses obtained in the previous real data analysis. We defined the signal strengths by $\mathrm{SNR}_f = \sum_{t=221}^{771} \theta_{fk}(t)^2/500$ and $\mathrm{SNR}_s = \sum_{t=221}^{771} \theta_{sk}(t)^2/500$. To define sensor noise levels, we independently sampled N_{fni} and N_{snj} from an n-dimensional standard Normal. We simulated 50 multiple trials datasets. Each contained two independent sets of observations: $n_f = 96$ sensor measurements, $Y_{fi}(t)$, $t = 0, 1, ..., 771$, $i = 1, 2, ..., 96$ for face stimuli and $n_s = 50$ sensor measurements, $Y_{sj}(t)$, $t = 0, 1, ..., 771$, $j = 1, 2, ..., 50$ for scrambled face stimuli. These observations followed the models:

$$\begin{aligned} Y_{fi}(t) &= \sum_{k=1}^{31} H(r_1)\theta_{fk}(t) + \varepsilon_{fi}(t), \\ Y_{sj}(t) &= \sum_{k=1}^{31} H(r_1)\theta_{sk}(t) + \varepsilon_{sj}(t), \end{aligned} \tag{2}$$

where $\varepsilon_{fi}(t) = N_{fni}\sqrt{\mathrm{SNR}_f}$ is the sensor noise vector for the ith trial under face stimuli and $\varepsilon_{sj}(t) = N_{snj}\sqrt{\mathrm{SNR}_s}$ is the sensor noise vector for the jth trial under scrambled face stimuli.

We took time instances $0, 1, ..., 221$ as the pre-stimulus time points and the remaining time instances as stimulus time points. We considered the following tasks in presence that we don't know these sources:

Task (i): Recover the source locations based on simulated sensor data $Y_{fi}(t)$, $t = 0, 1, ..., 771$, $i = 1, ..., 96$.

Task (ii): Recover the source locations based on simulated sensor data $Y_{si}(t)$, $s = 0, 1, ..., 771$, $i = 1, ..., 50$.

Task (iii): Reveal the source locations where the neuronal activities increase for the face stimuli relative to the scrambled face stimuli.

Scenario 2 (Sources with time-varying orientations). We adopted the same setting as Scenario 1 except re-defining

$(\beta_{fk}(t), \beta_{sk}(t)), 1 \le k \le 31$ as

$$\theta_{fk}(t) = \begin{cases} 0 & 0 \le t \le 221 \\ u_{fk}(t) & 221 < t \le 771 \end{cases},$$
$$\theta_{sk}(t) = \begin{cases} 0 & 0 \le t \le 221 \\ v_{sk}(t) & 221 < t \le 771 \end{cases}$$

where $u_{fk}(t) = \beta_{fk}(t)(10|\sin(\pi t/221)|, |\sin(2\pi t/221)|, |\sin(3\pi t/221)|)^T$ and $u_{sk}(t) = \beta_{sk}(t)(|\sin(4\pi t/221)|, |\sin(5\pi t/221)|, |\sin(6\pi t/221)|)^T$. We consider the same tasks as in scenario 1.

For each dataset generated above, we first applied BBFB to sub-datasets $\{Y_{fi}(t) : t = 0, 1, ..., 771, i = 1, 2, ..., 96\}$ and $\{Y_{sj}(t) : t = 0, 1, ..., 771, j = 1, 2, ..., 50\}$ separately. Carrying out tasks (i) and (ii) resulted in two estimated sets of source locations, say $\hat{B}_f$ and $\hat{B}_s$, where neural activity was predicted to increase against background noises. We calculated the L_1 biases $D(\hat{B}_f, B)$ and $D(\hat{B}_s, B)$, where B is the underlying source location set. We then carried out task (iii) by applying BBFB to the above subsets simultaneously to identify a set of source locations, namely $\hat{B}_{fs}$, where neural activity differed for faces and scrambled faces. We calculated the L_1 bias $D(\hat{B}_{fs}, B)$. To sum up, in each scenario, we obtained three sequences of simulated bias values, respectively, for the face stimuli against background noises, for the scrambled face stimuli against background noises, and for the face stimuli against the scrambled stimuli. Similarly, we performed LCMVFB and SAMFB on the above datasets. In each scenario, we also obtained three sequences of simulated bias values for each method. In scenario 1, for task (i), the mean biases (standard errors) of BBFB, LCMVFB and SAMFB were: 2.5987(0.0308), 3.1881(0.0349) and 2.6275(0.0353) for task (i); 2.7887(0.0330), 3.2356(0.0315) and 2.7206(0.0342) for task (ii); 2.7425(0.0424), 3.1231(0.0335) and 2.7763(0.0341) for task (iii). In scenario 2, these values were: 2.5956(0.0314), 3.1394(0.0324) and 2.5831(0.0315) for task (i); 2.7394(0.0352), 3.2981(0.0365) and 2.8213(0.0333) for task (ii); 2.7556(0.0396), 3.1975(0.0331) and 2.7513(0.0410) for task (iii). We displayed these localization bias values by multiple box-and-whisker plots in Figure 8. The results demonstrate that both BBFB and SAMFB can offer superior source localization results than LCMVFB due to their depth-invariant features. BBFB and SAMFB performed similarly in both scenarios.

[Put Figure 8 here.]

To compare these procedures when both localization bias and association were considered, we repeated the above experiment in scenario 2 for 50 times. We calculated the values of both L_1 bias and association index for each method in each experiment. These values were presented in Figure 9 by use of multiple box-and-whisker plots. In task (i), BBFB, LCMVFB and SAMFB respectively had the average localization biases (standard errors) 2.6963(0.0207), 3.0888(0.0306) and 2.5550(0.0190); the average association indices (standard errors) 0.7786(0.0044), 0.8153(0.0040) and 0.7663(0.0051). In task (ii), these values were 2.6831(0.0323), 3.3400(0.0499) and 2.8900(0.0364); 0.7264(0.0158), 0.7947(0.0040) and 0.5399(0.0168). In task (iii), these values were 2.7894(0.0255), 3.1675(0.0247) and 2.7969(0.0470); 0.7426(0.0036), 0.8303(0.0044) and 0.4834(0.0033) under face stimuli, and 0.7513(0.0053), 0.7568(0.0042) and 0.4217(0.0020) under scramble-face stimuli. These results suggest that although BBFB and SAMFB performed similarly in localization, the former outperformed the latter in association with the true source-courses. LCMVFB was the winner in terms of association index.

[Put Figure 9 here.]

4. DISCUSSION AND CONCLUSION

There are two major tasks for MEG neuroimaging: the determination of stimulus-specific cortical regions and the determination of their functional connectivity. In order to conduct a reliable source connectivity analysis, there is a need for spatial filters which can reduce the source depth bias and the source smearing in source reconstruction. For this purpose, we have proposed a family of forward beamformers for inverting electromagnetic models and detecting temporally correlated sources with MEG data: the Bregman-divergence-based, the SAM-based and the LCMV-based. All these methods are able to cope with the source smearing by the forward nulling the sources. In addition, the Bregman-divergence-based is scale-invariant and allows for time-varying source orientations. The SAM-based is only partially scale-invariant when orientations with evenly distributed weights in x, y and z axes and is restricted to fixed source orientation settings. Although the LCMV-based allows for time-varying source orientations, it is not scale-invariant. So, intuitively the Bregman-divergence-based procedure is expected to perform better than the other two. By the simulations, we have demonstrated that there is a remarkable improvement of localization bias if using the Bregman-divergence-based or the SAM-based procedure rather than the LCMV-based procedure. We have further developed a method for constructing a functional network of contrast between two stimuli. We have evaluated the performance of the proposed procedure by an analysis of MEG data derived from the face-perception experiment, revealing a novel functional connectivity network. Finally, following Zhang and Liu (2015) and Zhang and Su (2015), under certain regularity conditions, we can prove a consistency for the Bregman-divergence-based forward beamforming and use the autocovariance function to study time-varying sources. The details are beyond the scope of this paper.

APPENDIX: PROOFS

Proof of Proposition 1: It is straightforward and thus omitted.

Proof of Proposition 2: For any scale transformation $\tilde{S}_k = H_k S$ on H_k with $S = \text{diag}(s_1, s_2, s_3)$, we have

$$\begin{aligned}
&\hat{\sigma}_0^{-2}\left[\tilde{S}_k^T\hat{C}^{-2}\tilde{S}_k\right]^{-1}\left[\tilde{S}_k^T\hat{C}^{-1}\tilde{S}_k\right] \\
&= \hat{\sigma}_0^{-2}\left[SH_k^T\hat{C}^{-2}H_kS\right]^{-1}\left[SH_k^T\hat{C}^{-1}H_kS\right] \\
&= \hat{\sigma}_0^{-2}S^{-1}\left[H_k^T\hat{C}^{-2}H_k\right]^{-1}\left[H_k^T\hat{C}^{-1}H_kS\right].
\end{aligned}$$

Consequently,

$$\begin{aligned}
\text{NAI}_{bk} &= tr\left(\hat{\sigma}_0^{-2}S^{-1}\left[H_k^T\hat{C}^{-2}H_k\right]^{-1}H_k^T\hat{C}^{-1}H_kS\right) \\
&- \log\det\left(\hat{\sigma}_0^{-2}S^{-1}\left[H_k^T\hat{C}^{-2}H_k\right]^{-1}H_k^T\hat{C}^{-1}H_kS\right) \\
&-3 \\
&= tr\left(\hat{\sigma}_0^{-2}\left[\tilde{S}_k^T\hat{C}^{-2}\tilde{S}_k\right]^{-1}\tilde{S}_k^T\hat{C}^{-1}\tilde{S}_k\right) \\
&- \log\det\left(\hat{\sigma}_0^{-2}\left[\tilde{S}_k^T\hat{C}^{-2}\tilde{S}_k\right]^{-1}\tilde{S}_k^T\hat{C}^{-1}\tilde{S}_k\right) - 3.
\end{aligned}$$

This implies that the new index is scale-invariant. In particular, letting $S = D_k^{-1}$, we have

$$\begin{aligned}
\text{NAI}_{bk} &= tr\left(\hat{\sigma}_0^{-2}\left[\tilde{H}_k^T\hat{C}^{-2}\tilde{H}_k\right]^{-1}\tilde{H}_k^T\hat{C}^{-1}\tilde{H}_k\right) \\
&- \log\det\left(\hat{\sigma}_0^{-2}\left[\tilde{H}_k^T\hat{C}^{-2}\tilde{H}_k\right]^{-1}\tilde{H}_k^T\hat{C}^{-1}\tilde{H}_k\right) - 3.
\end{aligned}$$

This complete the proof.

Received 17 August 2020

REFERENCES

[1] Brookes, M.J., Woolrich, M., Luckhoo, H., Price, D., Hale, J.R., Stephenson, M. C., Barnes, G. R., Smith, S. M., Morris, P. (2011). Investigating the electrophysiological basis of resting state networks using magnetoencephalography. *Proc. Natl. Acad. Sci. USA*, **98**, 694-699.

[2] Chan, H., Chen, L., Chen, I., Chen, Y.S. (2011). Beamformer-based spatiotemporal imaging of linearly-related source components using electromagnetic neural signals. *NeuroImage*, **114**, 1-17.

[3] Davis, J.V., Kulis, B., Jain, P., Sra, S. and Dhillon, I.S. (2007). Information-theoretic metric learning. In *Proc. IEEE Intern. Conf. on Machine Learning.*

[4] Davies-Thompson, J. & Andrews, T.J. (2011). The localization and functional connectivity of face-selective regions in the human brain. *Jour. Vision*, **11**, 647.

[5] Drakesmith, M., El-Deredy, W., Welbourne, S. (2013). Reconstructing coherent networks from electroencephalography and magnetoencephalography with reduced contamination from volume conduction or magnetic field spread. *PLoS ONE*, **8**, e81553.

[6] Friston, K. (2011). Functional and effective connectivity: A review. *Brain Connectivity*, **1**, 13-36.

[7] Hämäläinen, M., Hari, R, Ilmoniemi, R.J., Knuutila, J., Lounasmaa, O.V. (1993). Magnetoencephalography theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev. Modern Phys.*, **21**, 413-460.

[8] Heller, L. and van Hulsteyn, D.B. (1992). Brain stimulation using electromagnetic sources: theoretical aspects. *Biophys J.*, **63**, 129-138.

[9] Henson, R.N., Wakeman,D.G., Litvak, V., Friston, K.J. (2011). A parametric empirical Bayesian framework for the EEG/MEG inverse problem: generative models for multi-subject and multi-modal integration. *Front. Hum. Neuroscience*, **5**, 1-16.

[10] Sarvas, J. (1987). Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Phys. Med. Biol.*, **32**, 11-22.

[11] Hillebrand, A., Barnes, G.R. (2002). A quantitative assessment of the sensitivity of whole-head MEG to activity in the adult human cortex. *Neuroimage*, *16*, 638-650.

[12] Hochberg, Y. (1988). A Sharper Bonferroni Procedure for Multiple Tests of Significance. *Biometrika,* **75**, 800-802.

[13] Huang, M.X., Shih, J.J., Lee, R.R., Harrington, D.L., Thoma, R.J., Weisend, M.P., Hanlon, F., Paulson, K.M., Li, T., Martin, G.A., Miller, G.A., Canive, J.M. (2004). Commonalities and differences among vectorized beamformers in electromagnetic source imaging. *Brain Topography*, **16**, 139-158.

[14] Oostenveld, R., Fries, P., Maris, E.,& Schoffelen, J.M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, **2011**, doi:10.1155/2011/156869.

[15] Robinson, S., Vrba, J. (1998). Functional neuroimaging by synthetic aperture magnetometry. In: Yoshimoto, T., Kotani, M., Kuriki, S. Karibe, H., Nakasato, N. (Eds.), *Recent Advances in Biomagnetism.* Tohoku University Press, Sendai, Japan, pp. 302-305.

[16] Sarvas, J. (1987). Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Phys. Med. Biol.* **32**, 11-22.

[17] Sekihara, K. & Nagarajan, S.S. (2010). Adaptive Spatial Filters for Electromagnetic Brain Imaging. *Springer-Verlag*, Berlin.

[18] Tsao, D.Y. and Livingstone, M.S. (2008). Mechanisms of face perception. *Annu. Rev. Neurosci.*, **31**, 411-437.

[19] Ishai A. (2008). Let's face it: it's a cortical network. *Neuroimage*, **40**, 415-419.

[20] Van Veen, B.D., Van Drongelen, W., Yuchtman, M., Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans. Biomed. Eng.*, **44**, 867-880.

[21] Yao, Z., Zhang, Y., Bai, Z. and Eddy, W.F. (2018). Estimating the Number of Sources in Magnetoencephalography Using Spiked Population Eigenvalues. *Jour. Ameri. Stat. Assoc.*, **113**, 505-518.

[22] Yao, Z., Fan, Z., Hayashi, M. and Eddy W.F. (2020). Quantifying time-varying sources in magnetoencephalography-A discrete approach. *Ann. Appl. Stat.*, **14**, 1379–1408.

[23] Zhang, J. and Liu, C. (2015). On Linearly Constrained Minimum Variance Beamforming. *Jour. Mach. Learn. Res.*, **16**, 2099-2145.

[24] Zhang, J., Liu, C., Green, G. (2014). Source localization with MEG data: A beamforming approach based on covariance thresholding. *Biometrics*, **70**, 121-131.

[25] Zhang, J. and Su, L. (2015). Temporal autocorrelation-based beamforming with MEG beuroimaging data. *Jour. Ameri. Stat. Assoc.*, **110**, 1375-1388.

[26] Zhen, Z., Fang, H., Liu, J. (2013). The hierarchical brain network for face recognition. *PLoS ONE*, **8**, e59886.

Table 1. The peak locations in the order which they been revealed and the clusters which they belong to.

Ind.	Loc.(cm)	Cl.	Ind.	Loc.(cm)	Cl.
v_1	(7,-3, 6)	3	v_{17}	(-5, 3, 8)	1
v_2	(8,3,5)	4	v_{18}	(-5,-1,9)	1
v_3	(7,0,8)	3	v_{19}	(-2,-2,-1)	2
v_4	(6,4,7)	4	v_{20}	(-4,2,6)	1
v_5	(4,5,8)	4	v_{21}	(3,-3,0)	2
v_6	(4,6,5)	4	v_{22}	(-1,-6,3)	2
v_7	(5,-5,7)	3	v_{23}	(3,-3,4)	3
v_8	(4,-5,7)	3	v_{24}	(-3,5,6)	1
v_9	(4,-3,9)	3	v_{25}	(-4,4,9)	1
v_{10}	(7,-3,7)	3	v_{26}	(2,-2,4)	2
v_{11}	(0,-7,4)	2	v_{27}	(3,-5,1)	2
v_{12}	(2,6,7)	4	v_{28}	(-3,-1,1)	2
v_{13}	(3,5,3)	4	v_{29}	(-1,5,6)	1
v_{14}	(1,-5,1)	2	v_{30}	(-1,1,9)	1
v_{15}	(-5,-5,4)	2	v_{31}	(-4,0,7)	1
v_{16}	(-6,-4,5)	2			

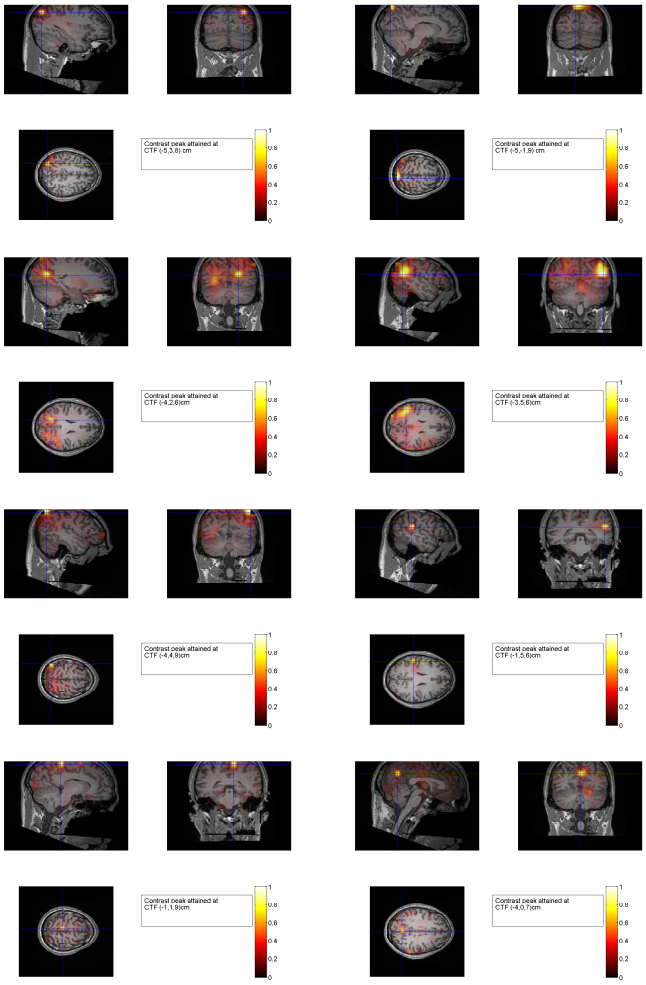

Figure 2. Nulled NAI maps for sources $v_{17}, v_{18}, v_{20}, v_{24}, v_{25}, v_{29}, v_{30}$ and v_{31} (in orders from the left to the right and from the upper to the bottom) in Cluster 1 derived from the face-perception data. Cluster 1 is around STS. Color will not be used in print.

Jian Zhang
School of Mathematics, Statistics and Actuarial Science,
University of Kent, Canterbury, Kent CT2 7NF, UK.
E-mail address: jz79@kent.ac.uk

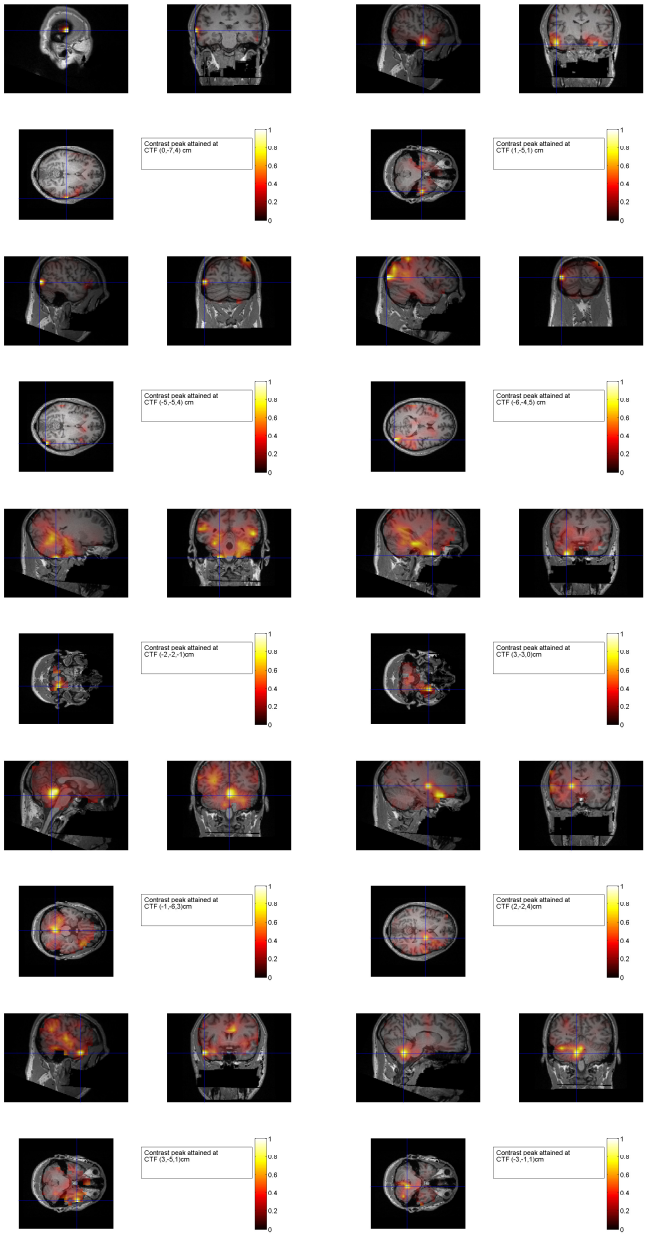

Figure 3. Nulled NAI maps for sources $v_{11}, v_{14}, v_{15}, v_{16}, v_{19}, v_{21}, v_{22}, v_{26}.v_{27}$ and v_{28} (in orders from the left to the right and from the upper to the bottom) in Cluster 2 derived from the face-perception data. Cluster 2 is around OFA and FFA. Color will not be used in print.

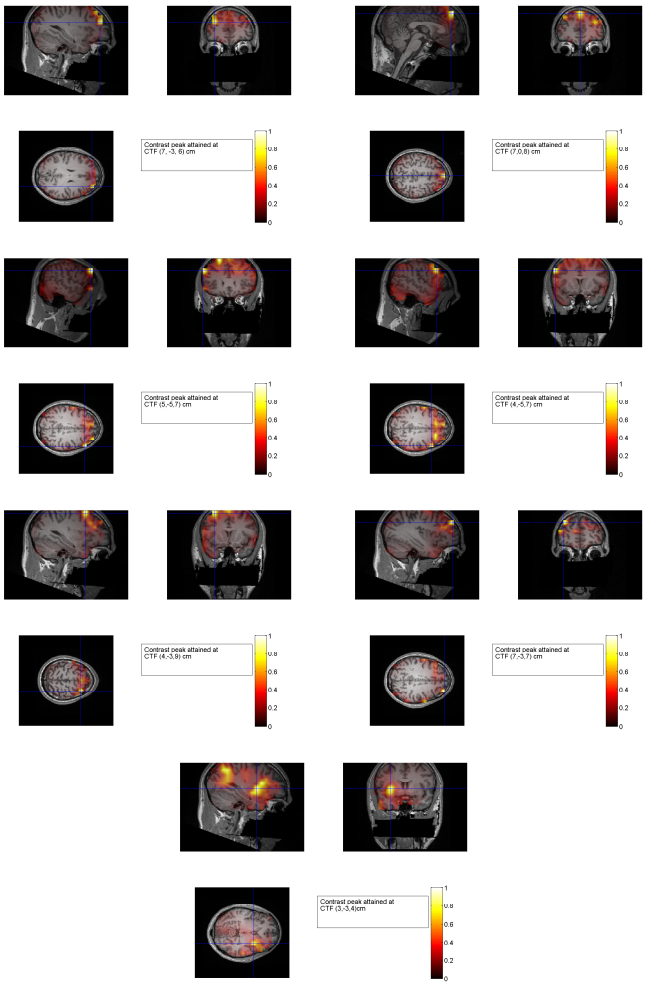

Figure 4. Nulled NAI maps for sources $v_1, v_3, v_7, v_8, v_9, v_{10}$ and v_{23} (in orders from the left to the right and from the upper to the bottom) in Cluster 3 derived from the face-perception data. Cluster 3 is around PCu. Color will not be used in print.

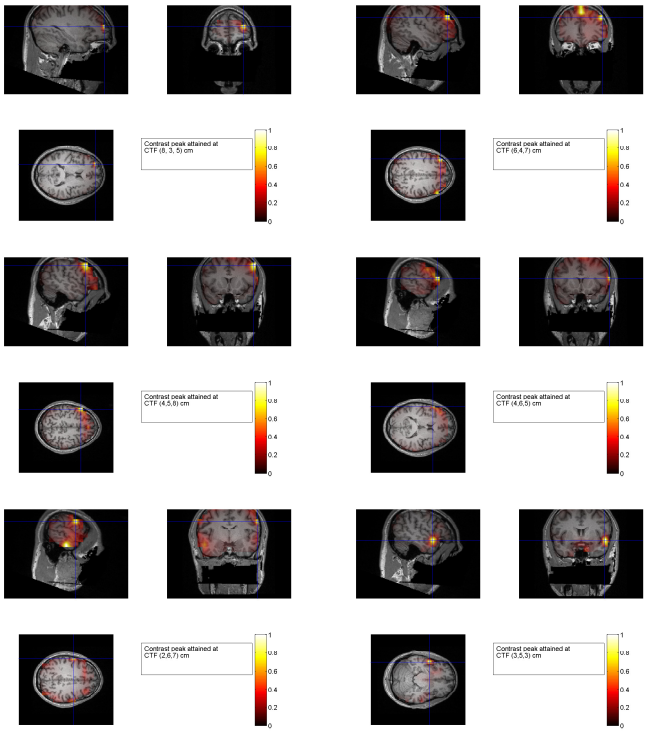

Figure 5. Nulled NAI maps for sources $v_2, v_4, v_5, v_6, v_{12}$ and v_{13} (in orders from the left to the right and from the upper to the bottom) in Cluster 4 derived from the face-perception data. Cluster 4 is around AMG and PCu. Color will not be used in print.

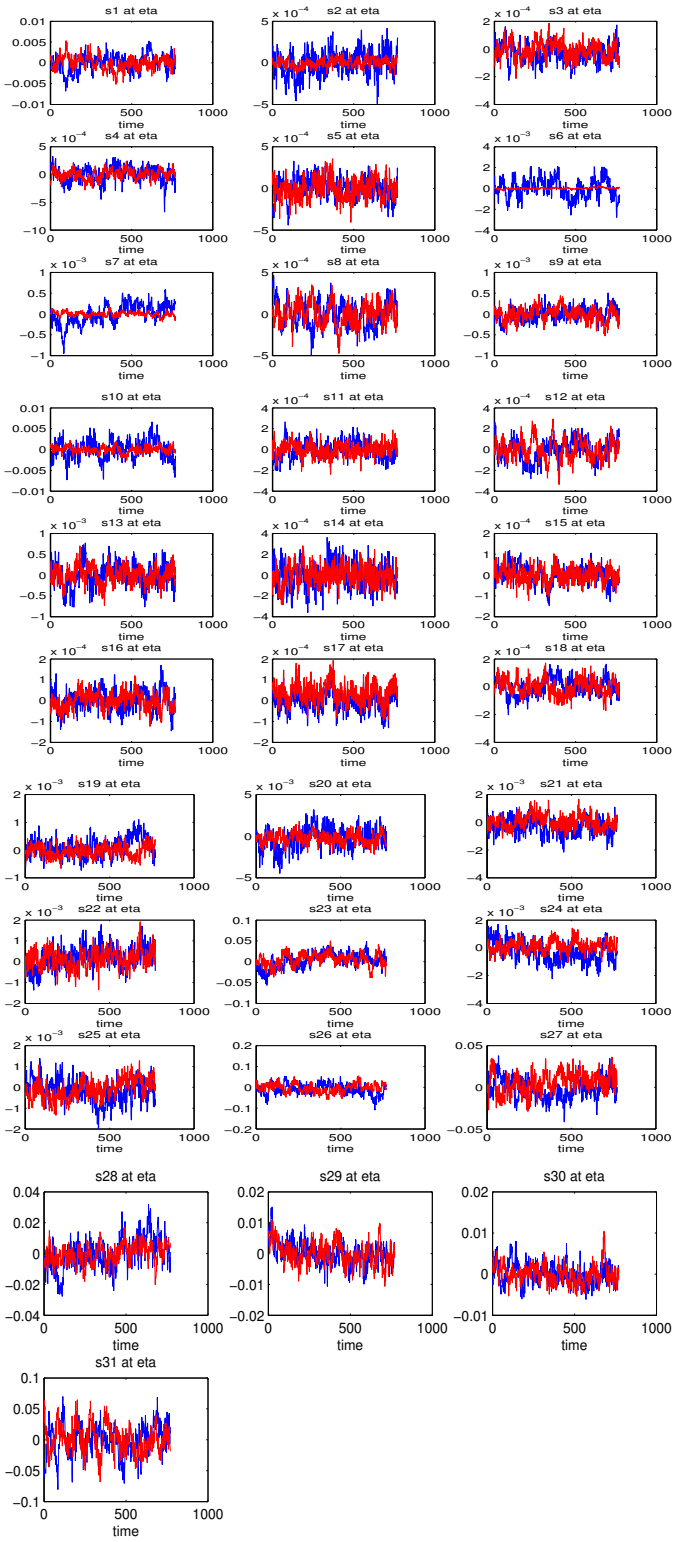

Figure 6. Estimated source time-course plots (average over the corresponding multiple trials) plots along the orientations with maximum eigenvalues for sources $v_1 \sim v_{31}$ in orders from left to right and from top to bottom. Blue colored curves are for face stimuli while red colored are for scrambled face stimuli. Color will not be used in print.

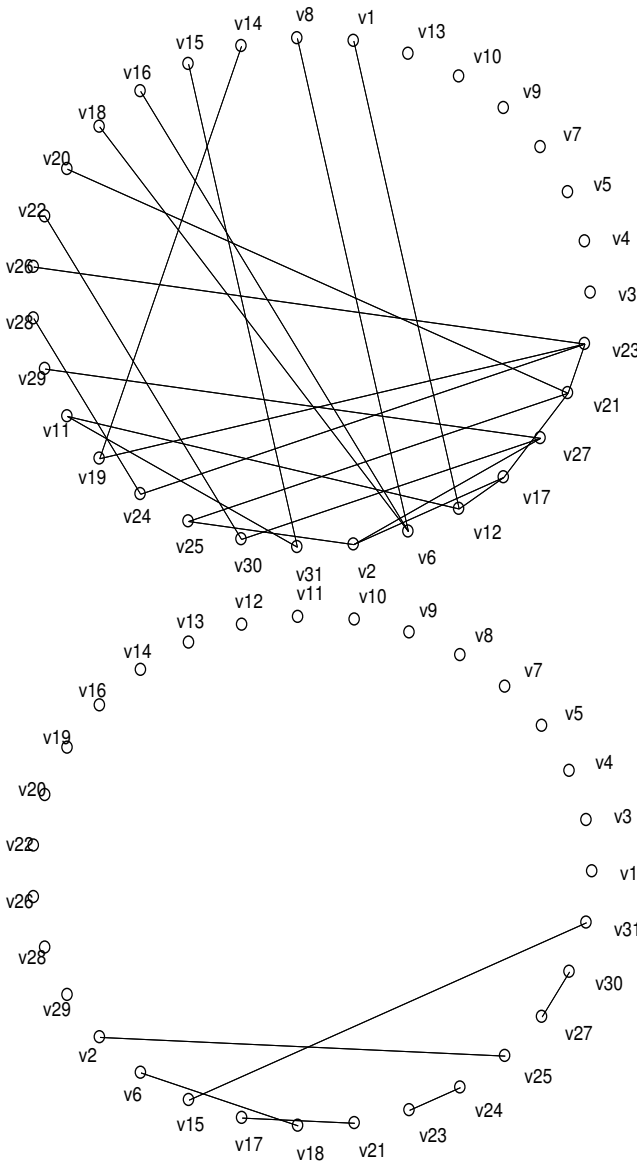

Figure 7. Functional networks: The plot at the top is a network with Hochberg's multiple testing adjustment at the 5% significance level while the plot in the bottom is one with Hochberg's multiple testing adjustment at the 1% significance level.

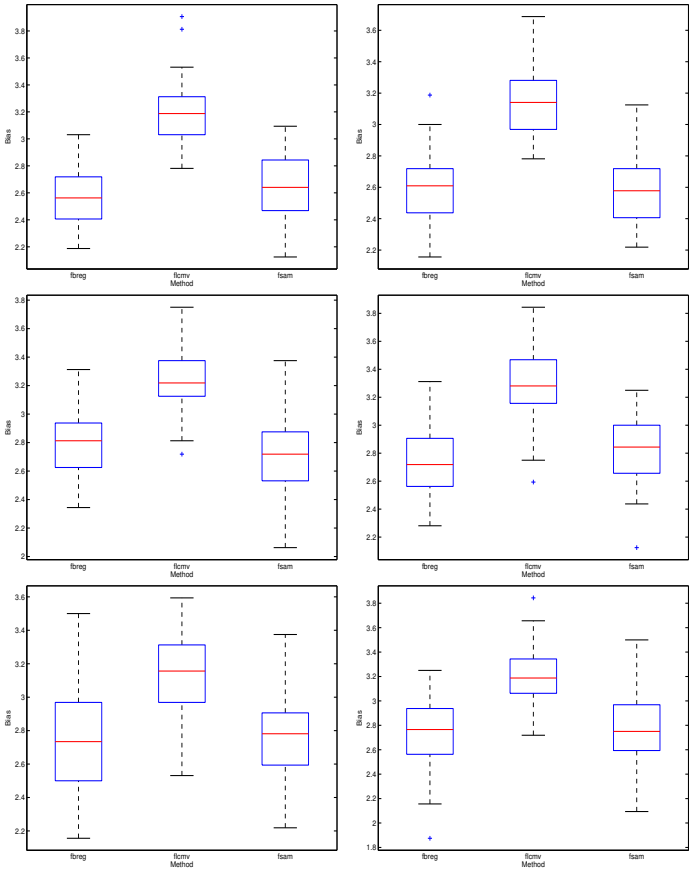

Figure 8. Localization bias over 50 replicates: Left column of plots for scenario 1 with tasks (i)-(iii). Right column of plots for scenario 2 with tasks (i)-(iii). Color will not be used in print.

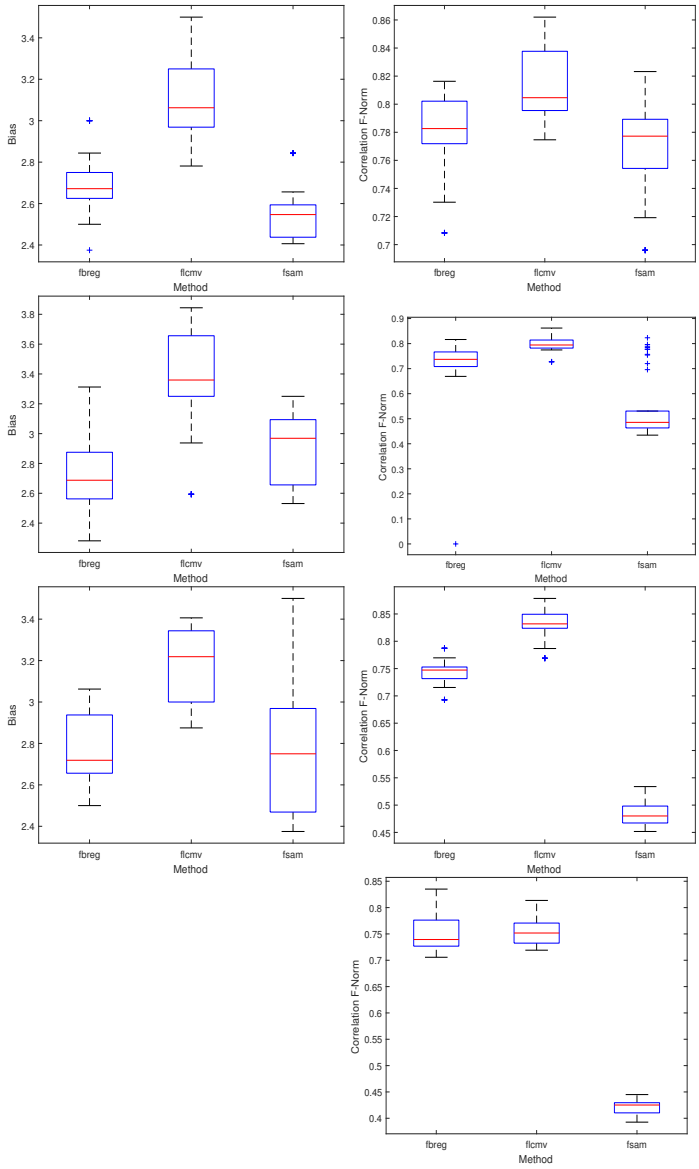

Figure 9. Plots of localization-biases and association indices over 50 *replicates in scenario 2: The left column shows the box-plots of localisation biases for tasks (i)-(iii) respectively. In the right column, the top two plots in the right column show the box-plots of the association indices for tasks (i)-(ii) respectively while the bottom three are source association plots in task (iii) for the face stimuli and the scramble face stimuli respectively. Color will not be used in print.*