

Cortical Localization of the Auditory Temporal Response Function from MEG via Non-convex Optimization

Proloy Das¹, Christian Brodbeck², Jonathan Z. Simon^{1 2 3} and Behtash Babadi^{1 2}

¹Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, USA

²Institute for Systems Research, University of Maryland, College Park, MD, USA

³Department of Biology, University of Maryland, College Park, MD, USA.

Email: proloy@umd.edu, brodbeck@umd.edu, jzsimon@umd.edu, behtash@umd.edu

Abstract—The magnetoencephalography (MEG) response to continuous auditory stimuli, such as speech, is commonly described using a linear filter, the auditory temporal response function (TRF). Though components of the sensor level TRFs have been well characterized, the underlying neural sources responsible for these components are not well understood. In this work, we provide a unified framework for determining the TRFs of neural sources directly from the MEG data, by integrating the TRF and distributed forward source models into one, and casting the joint estimation task as a Bayesian optimization problem. Though the resulting problem emerges as non-convex, we propose efficient solutions that leverage recent advances in evidence maximization. We demonstrate the effectiveness of the resulting algorithm in both simulated and experimentally recorded MEG data from humans.

I. INTRODUCTION

Neuroimaging experiments designed for probing sensory information processing in the brain traditionally use repeated trials of synthetic and transient stimuli [1]. In real world, however, the brain naturally processes more complex, informative sensory stimuli that evolves over time in a continuous fashion. As an example, consider an experiment in which the subject is listening to a series of repeated tones vs. a natural setting where the subject is listening to speech. In the former case, the neural phenomena of interest are isolated by experiment design, so that the evoked responses can be directly analyzed as the activation in response to the stimuli. In the later case, assessing the dependency of the neural response on the continuous stimuli is not straightforward and is usually carried out via model-based inference procedures. Previous studies have shown that for auditory MEG and EEG experiments under such naturalistic scenarios, the response can be modeled as a linear convolution of a continuous stimulus variable with a time-invariant filter [2]–[4], often referred to as the Temporal Response Function (TRF). The TRFs play a crucial role in characterizing the temporal structure of auditory information processing in the brain. For instance, the prominent negative peak observed in the TRF with a latency of ~ 100 ms has been shown to be modulated by the attentional state in a competing-speaker environment [5].

This work is supported in part by the National Science Foundation Awards No. 1552946 and 1734892.

Although the functional roles of specific components of the TRF have been studied over the sensor space, the cortical distributions of the underlying neural responses are not well-understood. Given that the spatial distribution of the TRF over neural sources can potentially unveil the brain dynamics that underlie continuous stimulus processing, existing studies either project the estimated sensor space TRF components to the source space using current dipole fitting methods [4], or use distributed source localization to map the MEG data to the cortical surface, followed by estimating the TRFs [6]. As a consequence of operating in this two-stage fashion, the resulting estimates are often highly biased and more importantly, suffer from spatial leakage (e.g., blurred reconstruction of focal sources), thus severely limiting the spatial resolution of the cortical TRF estimates. Even though the spatial resolution of the state-of-the-art source localization methods have significantly improved [7]–[10], they require various prior assumptions that inevitably increases the bias of the TRF estimates.

To mitigate these shortcomings, here we present a unified framework for determining the cortical localization of the TRFs *directly* from the MEG data, by integrating the TRF estimation and distributed forward source models into one. To this end, we model the neural response of each source location as a linear convolution of its local temporal response function with the continuous stimulus variables of interest, corrupted by background brain activity. We then cast the problem of cortical TRF localization as an optimization problem where the likelihood of the recorded MEG response is maximized over all the TRFs, thus eliminating the need for the aforementioned two-stage procedures. Though the resulting problem turns out to be non-convex, we provide an efficient coordinate-descent algorithm leveraging recent advances in evidence maximization that yields the solution in a fast and efficient manner. Finally, we demonstrate the utility of our proposed algorithm through application to simulated and experimentally recorded MEG data.

II. PRELIMINARIES AND NOTATIONS

Consider a task in which the subject is listening to a speech stream. We are interested in assessing the dependency of

the neural currents evoked at different anatomical locations in the brain on certain acoustic features of the speech. In particular, we consider the speech envelope (reflecting the momentary acoustic power of the speech stimuli) e_t as the main covariate, where $t = 1, \dots, T$ denotes the discrete time indices corresponding to a sampling frequency of f_s .

Typical MEG recordings consist of measurements of magnetic fields and gradients at different positions over the scalp, represented as a multidimensional time series. Let N denote the number of sensors. We denote the MEG observation at j^{th} channel at time t by $y_{j,t}$, for $1 \leq j \leq N$, $1 \leq t \leq T$, and the measurement vector at time t by $\mathbf{y}_t := [y_t^{(1)}, y_t^{(2)}, \dots, y_t^{(N)}]^\top$. Finally we denote multidimensional time series, consisting of all MEG measurements in the time window $[1, T]$ by a matrix $\mathbf{Y} := [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T]$.

In a distributed source model, the brain volume is divided into M voxels, while the primary current in each voxel is modeled by the dipole moment of a current source placed at the center of the voxel. Let the dipole moment at the m^{th} voxel at time t be denoted by a 3D vector $\mathbf{j}_{m,t} := [j_{m,0,t}, j_{m,1,t}, j_{m,2,t}]$, and the overall source current vector at time t by $\mathbf{j}_t = [j_{1,t}, j_{2,t}, \dots, j_{M,t}]^\top$. Similarly, we denote the neural currents in the time window $[1, T]$ by a matrix $\mathbf{J} := [\mathbf{j}_1, \mathbf{j}_2, \dots, \mathbf{j}_T]$. Then, the MEG observations \mathbf{Y} can be related to the neural sources via the following forward model:

$$\mathbf{Y} = \mathbf{L}\mathbf{J} + \mathbf{W}, \quad (1)$$

where $\mathbf{L} \in \mathbb{R}^{N \times 3M}$ is the *lead-field matrix*, a mapping from the *source space* to the *sensor space*, and \mathbf{W} is the measurement noise matrix. Typically, $M \sim 10^3 - 10^5$ and $N \sim 10^2$, which make the problem highly under-determined and necessitate appropriate priors for source estimation.

Following [2]–[4], we model the neural response (dipole activity) at the m^{th} source, in the d^{th} direction, and at time n , $j_{m,d,t}$, as a noise corrupted version of convolution between the continuous stimulus with a linear time-invariant filter:

$$j_{m,d,t} = (\boldsymbol{\tau}_{m,d})^\top \mathbf{e}_t + v_{m,d,t}, \quad (2)$$

for $m = 1, 2, \dots, M$ and $d = 0, 1, 2$, where $\boldsymbol{\tau}_{m,d} := [\tau_{m,d,1}, \tau_{m,d,2}, \dots, \tau_{m,d,p}]^\top$ is the TRF kernel of order l at the m^{th} source, $\mathbf{e}_t := [e_t, e_{t-1}, \dots, e_{t-l+1}]^\top$ is the stimulus history, and $v_{m,d,t}$ is a nuisance component corresponding to the background (stimulus-independent) brain activity. The 3D TRFs, $\boldsymbol{\tau}_m := [\tau_{m,0}, \tau_{m,1}, \tau_{m,2}]$, capture the temporal dependency of the neural currents on the stimulus. Let $\boldsymbol{\Phi} := [\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \dots, \boldsymbol{\tau}_M]^\top$ be the TRF matrix formed by stacking the 3D TRFs across the M brain sources, \mathbf{V} be the stimulus-independent neural current matrix and $\mathbf{S} := [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_T]$ be the stimulus covariate matrix formed by stacking the stimulus history vectors in the time window $[1, T]$, so that Eq. (2) takes the form of the following matrix equation:

$$\mathbf{J} = \boldsymbol{\Phi}\mathbf{S} + \mathbf{V}. \quad (3)$$

Our main goal in this work is to *directly* estimate the TRF matrix, $\boldsymbol{\Phi}$, given the MEG measurement matrix \mathbf{Y} , the lead-

field matrix \mathbf{L} , and the stimulus covariate matrix \mathbf{S} , without resorting to the existing two-stage procedures [4], [6].

III. PROBLEM FORMULATION

We first assume that the measurement noise has a temporally uncorrelated multivariate Gaussian distribution with covariance matrix $\boldsymbol{\Sigma}_w$:

$$p(\mathbf{Y}|\mathbf{J}) \propto |\boldsymbol{\Sigma}_w|^{-T/2} \exp\left\{-\frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{J}\|_{\boldsymbol{\Sigma}_w}^2\right\} \quad (4)$$

where $\|\mathbf{A}\|_{\mathbf{B}} := \sqrt{\text{tr}\{\mathbf{A}^\top \mathbf{B} \mathbf{A}\}}$. Next, we adopt the following prior on the stimulus-independent background activity \mathbf{V} :

$$p(\mathbf{V}|\boldsymbol{\Gamma}) \propto \prod_{m=1}^M |\boldsymbol{\Gamma}_m|^{-T/2} \exp\left\{-\frac{1}{2}\sum_{m=1}^M \|\mathbf{V}_m\|_{\boldsymbol{\Gamma}_m}^2\right\}, \quad (5)$$

implying that the background dipole activity at the m^{th} voxel has an independent zero mean Gaussian distribution with unknown 3×3 covariance matrix $\boldsymbol{\Gamma}_m$, so that,

$$p(\mathbf{J}|\boldsymbol{\Phi}, \boldsymbol{\Gamma}) \propto |\boldsymbol{\Gamma}|^{-T/2} \exp\left\{-\frac{1}{2}\|\mathbf{J} - \boldsymbol{\Phi}\mathbf{S}\|_{\boldsymbol{\Gamma}}^2\right\}, \quad (6)$$

where the $3M \times 3M$ matrix $\boldsymbol{\Gamma}$ consists of diagonal blocks given by $\boldsymbol{\Gamma}_m$'s. Under these assumptions, the joint distribution of the MEG observation and source currents is given by:

$$p(\mathbf{Y}, \mathbf{J}|\boldsymbol{\Phi}, \boldsymbol{\Gamma}) \propto |\boldsymbol{\Sigma}_w|^{-T/2} \exp\left\{-\frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{J}\|_{\boldsymbol{\Sigma}_w}^2\right\} \times |\boldsymbol{\Gamma}|^{-T/2} \exp\left\{-\frac{1}{2}\|\mathbf{J} - \boldsymbol{\Phi}\mathbf{S}\|_{\boldsymbol{\Gamma}}^2\right\}. \quad (7)$$

One can then integrate out \mathbf{J} to get the marginal distribution of the observed MEG matrix, parametrized by the TRF matrix $\boldsymbol{\Phi}$ and background covariance $\boldsymbol{\Gamma}$:

$$p(\mathbf{Y}|\boldsymbol{\Phi}, \boldsymbol{\Gamma}) \propto |\boldsymbol{\Sigma}_w + \mathbf{L}\boldsymbol{\Phi}\mathbf{L}^\top|^{-T/2} \times \exp\left\{-\frac{1}{2}\|\mathbf{Y} - \mathbf{L}\boldsymbol{\Phi}\mathbf{S}\|_{(\boldsymbol{\Sigma}_w + \mathbf{L}\boldsymbol{\Phi}\mathbf{L}^\top)^{-1}}^2\right\} \quad (8)$$

Under this model, if $\boldsymbol{\Gamma}$ were known, the TRF estimation would amount to the following optimization problem:

$$\min_{\boldsymbol{\Phi}} \frac{1}{2} \|\mathbf{Y} - \mathbf{L}\boldsymbol{\Phi}\mathbf{S}\|_{(\boldsymbol{\Sigma}_w + \mathbf{L}\boldsymbol{\Phi}\mathbf{L}^\top)^{-1}}^2. \quad (9)$$

Furthermore, to enforce temporal smoothness of the estimated TRFs, we represent $\boldsymbol{\tau}_{m,d}$ over a space spanned by Gabor atoms \mathbf{G} , i.e. $\boldsymbol{\tau}_{m,d} = \mathbf{G}\boldsymbol{\theta}_{m,d}$ or $\boldsymbol{\Phi} = \boldsymbol{\Theta}\mathbf{G}^\top$, where $\boldsymbol{\Theta} := [\boldsymbol{\theta}_{1,0}, \boldsymbol{\theta}_{1,1}, \boldsymbol{\theta}_{1,2}, \boldsymbol{\theta}_{2,0}, \dots, \boldsymbol{\theta}_{M,2}]^\top$. In order to promote temporal sparsity, the problem in (9) is replaced by the following mixed-norm regularized optimization problem:

$$\min_{\boldsymbol{\Theta}} \frac{1}{2} \|\mathbf{Y} - \mathbf{L}\boldsymbol{\Theta}\widehat{\mathbf{S}}\|_{(\boldsymbol{\Sigma}_w + \mathbf{L}\boldsymbol{\Theta}\widehat{\mathbf{S}}\mathbf{L}^\top)^{-1}}^2 + \eta \|\boldsymbol{\Theta}\|_{2,1,1}, \quad (10)$$

where $\widehat{\mathbf{S}} := \mathbf{G}^\top \mathbf{S}$, and the mixed $\ell_{2,1,1}$ norm is defined as

$$\|\boldsymbol{\Theta}\|_{2,1,1} := \sum_{m=1}^M \sum_{l=1}^L \left(\sum_{d=0}^2 \theta_{m,d,l}^2 \right)^{1/2}. \quad (11)$$

Note that the objective function in (10) is convex in $\boldsymbol{\Theta}$ and thus one can easily solve for $\boldsymbol{\Theta}$ by standard optimization techniques. But, this requires the knowledge of $\boldsymbol{\Gamma}$, which

is generally unknown. So, a suitable approximation to Γ is required. One principled way is to estimate both Θ and Γ from the observed MEG data by solving the following optimization problem:

$$\min_{\Theta, \Gamma} \frac{T}{2} \log(\Sigma_w + \mathbf{L}\Gamma\mathbf{L}^\top) + \frac{1}{2} \|\mathbf{Y} - \mathbf{L}\Theta\widehat{\mathbf{S}}\|_{(\Sigma_w + \mathbf{L}\Gamma\mathbf{L}^\top)^{-1}}^2 + \eta \|\Theta\|_{2,1,1}. \quad (12)$$

If Θ is known, the minimization in (12) is known as evidence maximization or empirical Bayes [11]. Unfortunately, the objective in (12) is not convex in Γ . In fact, the first term is a concave function of Γ . Nevertheless, there exist several EM-based algorithms [12] as well as the ‘Champagne’ algorithm [9], which aim at estimating Γ in problems similar to (12). In the next section, we present an efficient, recursive coordinate descent-based algorithm that leverages recent advances in evidence maximization and proximal gradient methods to solve the problem in (12).

IV. MAIN ALGORITHM

Since optimization with respect to both Θ, Γ is not straightforward, we instead aim at optimizing the objective in (12) by alternatingly updating Θ and Γ . The update rules at the $(r+1)^{\text{st}}$ step are given as follows:

1) Γ update: With $\Theta = \Theta^{(r)}$ fixed, the problem in (12) reduces to the following optimization problem:

$$\min_{\Gamma} \text{tr}(\Sigma_v^{-1} \mathbf{C}_v) + \log|\Sigma_v| =: \mathcal{L}_{\Theta^{(r)}}(\Gamma), \quad (13)$$

with $\mathbf{C}_v := T^{-1}(\mathbf{Y} - \mathbf{L}\Theta^{(r)}\widehat{\mathbf{S}})(\mathbf{Y} - \mathbf{L}\Theta^{(r)}\widehat{\mathbf{S}})^\top$ and $\Sigma_v := \Sigma_w + \mathbf{L}\Gamma\mathbf{L}^\top$. Although the problem is non-convex in Γ , it can be solved via another coordinate descent algorithm called Champagne [9], which solves for Γ by recursively updating a set of auxiliary variables. Though convergence to a global minimum is not guaranteed, the convergence rate is fast, the computation cost per iteration is linear in N , and most importantly, each pass is guaranteed to reduce the cost function in (13), or leave it unchanged.

2) Θ update: Minimizing over Θ while fixing $\Gamma = \Gamma^{(r+1)} = \arg \min_{\Gamma} \mathcal{L}_{\Theta^{(r)}}(\Gamma)$ is relatively straightforward. We note that the corresponding objective function

$$\frac{1}{2} \|\mathbf{L}\Theta\widehat{\mathbf{S}} - \mathbf{Y}\|_{\Sigma_v^{(r+1)}^{-1}}^2 + \eta \|\Theta\|_{2,1,1} =: \mathcal{L}_{\Gamma^{(r+1)}}(\Theta) \quad (14)$$

can be decomposed as the sum of a smooth differentiable function $f(\Theta) := \frac{1}{2} \|\mathbf{L}\Theta\widehat{\mathbf{S}} - \mathbf{Y}\|_{\Sigma_v^{(r+1)}^{-1}}^2$ and a non-smooth function $g(\Theta) := \eta \|\Theta\|_{2,1,1}$, whose proximal operator can be computed in an efficient manner. Optimization problems with this structure can be efficiently solved using an instance of the proximal gradient descent or forward-backward splitting algorithm [13], [14].

Although the objective function is not convex in (Θ, Γ) , the update steps described above together decrease the objective in (12) at every iteration, until a fixed-point or limit-cycle is reached. Intuitively, the Θ update step is akin to assigning sparse kernels to each neural sources that ‘best’ predict the MEG responses from the stimuli, allowing them to compete

Algorithm 1 TRF Localization from Multiple MEG Trials

```

1: Input: MEG observations  $\mathbf{Y}^k$  and Covariate matrices,  $\widehat{\mathbf{S}}^k$  for
    $k = 1, 2, \dots, K$ ; Lead-field matrix,  $\mathbf{L}$ ; Regularizing parameter
    $\eta$ ; initial guess  $\Theta^0$ ; tolerance  $\text{tol} \in (0, 10^{-3})$ , Maximum
   Number of outer iterations  $R_{\max} \in \mathbb{N}^+$ .
2:  $r = 0$ .
3: repeat
4:   for  $k = 1, 2, \dots, K$  do
5:     Compute  $\mathbf{C}_v^{k(r)} = \frac{1}{T} (\mathbf{Y}^k - \mathbf{L}\Theta^{(r)}\widehat{\mathbf{S}}^k)(\mathbf{Y}^k - \mathbf{L}\Theta^{(r)}\widehat{\mathbf{S}}^k)^\top$ 
6:      $\Gamma^{k(r+1)} = \arg \min_{\Gamma} \text{tr}(\Sigma_v^{-1} \mathbf{C}_v^{k(r)}) + \log|\Sigma_v|$ 
   s.t.  $\Sigma_v = \Sigma_w + \mathbf{L}^\top \Gamma \mathbf{L}$ 
7:     Compute  $\Sigma_v^{k(r+1)} = (\Sigma_w + \mathbf{L}\Gamma^{k(r+1)}\mathbf{L}^\top)^{-1}$ 
8:   end for
9:    $\Theta^{(r+1)} = \arg \min_{\Theta} \sum_{k=1}^K \frac{1}{2} \|\mathbf{L}\Theta\widehat{\mathbf{S}}^k - \mathbf{Y}\|_{\Sigma_v^{k(r+1)}^{-1}}^2 + \eta \|\Theta\|_{2,1,1}$ 
10: until  $\frac{\|\Theta^{(r+1)} - \Theta^{(r)}\|_2}{\|\Theta^{(r)}\|_2} < \text{tol}$  or  $r = R_{\max}$ .
11: Set  $r \leftarrow r + 1$ .
12: Output:  $\Theta^{(R)}$  where  $R$  is the index of the last outer iteration of
   the algorithm.

```

with one another, whereas the Γ update step adaptively changes the ‘best’ fitting noise normalization. In doing so, the algorithm decomposes the observed MEG data into stimulus-driven and stimulus-independent parts, taking into account the lead-field matrix, which in turn helps to limit the spatial leakage as we will demonstrate in the next section.

Next, consider the case where the data from K different trials corresponding to K different stimuli (i.e. different speech streams) are available. Suppose the covariate matrices and MEG observations corresponding to these trials are denoted by $\widehat{\mathbf{S}}^k$ and \mathbf{Y}^k , for $k = 1, 2, \dots, K$. Assuming that the background source activities correspond to different model parameters, Γ^k , we can extend (12) to obtain the following optimization problem:

$$\min_{\Theta, \Gamma} \sum_{k=1}^K \left[\frac{T}{2} \log(\Sigma_w + \mathbf{L}\Gamma^k\mathbf{L}^\top) + \frac{1}{2} \|\mathbf{Y} - \mathbf{L}\Theta\widehat{\mathbf{S}}^k\|_{(\Sigma_w + \mathbf{L}\Gamma^k\mathbf{L}^\top)^{-1}}^2 \right] + \eta \|\Theta\|_{2,1,1}. \quad (15)$$

In other words, we assume that the response functions remain unchanged across trials, which promotes the integration of relevant information across trials and obtaining consistent response functions. This optimization problem can also be solved via a coordinate descent method in a similar manner described earlier. The resulting algorithm including the update rules is summarized as Algorithm 1. In the spirit of easing reproducibility, a Python implementation of the algorithm is made available on the open source repository GitHub [15].

V. RESULTS

We applied the algorithm on a subset of MEG data collected from 17 adults (aged 18-27 years) under an auditory task described in [16]. During the task, the participants listened to 1 min long segments from an audiobook recording of *The*

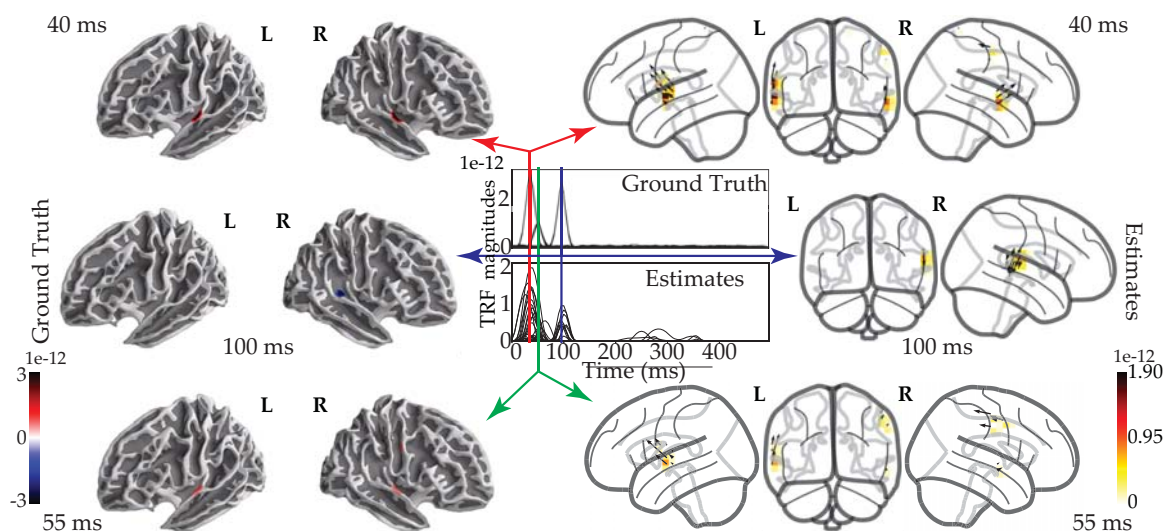


Fig. 1. TRF localization results from simulated MEG data. The left panel shows the active neural sources normal to cortical surface corresponding to ground truth TRFs (colorbar encodes intensity in each of the projections). The right panel shows the location and direction of the estimated TRFs (colorbar encodes intensity only). The middle panel compares the time-course of the estimated and ground truth TRF magnitudes.

Legend of Sleepy Hollow by Washington Irving (<https://librivox.org/the-legend-of-sleepy-hollow-by-washington-irving/>), narrated by a male speaker under different noise conditions. Here, we consider localizing the TRFs under the ‘quiet’ condition only, using a total of 6 min data from each participant. MNE-python 0.14 was used in preprocessing the raw data to automatically detect and discard flat channels, remove extraneous artifacts, and to band-pass filter the data in the range 1–80 Hz. The six 1 min long data epochs were then down-sampled to 200 Hz. As the stimulus variable, we used the speech envelope reflecting the momentary acoustic power, by averaging the auditory spectrogram representation (generated using a model of the auditory periphery [17]) across the frequency bands, sampled at 200 Hz.

The average head position over the entire recording period was used to obtain the head shape of individual subjects relative to the MEG sensors. Each subject’s head was co-registered to the ‘fsaverage’ brain using these digitized head shapes. A volume source space for individual subjects was defined on a 3D regular grid with a resolution of 7 mm in each direction. The lead-field matrix was then computed by placing free orientation virtual dipoles on the resulting 3322 grid points. From a wide range of regularizing parameters, the one resulting in the least generalization error in a 3-fold cross-validation procedure is chosen for individual subjects.

A. Application to Simulated MEG Data

To assess the performance of the proposed algorithm, we synthesized a 6 min long MEG data epoch for one of the participants from the aforementioned dataset according to the generative model (1)-(3), using the TRF kernels shown in Fig. 1. To emulate a realistic MEG experiment, we not only employed these specific time-courses and active sources, but also used MEG recordings during segment 2 while using stimulus variables of segment 1 in place of $\mathbf{W} + \mathbf{LV}$, and vice-versa. In addition, we simulated the cortical activity on a finely

discretized source space (namely, ico-5) using a direction-constrained lead-field matrix.

The 500 ms long TRF estimates obtained over the unconstrained volumetric source-space with 7 mm spacing is shown against the ground truth in Fig. 1. Not only our algorithm was able to recover the time-courses of the TRFs faithfully, but also the spatial extents of the active dipole sources closely resemble those of the simulated active regions. The most intriguing observation from Fig. 1 is that the estimated 3D TRF components closely align with the normal directions to the cortical surface, without the algorithm having any knowledge of said directions.

B. Application to Experimental MEG data

We estimated 1 s-long 3D TRFs for all 17 subjects from the aforementioned preprocessed MEG data using our proposed algorithm. The vector-valued TRF estimates were tested for consistent directionality [18] at each grid location, time bin, and across all subjects, against uniformity using a permutation test. To compensate for head misalignment and anatomical differences across subjects, the TRF estimates were first smoothed with a Gaussian kernel with standard deviation of 10 mm over the grid locations. Fig. 2 shows the group average of the estimated TRFs, masked by the significance level of $p = 0.05$.

The TRFs manifest three prominent peaks at around 20 ms, 45 ms, and 105 ms respectively, all of which are bilaterally centered at the auditory cortex (AC). Even though the two earlier peaks at 20 ms and 45 ms seem partially overlapping with nearly identical anatomical origins, the later peak exhibits increasing contribution from the inferior frontal cortex (IFC) and premotor cortex (PMC) as seen in the anatomical maps. Unlike these early responses, the peak at 105 ms seems to be much stronger at the right hemisphere and is localized slightly posterior to the earlier ones. In between the peaks, we observe responses from the PMC (at around 60 ms) and most

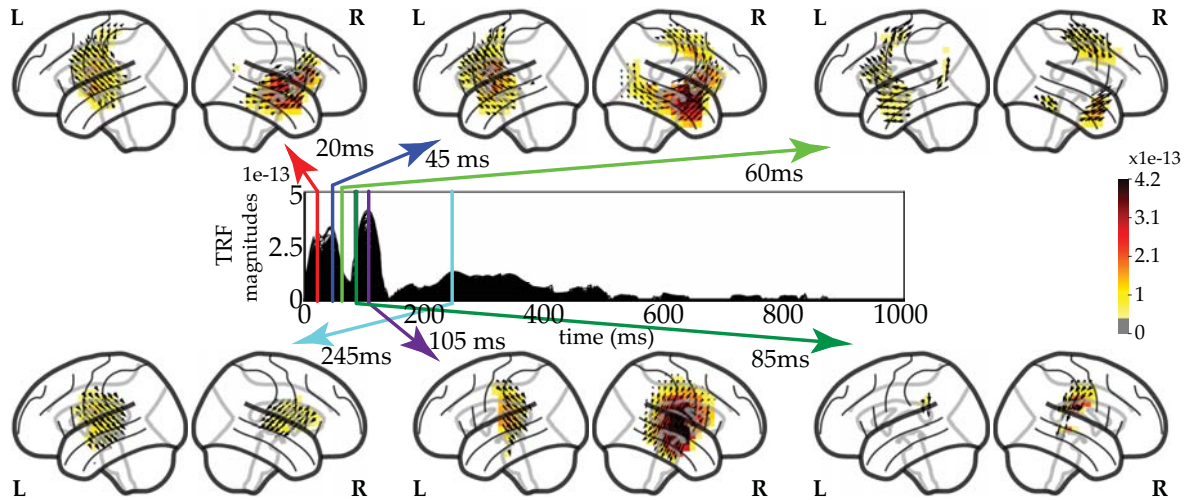


Fig. 2. TRF localization from experimentally recorded MEG data. The middle panel shows the time course of the estimated TRF magnitudes for all neural sources. Anatomical plots shows the active virtual dipoles at the visually prominent peaks, where the arrows indicate the estimated direction of the dipole activity for each source. The TRFs are masked by a significance level of $p = 0.05$ prior to visualization.

interestingly, a direction reversal at around 85 ms. It is worth noting that the spatiotemporal organization of the localized TRFs is consistent with the auditory processing stream of $AC \rightarrow IFC \rightarrow PMC \rightarrow AC$ [19].

VI. DISCUSSION AND FUTURE WORK

Characterizing neural response functions at the cortical level is key to unveiling the mechanisms of continuous information processing by segregating the contributions from different brain regions involved in complex information processing. Existing methods for this purpose perform a two-stage procedure, in which the cortical sources are first localized, followed by estimating the response functions per localized source. As a result, the localized response functions exhibit bias due to the assumptions of the underlying source localization method. In this work, we address this issue by introducing a novel framework for localizing the temporal response functions to continuous stimuli *directly* from the MEG responses in a one-stage fashion. Application of our proposed algorithm to synthetic and experimentally recorded MEG responses to continuous speech demonstrates its utility in providing new insights into the functional roles of various brain regions at different stages of auditory processing. Even though we presented our framework for a single stimulus variable (i.e., the speech envelope), our approach can be extended to incorporate multiple competing stimulus variables that can potentially explain the observed response more accurately.

REFERENCES

- [1] I. Gazzaniga and R. Ivry, *Cognitive Neuroscience: The Biology of the Mind*. New York: WW Norton and Company Press, 2009.
- [2] D. Ringach and R. Shapley, "Reverse correlation in neurophysiology," *Cognitive Science*, vol. 28, no. 2, pp. 147–166, 2004.
- [3] E. C. Lalor, B. A. Pearlmutter, R. B. Reilly, G. McDarby, and J. J. Foxe, "The VESPA: a method for the rapid estimation of a visual evoked potential," *Neuroimage*, vol. 32, no. 4, pp. 1549–1561, 2006.
- [4] E. C. Lalor, A. J. Power, R. B. Reilly, and J. J. Foxe, "Resolving precise temporal processing properties of the auditory system using continuous stimuli," *Journal of neurophysiology*, vol. 102, no. 1, pp. 349–359, 2009.
- [5] N. Ding and J. Z. Simon, "Neural coding of continuous speech in auditory cortex during monaural and dichotic listening," *Journal of neurophysiology*, vol. 107, no. 1, pp. 78–89, 2011.
- [6] C. Brodbeck, A. Presacco, and J. Z. Simon, "Neural source dynamics of brain responses to continuous stimuli: Speech processing from acoustics to comprehension," *NeuroImage*, vol. 172, pp. 162 – 174, 2018.
- [7] B. Babadi, G. Obregon-Henao, C. Lamus, M. S. Hämäläinen, E. N. Brown, and P. L. Purdon, "A subspace pursuit-based iterative greedy hierarchical solution to the neuromagnetic inverse problem," *NeuroImage*, vol. 87, pp. 427–443, 2014.
- [8] P. Krishnaswamy, G. Obregon-Henao, J. Ahveninen, S. Khan, B. Babadi, J. E. Iglesias, M. S. Hämäläinen, and P. L. Purdon, "Sparsity enables estimation of both subcortical and cortical activity from MEG and EEG," *Proceedings of the National Academy of Sciences*, p. 201705414, 2017.
- [9] D. P. Wipf, J. P. Owen, H. T. Attias, K. Sekihara, and S. S. Nagarajan, "Robust Bayesian estimation of the location, orientation, and time course of multiple correlated neural sources using MEG," *NeuroImage*, vol. 49, no. 1, pp. 641–655, 2010.
- [10] W. Wu, S. Nagarajan, and Z. Chen, "Bayesian machine learning: EEG/MEG signal processing measurements," *IEEE Signal Processing Magazine*, vol. 33, no. 1, pp. 14–36, 2016.
- [11] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*. Springer, 1985.
- [12] K. Friston, L. Harrison, J. Daunizeau, S. Kiebel, C. Phillips, N. Trujillo-Barreto, R. Henson, G. Flandin, and J. Mattout, "Multiple sparse priors for the M/EEG inverse problem," *NeuroImage*, vol. 39, no. 3, pp. 1104–1120, 2008.
- [13] Y. Nesterov, "Smooth minimization of non-smooth functions," *Mathematical programming*, vol. 103, no. 1, pp. 127–152, 2005.
- [14] T. Goldstein, C. Studer, and R. Baraniuk, "A field guide to forward-backward splitting with a FASTA implementation," *arXiv preprint arXiv:1411.3406*, 2014.
- [15] *Direct estimation of TRFs over the source space*. Available on GitHub Repository: <https://github.com/proloyd/DstRF>, 2018.
- [16] A. Presacco, J. Z. Simon, and S. Anderson, "Effect of informational content of noise on speech representation in the aging midbrain and cortex," *Journal of neurophysiology*, vol. 116, no. 5, pp. 2356–2367, 2016.
- [17] X. Yang, K. Wang, and S. A. Shamma, "Auditory representations of acoustic signals," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 824–839, March 1992.
- [18] K. V. Mardia and P. E. Jupp, *Directional statistics*. John Wiley & Sons, 2009, vol. 494.
- [19] J. P. Rauschecker and S. K. Scott, "Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing," *Nature neuroscience*, vol. 12, no. 6, p. 718, 2009.