Published in final edited form as:
Neuroimage. 2009 February 1; 44(3): 932–946. doi:10.1016/j.neuroimage.2008.05.063.

A Distributed Spatio-Temporal EEG/MEG Inverse Solver

Wanmei Ou[a], **Matti S. Hämäläinen**[b], and **Polina Golland**[a]

Wanmei Ou: wanmei@mit.edu; Matti S. Hämäläinen: msh@nmr.mgh.harvard.edu; Polina Golland: polina@csail.mit.edu

[a]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA

[b]Department of Radiology, Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Boston, MA

Abstract

We propose a novel $\ell_1\ell_2$-norm inverse solver for estimating the sources of EEG/MEG signals. Developed based on the standard ℓ_1-norm inverse solvers, this sparse distributed inverse solver integrates the ℓ_1-norm spatial model with a temporal model of the source signals in order to avoid unstable activation patterns and "spiky" reconstructed signals often produced by the currently used sparse solvers. The joint spatio-temporal model leads to a cost function with an $\ell_1\ell_2$-norm regularizer whose minimization can be reduced to a convex second-order cone programming (SOCP) problem and efficiently solved using the interior-point method. The efficient computation of the SOCP problem allows us to implement permutation tests for estimating statistical significance of the inverse solution. Validation with simulated and real MEG data shows that the proposed solver yields source time course estimates qualitatively similar to those obtained through dipole fitting, but without the need to specify the number of dipole sources in advance. Furthermore, the $\ell_1\ell_2$-norm solver achieves fewer false positives and a better representation of the source locations than the conventional ℓ_2 minimum-norm estimates.

Keywords

EEG; MEG; Inverse Solver; ℓ_1-norm; Temporal Basis Functions; Second-Order Cone Programming

1 Introduction

Electroencephalography (EEG) and magnetoencephalography (MEG) are widely used for functional brain mapping. With appropriate source estimation algorithms one can locate the activated regions, as well as estimate their dynamics. The non-invasive nature of EEG and MEG makes these methods particularly suitable for neuroscience research and clinical practice, such as surgical planning for epilepsy patients (Knake *et al.*, 2006).

Localizing activated regions from EEG/MEG data involves solving an electromagnetic inverse problem. Unfortunately, even with perfect knowledge of the electric and magnetic fields outside of the source region, this problem does not have a unique solution because there are currents which are either electrically or magnetically silent, or both. Moreover, solutions might

Correspondence to: Wanmei Ou, wanmei@mit.edu.

not depend continuously on the data without regularization, which means small errors in measurements might cause errors of arbitrary size in the estimated sources. These two characteristics make the inverse problem ill-posed in the sense of Hadamard (Hadamard, 1902). This paper introduces an integrated spatio-temporal regularizer to overcome the instabilities of standard sparse inverse solvers.

There are two main types of inverse solvers for EEG/MEG source estimation: discrete parametric solvers, also known as dipole fitting, and distributed inverse solvers. The standard dipole fitting algorithms estimate the location, orientation, and amplitudes of a fixed number of current dipoles (Wood, 1982; Scherg & Von Cramon, 1985; Mosher *et al.*, 1992; Uutela *et al.*, 1998). In contrast, distributed solvers discretize the source space into locations on the cortical surface or in the brain volume without explicitly controlling the number of current dipoles. The desired solution is computed by minimizing a cost function that depends on all sources in the source space, such as an overall minimum power or minimum current (Hämäläinen & Ilmoniemi,1984; Dale & Sereno, 1993; Wang *et al.*, 1993; Uutela *et al.*, 1999).

Dipole fitting usually provides robust estimates for activation signals, but localization is challenging when several sources are active because the associated cost function depends non-linearly on the dipole locations. Additionally, the quality of the results degrades when the assumed number of dipoles differs from the true number (see, e.g. (Wood, 1982; Hari & Forss, 1999)). Although it is possible to obtain an initial guess for the number of dipoles through singular-value decomposition (SVD) of the recordings (Huang *et al.*, 1998), this method is sensitive to the user-defined thresholds and is problematic in the presence of correlated source signals. Furthermore, it has been argued that a set of current dipoles may not be a good model for activations with relatively large spatial extents (Jerbi *et al.*, 2004).

Not restricted to a fixed number of dipoles, the distributed solvers estimate the amplitude of all possible source locations. The widely used minimum norm estimate (MNE) (Hämäläinen & Ilmoniemi,1984; Dale & Sereno, 1993; Wang *et al.*, 1993) recovers a source distribution with minimum overall energy (or minimum ℓ_2-norm) that produces data consistent with the measurements. Although the ℓ_2-norm method leads to an efficient linear inverse operator, the MNE solutions are often too diffuse. In particular, MNE is not appropriate for localization of early sensory activations and focal epilepsy, which have been shown to be focal in intracranial experiments (Barth *et al.*, 1982; Allison *et al.*, 1989). To overcome this property, the FOcal Underdetermined System Solver (FOCUSS) (Gorodnitsky & Rao, 1997) augments the MNE solver with a recursive weighting scheme. FOCUSS has been shown to be equivalent to a p-norm solver where $p \leq 1$ (Rao & Kreutz-Delgado, 1999). Other regularizers based on a norm penalty can provide bias towards sparsity. Among them, the minimum current estimate (MCE) (minimum ℓ_1-norm) is the most popular (Uutela *et al.*, 1999).

One of the drawbacks of the conventional ℓ_1-norm inverse solvers, as well as other focal solvers such as FOCUSS, is their sensitivity to noise. Similar to other distributed solvers, the conventional ℓ_1-norm solvers are typically applied to each time sample in the data separately. The solvers' sensitivity to noise causes the estimated activations to "jump" among neighboring spatial locations from one time instant to another. Equivalently, the time course at a particular location can show substantially "spiky" discontinuities when viewed over time. To avoid this problem one commonly averages the time courses across adjacent sites, at the expense of spatial resolution.

Two alternative approaches utilize temporal constraints to improve reconstruction accuracy: a direct application of the temporal constraint as a regularizer in the cost function and a use of temporal basis functions. In (Baillet & Sereno, 1997; Brooks *et al.*, 1999; Schmitt *et al.*, 2001; Galka *et al*., 2004; Zhang *et al*., 2005; Lamus *et al*., 2007), a regularizer is explicitly incorporated in the cost function to model the smoothness of the current source distributions between consecutive time instants. For example, (Baillet & Sereno, 1997) encourages small residuals in the least-squares estimates of the current sources between the current time point and the previous one. The studies of (Galka *et al*., 2004; Zhang *et al*., 2005; Lamus *et al*., 2007) propose a state-space model with smooth state transitions, i.e., the current source distributions between consecutive time instants. The temporal regularization terms in (Brooks *et al*., 1999; Schmitt *et al*., 2001) are expressed as the ℓ_2-norm of the output of the current sources passed through a pre-designed low-pass filter in the time domain. While conceptually, these methods address the problem of sensitivity to noise, their implementation requires substantial amount of computation, except for a limited number of low-pass filters. The work of (Zhang *et al.*, 2005) provides a comprehensive comparison among the above regularization methods. Taking a significantly different approach to reducing the sensitivity to noise, the vector-based spatio-temporal minimum ℓ_1-norm solver (VESTAL) projects the sample-wise ℓ_1-norm estimates onto a set of temporal basis functions (Huang *et al*., 2006). Models based on temporal basis functions have also been proposed for other types of inverse solvers. For instance, Geva (Geva, 1998) constructed a basis set using wavelets and computed inverse solutions for each basis function separately using dipole fitting. Trujillo-Barreto *et al*. (Trujillo-Barreto *et al*., 2007) explored the use of wavelets as a temporal model in the context of distributed solutions.

Similar to MCE and VESTAL, we employ the ℓ_1-norm regularizer to encourage spatial sparsity. We reduce MCE's sensitivity to noise by incorporating our knowledge of the temporal characteristics of the source signals. Specifically, we assume that the source signals are linear combinations of *multiple* temporal basis functions, and apply the distributed inverse solver to the coefficients of all basis functions simultaneously. We utilize the conventional definition of amplitude, the ℓ_2-norm, to summarize the activation strength at each location. Since the ℓ_2-norm does not encourage sparsity, many coefficients for an active location are usually non-zero in the inverse solution.

This integrated spatio-temporal regularizer is at the core of our $\ell_1\ell_2$-norm inverse solver. The $\ell_1\ell_2$-norm regularizer was suggested in farfield narrowband sensor array applications (Malioutov *et al*., 2005) to model the diffuse temporal structure of the source signals. Although we focus on the EEG/MEG application, the proposed framework is also applicable to computed tomography reconstruction, with modifications to the spatial model so as to encourage piece-wise constant solutions (i.e., ℓ_1-norm on spatial derivatives).

To summarize, the proposed solver imposes ℓ_1-norm regularization in space and ℓ_2-norm regularization in the temporal domain. The resulting inverse problem can be formulated as a second-order cone programming (SOCP) problem and solved efficiently using the interior-point method (Alizadeh & Goldfarb, 2001). In contrast to VESTAL, which uses the spatial and the temporal models separately in a two-step estimation procedure, our solver unifies the two models into a single regularizer in order to avoid error propagation from the first estimation step to the second one. Experimental comparisons in Section 4.1 reveal that the joint spatio-temporal model implicitly increases the signal-to-noise ratio (SNR) and achieves a more accurate reconstruction.

There are various ways to obtain the temporal basis functions to represent the source signals, including Fourier and wavelet decompositions. However, a compact representation of the signals, i.e., a small number of basis functions, can significantly reduce computation requirements. In this work, we generate the basis set through the singular-value decomposition of the sensor data, which often closely reflect the temporal structure of the source signals. We examine the effects of the basis selection on the resulting reconstruction by varying the singular-value cutoff and the noise amplitude in simulation experiments.

Existing ℓ_1-norm related solvers also lack a consensus in handling the free-orientation source reconstruction. In the conventional MCE (Uutela *et al*., 1999) and its cortically-constrained variant (Lin *et al*., 2006), the orientations of the sources are determined prior to invoking the ℓ_1-norm minimizer. Uutela *et al*. estimated the orientations from an initial MNE solution, while Lin *et al*. utilized both the MNE solution and anatomical information. The method proposed in (Matsuura & Okabe, 1999) alternates between computing the inverse solution and estimating the source orientation, but it suffers from convergence issues and requires intensive computations. VESTAL (Huang *et al*., 2006) applies the ℓ_1-norm to each source component via a bias-reduction scheme in the free-orientation case. Since the ℓ_2-norm is invariant with respect to rotations of the local coordinate system at each source, it is straightforward to extend our method to include free orientations, as we demonstrate in this paper.

We also construct a permutation test for the $\ell_1\ell_2$-norm inverse solution. Here, we follow the framework proposed in (Pantazis *et al*., 2005) for constructing the null hypothesis distribution. In their work, permutations can be applied either before or after the reconstruction due to linearity of the MNE inverse operator. In contrast, permutations must be performed prior to the $\ell_1\ell_2$-norm inverse operation. Although a large number of samples is required to yield an accurate estimate of the null distribution, we can still apply the permutation testing since the interior-point method for SOCP is quite efficient.

The remainder of this paper is organized as follows. Section 2 describes the $\ell_1\ell_2$-norm inverse solver. Section 3 briefly addresses implementation issues. Section 4 presents experimental results using simulated and real MEG data, followed by a discussion and conclusions.

2 Methods

In this section, we first provide background and define our notation, and then we describe our spatio-temporal model in Section 2.2. In Section 2.3, we formulate the $\ell_1\ell_2$-norm inverse solver as an SOCP problem for a fixed-orientation source model. Section 2.4 explains the temporal basis construction scheme employed by the current solver. Section 2.5 extends the solver to free-orientation cases. At the end of this section, we present a permutation test for assessing the significance of the resulting inverse solutions.

2.1 Background and Notation

Under the *quasi-static* approximation of Maxwell's equations, the observed EEG/MEG signals $\mathbf{y}(t)$ at time t are linear functions of the current sources $\mathbf{s}(t)$:

$$\mathbf{y}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{e}(t) \tag{1}$$

where $\mathbf{A}$ is the $N \times M$ lead-field matrix. $\mathbf{e}(t) \sim N(\mathbf{0}, \Sigma)$ is the measurement noise; the noise covariance Σ can be estimated from pre-stimulus data. $\mathbf{s}(t)$, $N \times 1$, and $\mathbf{y}(t)$, $M \times 1$, are vectors in the source space and the signal space, respectively. The number of sources N ($\sim 10^3 - 10^4$) is much larger than the number of measurements M ($\sim 10^2$), leading to an infinite number of solutions satisfying Eq. (1) even for $\mathbf{e}(t) = \mathbf{0}$. Without loss of generality, we apply spatial whitening based on the estimated noise covariance Σ to both the data and the lead-field matrix, leading to $\mathbf{e}(t) \sim N(\mathbf{0}, \mathbf{I})$ in the derivations.

2.2 Spatio-Temporal Model

The quasi-static assumption allows us to conduct inverse estimation for each time instant independently. However, this often results in highly variable source time courses. The large variability is particularly prominent in the focal solvers, such as the MCE, due to their non-linear nature. To mitigate this problem, we utilize the knowledge of the temporal properties of the source signals to further constrain the solution. To this end, we express the data model in Eq. (1) for all time instants as:

$$\boldsymbol{Y}=\mathbf{AS}+\mathbf{E} \tag{2}$$

where $\mathbf{Y} = [\mathbf{y}(1), \mathbf{y}(2), \dots ,\mathbf{y}(T)]$ is an $M \times T$ matrix that contains EEG/MEG measurements for all T temporal samples, and $\mathbf{S}$ is an $N \times T$ matrix that represents the source signals. Here, we assume that noise $\mathbf{E}$ is independent in time, i.e., $E[\mathbf{E}^T\mathbf{E}] = \mathbf{I}$. Time-dependent noise models as those suggested in (Huizenga *et al*., 2002; Bijma *et al*., 2005) can be incorporated into the estimation procedure as well; this is a topic of future work.

The underlying sources of EEG/MEG measurements, closely related to the postsynaptic potentials (Hämäläinen *et al*., 1993), are relatively smooth with occasional deflections. For example, a typical response from the contralateral primary somatosensory area has relatively strong deflections immediately after the stimulus (20-40 msec) followed by a smoother time course (Weerd & Kap, 1981). Hence, the activation signals are neither sparse nor diffuse in time. Direct temporal regularization using the ℓ_1-norm or the ℓ_2-norm is therefore not appropriate. To model the time-varying frequency content of the signals, we assume that the source signals are linear combinations of multiple orthonormal temporal basis functions, $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots ,\mathbf{v}_K]$, that collectively capture the temporal properties of the source signals. $\mathbf{v}_k$, $T \times 1$, denotes the k^{th} basis function. In Section 2.4, we will discuss how to obtain the basis appropriate for the reconstruction. We assume that the basis functions are orthonormal; if they are not, minor modifications of the remaining derivations are needed, as we addressed in Section 5.

Projecting both the sensor recordings and the source signals onto the basis functions, the new variables $\tilde{\mathbf{Y}} = \mathbf{YV}$ and $\tilde{\mathbf{S}} = \mathbf{SV}$ are the corresponding projection coefficients. $\tilde{\mathbf{Y}}$ and $\tilde{\mathbf{S}}$ are of size $M \times K$ and $N \times K$, respectively. The (n, k) element of $\tilde{\mathbf{S}}$, $\tilde{s}_{nk}$, indicates the k^{th} coefficient for the source signal at location n. We can rewrite the original data model in Eq. (2) in the transformed domain:

$$\tilde{\boldsymbol{Y}} =\mathbf{A}\tilde{\mathbf{S}} + \tilde{\mathbf{E}} \tag{3}$$

where $\tilde{\mathbf{E}} = \mathbf{EV}$. We use $\tilde{\mathbf{e}}_k$ to denote the k^{th} column of $\tilde{\mathbf{E}}$. The temporal independence assumption on $\mathbf{E}$ and orthonormality of $\mathbf{V}$ imply that $\tilde{\mathbf{e}}_k$ and $\tilde{\mathbf{e}}_{k'}$ are independent for $k \neq k'$ and that $\tilde{\mathbf{e}}_k \sim N(\mathbf{0}; \mathbf{I})$. Eq. (3) is still under-determined, containing MK equations with NK variables.

To compute inverse solutions for all K basis functions simultaneously, we extend the existing regularizers to use the signal magnitude in the subspace spanned by $\mathbf{V}$, $\sqrt{\sum_{k=1}^{K} \tilde{s}_{nk}^2}$, as an indicator of the activation status for location n. In other words, we apply ℓ_2-norm regularization to the K coefficients for each source location. Because we choose to work with orthonormal basis functions, the ℓ_2-norm of the reconstructed source signal in the temporal domain is equal to the ℓ_2-norm in the transformed domain. However, we find it more intuitive to present the model in the transformed domain.

In addition, we assume that the sources exhibit a spatially sparse pattern. This assumption represents the relatively compact source regions typically activated in the sensory areas. To obtain a focal inverse solution, we should ideally employ the ℓ_0-norm as the spatial regularizer. However, the ℓ_0-norm regularization leads to an NP-hard optimization problem. In practice, under some regularity conditions (Donoho & Elad, 2003), the ℓ_1-norm regularizer leads to solutions identical to those produced by the ℓ_0-norm regularizer. Even when the solution obtained through the ℓ_1-norm regularization is different from the one produced by the ℓ_0-norm regularization, it is still more sparse than that obtained with the ℓ_2-norm regularizer.

With the ℓ_1-norm regularizer in the spatial domain and the ℓ_2-norm regularizer in the temporal domain, we incorporate the integrated spatio-temporal $\ell_1\ell_2$-norm regularizer

$$| \tilde{\mathbf{S}} |_{\ell_1}^{\ell_2} = \sum_{n=1}^{N} \sqrt{\sum_{k=1}^{K} \tilde{s}_{nk}^2} \tag{4}$$

into the estimation problem:

$$\tilde{\mathbf{S}}^* = \arg\min_{\tilde{\mathbf{S}}} \| \tilde{Y} - \mathbf{A}\tilde{\mathbf{S}} \|_F^2 + \lambda | \tilde{\mathbf{S}} |_{\ell_1}^{\ell_2} \tag{5}$$

$$= \arg\min_{\tilde{\mathbf{S}}} \sum_{k=1}^{K} \| \tilde{\mathbf{y}}_k - \mathbf{A}\tilde{\mathbf{s}}_k \|_{\ell_2}^2 + \lambda | \tilde{\mathbf{S}} |_{\ell_1}^{\ell_2} \tag{6}$$

where $\tilde{\mathbf{s}}_k$ and $\tilde{\mathbf{y}}_k$ are the k^{th} column vectors in $\tilde{\mathbf{S}}$ and $\tilde{\mathbf{Y}}$. $\| \cdot \|_F$ and $\| \cdot \|_{\ell 2}$ (i.e., $\|\mathbf{x}\|_{\ell_2} = \sqrt{\mathbf{x}^T\mathbf{x}}$) denote the Frobenius norm of a matrix and the standard ℓ_2-norm of a vector, respectively. λ controls the regularization strength. We will discuss how to select this parameter in Section 4.1.4. After we obtain the optimal coefficients $\tilde{\mathbf{S}}^*$, the reconstructed source signals are linear combinations of the temporal basis functions:

$$\mathbf{S}^* = \tilde{\mathbf{S}}^* \mathbf{V}^{\mathbf{T}}. \tag{7}$$

In this paper, we formulate the inverse problem as a regularized optimization. It also has an equivalent Bayesian interpretation. The first term in Eq.(5) can be considered as the negative log likelihood under white Gaussian noise. The second term corresponds to the negative log prior of the source signals, which in our case is Laplacian in space and Gaussian in time.

2.3 From the $\ell_1\ell_2$-Norm Regularizer to Second-Order Cone Programming (SOCP)

We cannot directly apply gradient based methods to the optimization problem specified by Eq. (6) since the $\ell_1\ell_2$-norm penalty term is not differentiable at zero. However, Eq. (6) can be reduced to the SOCP problem by converting the original unconstrained optimization problem to a constrained one:

$$<\tilde{\mathbf{S}}^*, q^*, z^*, \mathbf{w}^*, \mathbf{r}^*> = \arg\min_{<\tilde{\mathbf{S}}, q, z, \mathbf{w}, \mathbf{r}>} (q + \lambda z) \quad (8)$$

$$\text{s.t. } \|\tilde{\mathbf{y}}_k - \mathbf{A}\tilde{\mathbf{s}}_k\|_{\ell_2}^2 \le w_k \forall k = 1, \cdots, K \quad (9)$$

$$\sum_{k=1}^{K} w_k \le q \quad (10)$$

$$\sqrt{\sum_{k=1}^{K} \tilde{s}_{nk}^2} \le r_n \forall n = 1, \cdots, N \quad (11)$$

$$\sum_{n=1}^{N} r_n \le z \quad (12)$$

New variables, q, z, $\{w_k\}_{k=1}^{K}$, and $\{r_n\}_{n=1}^{N}$, are introduced in the conversion procedure. w_k is an upper bound on the discrepancy between the measurements and the signals predicted by the estimated sources in the projection onto $\mathbf{v}_k$. q is an upper bound on all w_k's. r_n is an upper bound on the activation strength for location n. z is an upper bound on the ℓ_1-norm of the activation strength of all N locations. At the minimum, the inequality constraints in Eq. (9)-(12) are satisfied with equality; otherwise, the objective function can be further reduced.

Mathematically, a second-order cone of dimension D is defined as

$$Q_D := \left\{ \mathbf{x} = (x_0, \bar{\mathbf{x}}) \in R^D : x_0 \ge \|\bar{\mathbf{x}}\|_{\ell_2} \right\} \quad (13)$$

where x_0 and $\bar{\mathbf{x}}$ denote the first element and the remaining elements of $\mathbf{x}$, respectively. We can see that Eq. (11) matches with the second-order cone definition. As shown in (Alizadeh & Goldfarb, 2001), a wide range of constrained formulations, including the quadratic constraint in Eq. (9), can be reduced to the canonical form of a second-order cone. For completeness, we provide the corresponding derivations in Appendix A.

An SOCP problem can be expressed in the canonical form that contains a linear objective function and the feasible set specified as an intersection of an affine linear manifold and the Cartesian product of second-order cones. Since the second-order cone defines a convex set, the feasible set of SOCP is convex. Therefore, SOCP is a convex optimization problem and its local minimum is the global minimum. In fact, for one-dimensional and two-dimensional cones, the second-order cone constraint in Eq. (13) can be reduced to linear constraints. As a result, the corresponding SOCP problem is reduced to a linear programming problem. It is also straightforward to show that the quadratically constrained quadratic programs are a subset of the SOCP problems. Furthermore, the SOCP problem is a special case of a semi-definite program. Therefore, SOCP can be solved efficiently using the primal-dual interior-point method (Alizadeh & Goldfarb, 2001), where Newton's method is employed to reduce the duality gap. Appendix B reviews the primal-dual interior-point method in application to SOCP.

2.4 Temporal Basis Selection

The formulation of our inverse solver is independent of the selected basis $\mathbf{V}$, but a compact representation of the signals can significantly reduce computation. We estimate the basis using the singular-value decomposition (SVD) of the measurements, which is often able to compactly capture the time-varying frequency content and significant differences in source signals from different regions. Another advantage of using data-adaptive temporal basis functions is that it avoids the difficulty of setting a set of basis functions to accommodate highly variable source signals due to different experimental tasks and subject-to-subject variations. Since according to Eq. (2) the sensor signals are linear combinations of the source signals, the temporal pattern of the source signals is present in the sensor signals as well. In fact, the standard dipole fitting procedure (Mosher *et al*., 1992) also performs fitting of the K largest SVD components of the measurements that "adequately" describe the data.

The singular-value decomposition of data $\mathbf{Y}$ is expressed as

$$\boldsymbol{Y}=\mathbf{U}_{Y}\boldsymbol{\Lambda}_{Y}\mathbf{V}_{Y}^{\mathrm{T}} \tag{14}$$

each column in $\mathbf{U_Y} = [\mathbf{u}_1, \mathbf{u}_2, \dots , \mathbf{u}_M]$ denotes the electromagnetic field pattern; each column in $\mathbf{V_Y} = [\mathbf{v}_1, \mathbf{v}_2, \dots , \mathbf{v}_M]$ denotes the temporal pattern. $\boldsymbol{\Lambda}_\mathbf{Y}$ is a diagonal matrix of the singular values in a descending order. We assume more temporal samples than EEG/MEG sensors, which is usually true in practice due to fast sampling rates.

As mentioned before, we further assume that activation signals only lie in the subspace spanned by $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots , \mathbf{v}_K]$, but not in the subspace spanned by $\mathbf{V}^{\perp} = [\mathbf{v}_{K+1}, \mathbf{v}_{K+2}, \dots , \mathbf{v}_M]$. Performing reconstruction in the signal subspace helps to stabilize the reconstructed source signals since they are constructed as linear combinations of relatively smooth temporal basis functions. In our experiments, the number of basis functions K is fixed in each reconstruction. We examine the performance of the proposed solver with varying K in Section 4.1.3 and discuss alternative approaches to basis function selection in Section 5.

2.5 $\ell_1\ell_2$-norm Reconstruction for the Free-Orientation Source Model

The free-orientation source model has been used both with volumetric source spaces covering the entire brain and with source locations restricted to the cortex only (Dale & Sereno, 1993) (Pascual-Marqui *et al*., 1994). The results of the direct application of the ℓ_1-norm regularizer to the three dipole moment coordinates depend on the parametrization of the local coordinates.

To extend our solver to free-orientation sources, we model the current dipole moment magnitude as the ℓ_2-norm of the current dipole moments along the three coordinates. This model agrees with the conventional definition of magnitude. The resulting inverse problem is independent of the local coordinate system since the ℓ_2-norm is invariant to rotations of orthogonal coordinates. In other words, our method models the spatially sparse activation pattern, but does not enforce sparsity on individual components of the dipole moments. This idea is analogous to the sensor array application (Malioutov *et al*., 2005) where signals are complex numbers; it was also independently developed and thoroughly evaluated by Ding and He for EEG source localization (Ding & He, 2007).

Extending our formulation in Eq. (4), $\tilde{s}_{nk}$ is replaced by a three-dimensional vector, denoting the current dipole moments in the three coordinates $\tilde{s}_{x_{nk}}$, $\tilde{s}_{y_{nk}}$, and $\tilde{s}_{z_{nk}}$. The optimization problem in Eq. (8)-(12) remains the same except that Eq. (11) is replaced by a constraint on the three coordinates:

$$\sqrt{\sum_{k=1}^{K}\left(\tilde{s}_{x_{nk}}^{2}+\tilde{s}_{y_{nk}}^{2}+\tilde{s}_{z_{nk}}^{2}\right)} \leq r_n \forall n=1:N \tag{15}$$

In the original problem, each cone specified in Eq. (11) lies in a K + 1-dimension subspace; in the free-orientation case, the corresponding cone is extended to a 3K + 1-dimension subspace. Since the feasible region is an intersection of hyper-cones and hyper-planes, the new formulation is still consistent with the SOCP structure.

2.6 Statistical Significance Test

The non-linear nature of the ℓ_1-norm related inverse operators, including the $\ell_1\ell_2$-norm inverse solver, presents a challenge in obtaining a sufficient statistic for hypothesis testing. Since there is no closed-form solution for the $\ell_1\ell_2$-norm solver, we employ a permutation test. We construct the null distribution by permuting equal-length pre-stimulus and post-stimulus single-trial recordings. Under the null hypothesis defined as the absence of activation, the pre-stimulus and the post-stimulus recordings are equivalent. As described in (Pantazis *et al*., 2005), in each permutation, we randomly select trials; for each selected trial, we swap its pre-stimulus and post-stimulus recordings. Then we apply the inverse solver to the average data. This procedure preserves the noise covariance structure.

All the results presented in this paper are based on 5000 permutations. In this work, we control the false discovery rate (FDR) (Genovese *et al*., 2002; Efron & Tibshirani, 2002), over an amplitude-normalized source space. We first convert the source estimates into *p*-values, the ratio of permutations whose corresponding amplitude exceeds the original estimate, for each vertex and for each time instant separately. We then compute the FDR threshold (Genovese *et al*., 2002). We choose to use the *p*-values instead of the estimated amplitudes since source strength varies among activation regions and varies over time. For instance, the contralateral primary somatosensory (cSI) response is usually substantially larger than that in the ipsolateral secondary somatosensory (iSII) region. In addition, the N20 deflection is often weaker than the later deflections from the cSI area.

We cannot directly compare the activation maps created based on the permutation method with the corresponding statistics for MNE, the dynamic statistical parametric map (dSPM) (Dale & Sereno, 1993), because dSPM typically exhibits higher statistical power due to the quite restrictive Gaussian distribution assumption. On the other hand, the permutation method can capture activations for which the Gaussian assumption is not valid. Since dSPM is one of the most popular estimates in the EEG/MEG inverse community, we visually compare our results with dSPM side-by-side in the experimental section.

3 Implementation

3.1 Source Space and Lead-Field Matrix

For the computation of the lead-field matrix, we need a specification of the conductivity structure of the head, i.e., the forward model and the source space. In the forward computations for MEG, we employ the single-compartment boundary-element model (Hämäläinen & Sarvas, 1989) (Oostendorp & Van Oosterom, 1989). For the source space, we restrict the locations of the sources to the cortical surface, which, in this work, is extracted using Freesurfer (Dale *et al*., 1999; Fischl *et al*., 1999). Due to the organization of the cortex, we can further constrain the source orientation to be perpendicular to the cortical surface. Independent of the choice of source space resolution, the orientation at each vertex is computed from the original triangulation of the cortical surface with a 0.65-mm grid spacing. Similar to other inverse solvers with orientation constraint, the sparse spacing of the source space may result in localization errors (Lin *et al*., 2006) which could be avoided by denser sampling. Moreover, it is straightforward to allivate this effect in our method by applying the free-orientation model presented in Section 2.5.

In practise, the lead-field matrix $\mathbf{A}$ is often ill-conditioned. That means some of its M singular values are close to zero. It is common to improve the conditional number of $\mathbf{A}$ by employing the truncated SVD regularization. We use $\mathbf{A}^{(m)}$, a rank-m approximation of $\mathbf{A}$ (Kaipio & Somersalo, 2004). In our experience, the inverse solutions obtained using $\mathbf{A}$ and $\mathbf{A}^{(m)}$ are almost identical, which reflects the robustness of our solver. Working with $\mathbf{A}^{(m)}$ further reduces the number of variables in the optimization problem by reducing M to m and significantly accelerates computations. Therefore, all the results reported in this article are based on $\mathbf{A}^{(m)}$ with $m = 100$. On the other hand, to obtain realistic simulated data, the forward calculations of the simulated signals are based on the full matrix $\mathbf{A}$.

3.2 Pre-processing for Temporal Basis Function Construction

Due to different types of sensors, gradiometers and magnetometers in MEG and electrodes in EEG, the measurements have different units and different ranges of recordings. To construct a set of temporal basis functions, we must first whiten the measurements in the sensor space according to the estimated noise covariance matrix. Without this whitening procedure, some subsets of the sensor recordings, such as the magnetometers, would have been ignored in the construction of basis functions. In addition, we need to exclude eventual stimulus artifacts when computing the SVD of $\mathbf{Y}$; otherwise, most of the basis functions in $\mathbf{V}_{\mathbf{Y}}$ would mainly explain the artifacts. For example, in our analysis of median-nerve experiments, measurements from the first 5 msec after the stimulus onset are excluded in the basis function construction.

3.3 Multi-Resolution Approach

In this work, the estimates of the source locations are confined to a mesh. In order to reduce computational complexity, we employ a multi-resolution scheme. We first perform source estimation on a coarse mesh, then we adaptively refine the mesh around the activation regions. In other words, the forward model at a high resolution level includes all the vertices at one level below and the newly introduced vertices around the detected regions.

The ℓ_1-norm regularization often produces focal estimates, which is more appropriate to model activations in the sensory regions. However, for a spatially extended source, the corresponding source estimates may appear as several activated vertices in the extended patch if the estimation is conducted on a much finer mesh (Uutela *et al*., 1999). In most of our experiments, we used a double resolution scheme, 20- and 10-mm spacing between vertices. Our reconstruction results for the median-nerve experiments show that the activations in the primary sensory cortex can be accurately represented using this multi-resolution scheme. For auditory experiments, where the sources are slightly more diffuse, our solver detects several adjacent vertices in the auditory areas.

3.4 Computation Requirements

Compared with the MCE, which solves T linear programming problems with N variables each, our solver performs a single SOCP optimization over NK variables. As described in Eq. (8) and Appendix A, we increase the number of variables to approximately $(N + M)K$ in order to convert the $\ell_1\ell_2$-norm solver into the SOCP formulation. In this work, we use the Self-Dual-Minimization software package (SeDuMi) (Sturm *et al*., 2001) that implements the primal-dual interior-point method with logarithm barrier functions, to solve for the SOCP problem. The primal-dual interior-point method, employed in SeDuMi, has run time $\Theta(((N + M)K)^3)$ per iteration. It converges within thirty iterations in most of our experiments. For $M = 100$ (with the truncated SVD-regularized lead-field matrix), $K = 3$, and $N \approx 500$, our current implementation takes about 10 seconds with a standard PC (2.8 GHz CPU and 8 GB RAM) to compute the inverse solution. When $N \approx 2000$, it takes about 100 seconds. Combining with a two-level multi-resolution scheme, with $N \approx 500$ for the first level and $N \approx 700$ for the second level, it takes 25 seconds. When K increases to 6, the multi-resolution scheme takes about 100 seconds.

4 Results

Due to a lack of ground truth in real EEG/MEG experiments, we first study the behavior of the method and its sensitivity to parameter settings and to noise using simulated data. We then compare the method to standard inverse solvers using real MEG data from somatosensory and auditory studies.

4.1 Simulation Studies

To simulate MEG measurements, we created active vertices A, B, and C (Fig. 1, top) on the cortical sheet at source spacing of 20 mm, with current source orientation along the normal to the cortical surface. In all experiments in this paper, we scaled the reconstructed surfaces to 10^5-mm^2 surface area per hemisphere. Vertex A is located in the lateral frontal region, Vertex B is located at the pre-central gyrus, and Vertex C is located at the Sylvian fissure. The time courses of these three vertices are shown in Fig. 2a-c (black solid curves). We chose the signals to have similar temporal characteristics to those of the auditory evoked responses, but with temporal translation and scaling. The source signals of vertices B and C show activation during overlapping time intervals, which makes the inverse problem difficult.

For the forward calculations, we employed the sensor configuration of the 306-channel Neuromag VectorView MEG system (204 gradiometers and 102 magnetometers) used in our experimental studies. The location of the array with respect to the head and the noise covariance matrix were obtained from real MEG experiments. A single-compartment homogeneous forward model was employed. With Gaussian noise added, the resulting signals have an SNR = 3 dB, where the SNR is defined as $10\log_{10}\frac{\|\mathbf{AS}\|_F^2}{MT\sigma^2}$, where σ^2 is the noise variance. The resolution of the source space is relatively coarse; nevertheless, this example serves as a good illustration for the method.

In the inverse estimation, we fixed the orientation of the estimated currents to be perpendicular to the cortical mesh. Fig. 1b depicts the inverse solutions at three time frames obtained from the $\ell_1\ell_2$-norm solver using three basis functions and $\lambda = 10^9$. The parameter values were selected based on our validation experiments presented in Section 4.1.3 and 4.1.4. Curves marked with 'o' in Fig. 2a-c correspond to the source signals estimated by the method. The resulting spatial maps and source time courses match well with the ground truth.

4.1.1 $\ell_1\ell_2$-norm vs. ℓ_1-per-coefficient—To further explore the behavior of the $\ell_1\ell_2$-norm regularizer, we compared its reconstructions with those obtained by applying the ℓ_1-norm regularizer to the coefficients of each basis function separately. We will refer to this solver as ℓ_1-per-coefficient. This comparison reveals the effect of the ℓ_2-norm regularization for all coefficients. For each basis function, ℓ_1-per-coefficient computes the least ℓ_1-norm solution for the coefficients independent of other basis functions. It also achieves stable reconstruction due to the use of temporal basis functions. We applied ℓ_1-per-coefficient, also employing the interior-point method implemented in (Sturm *et al.*, 2001), to the data described above. Fig. 1c depicts the reconstruction results, and the corresponding time courses of the three active vertices are presented in Fig. 2a-c (marked with '+'). Both $\ell_1\ell_2$-norm and ℓ_1-per-coefficient detect activations in the three vertices (Fig. 1), but the $\ell_1\ell_2$-norm solution contains fewer false positive activations than that obtained through the ℓ_1-per-coefficient method. Moreover, the reconstructed time courses of the $\ell_1\ell_2$-norm solution match the ground truth time courses slightly better than those of the ℓ_1-per-coefficient solution.

The three selected basis functions are illustrated in Fig. 2d, and the reconstructed coefficients from these two algorithms are shown in Fig. 3. The projection coefficients of the simulated signals are also presented for comparison purpose (Fig. 3a). Although the spatial pattern of the projection coefficients are similar for the two methods, the spatial pattern for all three coefficients is more sparse in the $\ell_1\ell_2$-norm solution (Fig. 3b) than in the ℓ_1-per-coefficient solution (Fig. 3c). Since ℓ_1-per-coefficient models the coefficients of each basis function separately, vertex with large coefficient for one basis function may have zero coefficient for another basis function. On the other hand, $\ell_1\ell_2$-norm considers all coefficients jointly in sparsity determination. This method is particularly helpful for basis functions which have a smaller SNR, such as $\mathbf{v}_2$ and $\mathbf{v}_3$. That is illustrated by a more sparse distribution of the coefficients in Fig. 3b than that in Fig. 3c, a missing $\mathbf{v}_2$ component for Vertex A (Fig. 3c), and a false detection for a vertex close to Vertex B (Fig. 3c). The ℓ_2-norm regularizer essentially helps bundle basis functions $\mathbf{v}_2$ and $\mathbf{v}_3$ with those that are aligned well with the signal subspace, such as $\mathbf{v}_1$, to jointly determine an activation map. Therefore, we can see that sparsity defined by all coefficients is more suitable for the current basis construction method in conjunction with complex neural signals.

Fig. 2e presents the projection coefficients of the simulated source signals onto the temporal basis functions. The coefficients that correspond to basis functions $\{\mathbf{v}_k : k \geq 4\}$ are close to zero. We only displayed the coefficients corresponding to the first ten basis functions. For this simple example, Fig. 2e verifies that the selected basis well approximates the signal subspace of the simulated signals.

4.1.2 Comparison with MNE, MCE, and VESTAL—We also compared the proposed method with the standard MNE, MCE, and VESTAL (Fig. 4). The estimates from the standard MNE are smaller than the simulated signals, and it is caused by the diffuse property of the ℓ_2-norm regularization. The estimated time courses from MCE exhibit substantially "spiky" discontinuities due to the solver's sensitivity to noise. Projecting MCE results to a set of basis functions, VESTAL removes the discontinuities; however, the amplitude of the estimated time courses is smaller than the true activation signals since the two-step procedure cannot fully compensate for the errors in the original MCE solutions. Therefore, in the rest of the stimulated experiments, we focus on the performance of $\ell_1\ell_2$-norm and ℓ_1-per-coefficient.

4.1.3 Sensitivity to Noise and Basis Selection—To examine the sensitivity of the methods to noise and basis selection, we computed inverse solutions for 100 independently generated data sets for each noise setting (varying from SNR 1 dB to 8 dB) and basis selection cutoff (*K* varying from 1 to 6). The relative mean-square error (MSE) [1] for the three active vertices and for all vertices of the ℓ_1-per-coefficient and the $\ell_1\ell_2$-norm inverse solutions are shown in Fig. 5.

[1]We define the relative MSE as $\dfrac{\|\text{reconstructed signals} - \text{ground truth signals}\|_F^2}{\|\text{ground truth signals}\|_F^2}$.

The $\ell_1\ell_2$-norm outperforms ℓ_1-per-coefficient under all SNR settings and basis selection cutoffs we examined. The improvement of the relative MSE varies from 4% to 10%, with larger improvement for noisier data. The large improvement in the low SNR cases again demonstrates the importance of the ℓ_2-norm regularization on the coefficients of the representation. The standard deviation of the reconstructions estimated over the 100 simulated data sets is similar for the two solvers. It varies between 0.3% and 2.5% for $K \geq 2$ and all three selected SNR settings. For $K = 1$, the standard deviation is between 5% and 10% due to more variability in representing the signals using a single temporal basis function.

In general, both solvers achieve the best performance for $K = 3$ basis functions. If the chosen number of basis functions is too high, some basis functions represent noise, resulting in slight degradation of the result quality as reflected by the gentle slope on the logarithmic scale. On the other hand, the ℓ_1-per-coefficient's performance is not affected by including too many basis functions, because its estimated sources from the noisy basis functions are usually small. Including too few basis functions leads to a significant loss of signals; both solvers fail to recover the missing signals.

4.1.4 Sensitivity to Regularization Strength—We also investigated the methods' sensitivity to the value of the regularization parameter λ. Large λ corresponds to a high penalty on the strength of the current sources, in terms of the $\ell_1\ell_2$-norm; small λ emphasizes the data fidelity term. Due to whitening, the first term in Eq. (5) is on the order of MK, where M is on the order of 10^2. For an activated vertex in our experiments, $\tilde{s}_{nk}$ is on the order of 10^{-8} Hence, $|\tilde{\mathbf{S}}|_{\ell_1}^{\ell_2}$ is approximately $10^{-7}K$. Therefore, $\lambda = 10^9$ roughly balances between the data fidelity and the regularization terms in Eq. (5). In the experiments using real MEG data, we set $\lambda = 10^9$. Since the values in the data fidelity and regularization terms are both linearly proportional to K, the regularization strength should be independent of the number of basis functions participating in the inverse calculation. That means the sensitivity of ℓ_1-per-coefficient and $\ell_1\ell_2$-norm to λ should be the same. Hence, we only present the relative MSE obtained using $\ell_1\ell_2$-norm for all vertices and for the three active vertices for various values of λ (Fig. 6). As we can see, λ around 10^9 provides accurate reconstruction results. The regularization shows no effect for $\lambda < 10^3$; when $\lambda > 10^{10}$, the data fidelity term is effectively ignored in the optimization process. For $\lambda = 10^9$, the standard deviation of the MSE estimated from the 100 data sets is less than 1%.

4.1.5 Different Spatial Resolutions—To further examine the $\ell_1\ell_2$-norm inverse solutions at different spatial resolutions, we extended the simulated sources at vertices A, B, and C, described in Section 4.1.1, to three patches at 0.65 mm tessellation resolution. The patches have spatial extent of approximately 15-mm in diameter (200 to 500 vertices at a 0.65 mm resolution), indicated by the colored patches in the first row of Fig. 7. The ground truth source signals are identical to those employed in Section 4.1.1, shown as solid curves in Fig. 7a. To generate the sensor signals, we added Gaussian noise with covariance matrix estimated from the pre-stimulus recordings of a real MEG data set, with a resulting SNR = 3 dB. To avoid an "inverse crime," the inverse solutions were calculated at the resolution lower than the resolution used in the simulations, including a single-level mesh at 20 mm, a two-level multi-resolution scheme at 20 and 10 mm, and a three-level multi-resolution scheme at 20, 10, and 5 mm. We set $\lambda = 10^9$ and $K = 3$ in this experiment.

Fig. 7 shows the inverse solution from each of the three multi-resolution schemes at 30, 65, and 72 msec, corresponding to the peaks of the three simulated source signals. The detected areas are either in bluecyan or red-yellow corresponding to current flowing inward or outward with respect to the cortex. Each reconstruction was thresholded such that all three areas were detected at their peak times. Good performance is indicated by fewer false positives. A smaller amplitude in the dipole fitted time course for patch C (Fig. 7a row one) indicates that some of the vertices in this patch have current orientations silent with respect to the MEG sensors. Smaller amplitudes in the reconstructed time courses from our solver are expected due to the magnetically silent sources, as well as the distributed nature of the model where some nearby vertices are detected despite the regularization promoting spatial sparsity. We scaled all the reconstructed time courses by a factor of four for illustration purposes.

We observe that 20 mm resolution is too coarse for reconstruction, as reflected by more ambiguity in the source locations. The reconstruction results at the 5 mm resolution are too focal. At this reconstruction resolution, the vertices in a simulated patch are close, and some of them have similar orientation. Mathematically speaking, the signal distribution corresponding to those vertices, indicated by the column vectors in the lead-field matrix, are almost linearly dependent. The ℓ_1-norm encourages spatial sparsity, and it usually allocates all source current to one of those vertices. Reconstruction at the 10 mm resolution achieves the most accurate results: fewer false positives and a better representation of the spatial extent of the simulated patches. Therefore, we employed the two-level multi-resolution scheme in the experiments using real MEG data.

4.2 Real MEG Experiments

Experiments with synthetic data reveal the potential of the $\ell_1\ell_2$-norm solver to provide accurate and stable solutions when handling focal and correlated sources, even in a noisy environment. Next, we compare the performance of the solver to the MNE and dipole fitting using two real MEG data sets from median-nerve and auditory experiments. Both experiments were acquired using a 306-channel Neuromag VectorView system. The anatomical images were collected with a Siemens Avanto 1.5 T scanner with a T1-weighted sagittal MPRAGE protocol, which were employed for cortical surface reconstruction (Dale *et al*., 1999; Fischl *et al*., 1999). A multi-echo 3D Flash acquisition was performed to extract the inner skull surface for the boundary-element model (Hämäläinen & Sarvas, 1989; Oostendorp & Van Oosterom, 1989; Hämäläinen, 2005). Informed consent in accordance with the Massachusetts General Hospital ethical committee was obtained from subjects prior to participation.

4.2.1 Median-Nerve Experiments—We present results for one subject, a 40-years old male, in the study. The median nerve was stimulated at the left wrist according to an event-related protocol, with a random inter-stimulus-interval ranging from 1.5 to 2 seconds. Data were acquired at sampling rate of 2 KHz; a 200-msec baseline before the stimulus was used to estimate the noise covariance matrix. Approximately 300 trials remained after rejecting trials with eye-movements and other artifacts [2], from which we computed the average signal used as the input to the inverse solvers. We first applied baseline correction and whitened the data spatially based on the pre-stimulus measurements. For the $\ell_1\ell_2$-norm solver, we used six basis functions shown in Fig. 8. SVD was performed on signals between 6 msec and 200 msec after stimulus onset to avoid post-stimulus artifacts.

It has been shown that the median nerve stimulus activates a complex cortical network (Hari & Forss 1999), The first activation of the contralateral primary somatosensory cortex (cSI) peaks around 20 msec and continues over 100 msec; then the secondary somatosensory cortex (SII) activates bilaterally around 70 msec and lasts up to 200 msec. Whether SI and SII form a sequential or parallel architecture is still a topic of active debate (Kass *et al*., 1979; Rowe *et al*., 1996). The posterior parietal cortex (PPC), located on the wall of the post-central sulcus, medial and posterior to the SI cortex hand area, activates around 70-110 ms. This area, also

[2]Trials with peak-to-peak amplitude of the EOG signals exceeding 150 μV, gradiometer signals exceeding 3000 fT/cm, or magnetometer signals exceeding 3.5 pT were rejected. These rejection criteria are the same for the auditory experiment.

known as the parietal association area, most probably functions as an integrator between sensory and motor processing. Although the SI-SII network exhibits robust activation, there is significant variation from subject to subject especially in the time courses of SII activations.

In this experiment, we controlled FDR at 0.05, computed from 5000 permutations as described in Section 2.6. We also compared our results with the MNE computed using a standard software package (Hämäläinen, 2005). In practice, experts often interpret MNE through its statistics, dSPM, with manually adjusted thresholds. For the purpose of the comparison, we selected the threshold for dSPM so that all four regions of interest, cSI, cSII, iSII, and PPC, were included.

Fig. 9 presents the activation maps obtained using $\ell_1\ell_2$-norm and MNE. At 20 msec, $\ell_1\ell_2$-norm pinpoints cSI on the postal wall of the central sulcus. MNE produces a more diffuse solution leading to false positives in the post-central sulcus. The $\ell_1\ell_2$-norm clearly demonstrates change of polarity in cSI, reflected by the change of current direction between 20 and 35 msec. The polarities estimated using the $\ell_1\ell_2$-norm solver agree with the literature (Wikström *et al*., 1996): outwards intra-cellular current at 20 msec associated with N20 and inwards intra-cellular current at 35 msec associated with P35. At 75 msec, both MNE and $\ell_1\ell_2$-norm capture signals from cSII. $\ell_1\ell_2$-norm successfully localizes PPC at the post-central sulcus, but the location of PPC is ambiguous in the MNE results. According to the findings reported in (Forss *et al*., 1994;Hari & Forss, 1999), the signals from iSII is weaker than those from cSII. By controlling FDR at 0.05, the $\ell_1\ell_2$-norm solver detects iSII activation at 85 msec, but places it at the superior temporal lobe instead of the inferior parietal lobe. As shown in the volumetric display (Fig. 11), these two regions are very close, making the inverse problem challenging. MNE also presents weak iSII signals; the location is ambiguously spread between the iSII region and the superior temporal lobe.

We estimated the current source dipoles and their corresponding time courses through the standard dipole fitting procedure (Nelder & Mead, 1965; xfit software). Dipole fitting was performed using the corresponding channels at 20-38 msec, 75 msec, and 85 msec after the stimulus onset. The source estimates are summarized in one map as illustrated in Fig. 10, and the corresponding time courses are presented in Fig. 12(b). Dipole fitting did not correctly localize PPC from these data because PPC is very close to cSI. The locations for cSI and cSII identified by our solver (Fig. 9) match with the dipole fitting results. The correct localization of iSII using dipole fitting required manual intervention in selecting appropriate channels in contrast to the automatic $\ell_1\ell_2$-norm solver. The highly folded cortical pattern along the Sylvian fissure presents a significant challenge for most inverse solvers, including both distributed and discrete parametrization approaches. One way to resolve this problem is to utilize measurements from other modalities, such as fMRI, to further constrain the solution (Liu *et al*., 1998). We defer such extensions to a future study.

Fig. 12 shows the time courses of the activated regions detected by our solver. The general shape of these time courses agrees with the neuroscience literature (Forss *et al*., 1994;Hari & Forss, 1999). Our method yields stable time courses that capture the main deflections precisely. The first deflection in cSI occurs at 20 msec. cSI soon changes its polarity and reaches its maximum at 35 msec. Although cSII has stronger signal than iSII, they have similar temporal signature: onset at 60 msec and peak at 82-85 msec. The time courses are quite similar to those estimated through dipole fitting (Fig. 12(b)), except for the cSI activation between 70 and 150 msec. This is most likely because the PPC activation was missed by the dipole fitting and its time course incorporated into the estimate of cSI. Even though the magnitude of the time courses obtained from the two solvers are not directly comparable, this comparison demonstrates the ability of the $\ell_1\ell_2$-norm regularization to achieve high-quality reconstructions of source signals. Furthermore, $\ell_1\ell_2$-norm does not restrict itself to a fixed number of dipole sources.

4.2.2 Auditory Experiments—In the auditory experiments, 500 Hz tone bursts were presented to either the right ear or the left ear of the subject according to an event-related paradigm, with a random inter-stimulus-interval between 1.2 and 1.5 seconds. Temporal sampling rate of these MEG data was 1.25 KHz. As before, a 200 msec baseline period was used for noise estimation.

After standard pre-processing, described in the last section, we applied $\ell_1\ell_2$-norm, MNE, and dipole fitting to the average data. Fig. 13 illustrates one frame of the reconstructed signals, at 110 msec after the stimulus onset. The statistics and the thresholds for $\ell_1\ell_2$-norm were computed using the same permutation method as before. Both the $\ell_1\ell_2$-norm and the dSPM detected auditory activations in both hemispheres. Due to close distance between the inferior parietal and the superior temporal regions, all three solvers have false positives in the parietal lobe. The false positives are weaker in the $\ell_1\ell_2$-norm solutions than in the MNE solutions. Given the MNE results, it is more ambiguous whether the sources originate from the auditory region or from the SII region. We also examined the polarities of the estimated sources (results not shown), and they all agree with the findings reported in the literature (Tuomisto *et al*., 1983).

The corresponding estimated source signals from the $\ell_1\ell_2$-norm solver and dipole fitting are depicted in Fig. 14. Both methods detected that the early auditory response occurs around 60 msec and that the contralateral auditory region activates slightly stronger than the ipsolateral region. Compared with the $\ell_1\ell_2$-norm solution, the dipole fitting solution captures the temporal details slightly more accurately, as reflected by the 6-8 msec difference between the contralateral and the ipsolateral activations. Nevertheless, the $\ell_1\ell_2$-norm solver is more flexible than dipole fitting in capturing the spatial extent of the activation regions.

5 Discussion

The proposed inverse solver utilizes ℓ_1-norm regularization to capture spatial sparsity of the activations and ℓ_2-norm regularization on the projection coefficients in the signal subspace to model the time-varying frequency content in the activation signals. While considering all vertices in the brain as candidate activation sources, our solver can still obtain focal activation maps and capture activation signals with precise deflection signatures. The $\ell_1\ell_2$-norm inverse solutions share some similar characteristics with dipole fitting results; however, the number of dipole sources is not required to be known *a priori* for $\ell_1\ell_2$-norm. As demonstrated in the simulations, the performance of the $\ell_1\ell_2$-norm solver is robust to the chosen number of basis functions. This feature makes the method particularly suitable for neuroscience applications where the number of dipole sources is usually unknown.

We model the activation signals as linear combinations of multiple temporal basis functions. There are various approaches to obtain the basis functions such as the Fourier and wavelet decompositions. If the Fourier decomposition is employed, the selected basis functions must capture the frequency components of the neural signals. If wavelets are used, we need to choose a wavelet family appropriately. If the temporal structure of the source signals at a particular region were known, we would incorporate it as one of the temporal basis functions. In this case, the assumption of the linear combinations of multiple basis functions would need to be modified. Furthermore, we chose to work with orthonormal basis functions. If the basis functions are not orthonormal, the general idea of this paper is still valid, but the ℓ_2-norm would have to be replaced with the Mahalanobis distance.

Obtaining a compact representation of the signals can significantly reduce the computational requirements. Because of the time-varying frequency content and substantial variability in the signals across activation regions, subjects, and tasks, data-independent basis sets such as

Fourier and wavelets may not be the best choice to compactly represent the signals. In this work, we constructed the temporal basis functions through the SVD decomposition of the data. We chose a set of basis functions that correspond to the largest singular values. The cutoff was determined by our knowledge of the source signals and the temporal structure of the singular vectors. Fixing the cutoff may lead to a loss of signals that lie in the orthogonal subspace spanned by $\mathbf{V}^{\perp}$. A possible improvement is to alternate between modifying the basis functions and performing reconstruction.

Accurate estimation of the spatial extent of the sources is one of the main challenges for any source modeling approach. Compared with dipole fitting, the $\ell_1\ell_2$-norm solver demonstrated a better ability to capture the auditory activations as shown in Section 4.2.2. To further improve the $\ell_1\ell_2$-norm solver performance for extended sources, we can apply the ℓ_1-norm regularizer to model the difference among neighboring vertices, rather than the vertices directly. Thus, the new model would favor a piece-wise constant activation pattern. This idea combined with an MNE estimator is at the core of the LORETA reconstruction method (Pascual-Marqui *et al*., 1994). The work of (Auranen *et al*., 2007) proposed a Bayesian approach in which the measurements and the hyper-prior determine the spatial extent of the activations through estimating the joint posterior distribution of the inverse solution and the exponent in the regularizer. Alternatively, the activation pattern could be expressed using a set of spatial basis functions (Limpiti *et al*., 2006) or the current multipolar expansions (Cottereau *et al*., 2007).

6 Conclusions

The proposed inverse solver takes advantage of the relatively smooth nature of the underlying EEG/MEG source signals through performing inverse operation for all temporal samples simultaneously. To overcome the overly diffuse inverse solutions, the $\ell_1\ell_2$-norm captures spatial sparsity through ℓ_1-norm regularization. It also applies an ℓ_2-norm regularizer to the projection coefficients of the temporal basis functions spanning the signal subspace. Performing reconstruction in the signal subspace while jointly considering the coefficients for all selected basis functions leads to stable estimates with a smaller number of false positives as confirmed by our experiments using simulated and real MEG data. The $\ell_1\ell_2$-norm solver is formulated as an SOCP problem. Its fast optimization enables us to perform a statistical significance test for the $\ell_1\ell_2$-norm inverse solutions via a permutation method. The $\ell_1\ell_2$-norm can be straightforwardly applied with and without orientation constraints. Its flexible formulation should also allow incorporation of fMRI information as a constraint.

Appendix

Appendix A

A quadratic inequality constraint can be revised into the canonical form of a second-order cone. We start with the standard form of a quadratic inequality constraint:

$$\|\mathbf{y} - A\mathbf{s}\|_{\ell_2}^2 \leq w \tag{16}$$

With straightforward derivations including expanding the ℓ_2-norm and completing squares, we can show that Eq. (16) is equivalent to

$$\mathbf{s}^{\mathrm{T}}A^{\mathrm{T}}A\mathbf{s}+\left(\frac{1-2\mathbf{y}^{\mathrm{T}}A\mathbf{s}+\mathbf{y}^{\mathrm{T}}\mathbf{y}-w}{2}\right)^{2}\leq\left(\frac{1+2\mathbf{y}^{\mathrm{T}}A\mathbf{s}-\mathbf{y}^{\mathrm{T}}\mathbf{y}+w}{2}\right)^{2} \tag{17}$$

Eq. (16) also implies $1 + 2\mathbf{y}^{\mathrm{T}}A\mathbf{s} - \mathbf{y}^{\mathrm{T}}\mathbf{y} + w \geq 0$. Therefore, setting $x_0=\frac{1+2\mathbf{y}^{\mathrm{T}}A\mathbf{s}-\mathbf{y}^{\mathrm{T}}\mathbf{y}+w}{2}$ and $\bar{\mathbf{x}}=\begin{bmatrix} A\mathbf{s} \\ \frac{1-2\mathbf{y}^{\mathrm{T}}A\mathbf{s}+\mathbf{y}^{\mathrm{T}}\mathbf{y}-w}{2} \end{bmatrix}$, we arrive at the equivalent second-order cone constraint $x_0 \geq \|\bar{\mathbf{x}}\|_{\ell 2}$. As we can see that the conversion introduces a set of new variables $\mathbf{x}$, of size $M + 2$.

Appendix

Appendix B

Second-order cone programming (SOCP) problems are defined by: (1) a linear objective function, (2) a feasible set that is an intersection of an affine linear manifold with the Cartesian product of second-order cones. Since the linear objective function and the feasible set are convex, SOCP problems can be solved by convex optimization techniques. The canonical primal form of SOCP is as follows:

$$\min\ \mathbf{c}^{\mathrm{T}}\mathbf{x} \tag{18}$$

$$\text{s.t. } A\mathbf{x}=\mathbf{b} \tag{19}$$

$$\mathbf{x} \in Q \tag{20}$$

where $Q = \{\mathbf{x} : \mathbf{x}_0 \geq \|\bar{\mathbf{x}}\|_{\ell 2}\}$. Q is also referred to as the Lorentz cone; it is self-dual. The dual cone Q^C is defined as

$$Q^{C}:=\left\{\mathbf{y}:\forall\mathbf{x}\in Q,\mathbf{y}^{\mathrm{T}}\mathbf{x}\geq 0\right\} \tag{21}$$

It is straightforward to prove that the self-dual property, $Q = Q^C$, using proof by contradiction.

Similar to linear programming, SOCP problems can be solved using the interior-point method with the logarithmic barrier function for the constraints. Even though the primal or dual interior-point methods developed for linear programming can be directly extended to SOCP, as described in (Nesterov & Nemirovski, 1994), the primal-dual interior-point method is preferred due to its numerical robustness.

The dual form of Eq. (18)-(20) is defined as follows:

$$\max\ \mathbf{b}^{\mathrm{T}}\mathbf{y} \tag{22}$$

$$\text{s.t. } A^{\mathrm{T}}\mathbf{y}+\mathbf{z}=\mathbf{c} \tag{23}$$

$$\mathbf{z} \in Q^{c}=Q \tag{24}$$

The general procedure of the primal-dual interior-point algorithm combines the primal and dual feasibility and the complementarity conditions and yields

$$A\mathbf{x}=\mathbf{b} \quad \mathbf{x} \in Q \tag{25}$$

$$A^{\mathrm{T}}\mathbf{y}+\mathbf{z}=\mathbf{c} \quad \mathbf{z} \in Q \tag{26}$$

$$\mathbf{x}^{\mathrm{T}}\mathbf{z}=0 \tag{27}$$

The above system of linear equations is almost identical to the corresponding one for linear programming, except for the extra conic constraints in the primal and dual feasibility equations in Eq. (25) and Eq. (26). In fact, one can combine the two conic constraints and the complementary slackness condition in Eq. (27), and reduce them to a more suitable form (Eq. (30)) for numerical solvers. The revised system of linear equations becomes

$$A\mathbf{x}=\mathbf{b} \tag{28}$$

$$A^{\mathrm{T}}\mathbf{y}+\mathbf{z}=\mathbf{c} \tag{29}$$

$$\mathbf{x} \circ \mathbf{z}=\mathbf{0} \tag{30}$$

$$\text{where } \mathbf{x} \circ \mathbf{z}:=\begin{bmatrix} \mathbf{x}^{\mathrm{T}}\mathbf{z} \\ x_0\bar{\mathbf{z}}+z_0\bar{\mathbf{x}} \end{bmatrix}=\begin{bmatrix} \mathbf{x}^{\mathrm{T}}\mathbf{z} \\ x_0 z_1+z_0 x_1 \\ \vdots \\ x_0 z_n+z_0 x_n \end{bmatrix} \tag{31}$$

We refer readers to (Alizadeh & Goldfarb, 2001) for detailed derivations. The primal-dual interior-point method solves this system of linear equations (Eq. (28)-(30)) using Newton's method. The optimization begins with a relaxed version of the complementary condition (Eq. (30)), and slowly strengthens this condition as iterations proceed. Iterations stop once the residual is less than a preselected threshold.

Acknowledgments

We thank Dmitry Malioutov and Xin Huang for interesting discussion on SOCP. We also thank Biswajit Bose, Hesheng Liu, and Thomas Yeo for helpful comments on this work, and Dr. Van Leemput for proof reading this draft. This work was supported in part by NIH NAMIC U54-EB005149, NCRR mBIRN U24-RR021382, P41-RR13218 and P41-RR14075 grants, by the NSF CAREER Award 0642971, and by U.S. DOE Award Number DE-FG02-99ER62764 to the MIND Institute. Wanmei Ou is partially supported by an NSF graduate fellowship.

References

Alizadeh, F.; Goldfarb, D. Technical Report. RRR Report number 51-2001, RUTCOR. Rutgers University; 2001. Second-order cone programming.

Allison T, McCarthy G, Wood CC, Darcey TM, Spencer DD, Williamson PD. Human cortical potentials evoked by stimulation of the median nerve. I. cytoarchitectonic areas generating short-latency activity. J Neurophysiol 1989;62:694–710. [PubMed: 2769354]

Auranen T, Nummenmaa A, Hämäläinen MS, Jääskeläinen IP, Lampinen J, Vehtari A, Sams M. Bayesian analysis of the neuromagnetic inverse problem with l^p-norm priors. NeuroImage 2007;26:870–884. [PubMed: 15955497]

Baillet S, Garnero L. A Bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem. IEEE Trans Biomed Engr 1997;44:374–385.

Barth DS, Sutherling W, Engel J, Beatty J. Neuromagnetic localization of epileptiform spike activity in the human brain. Science 1982;218:891–894. [PubMed: 6813968]

Bijma F, de Munck JC, Heethaar RM. The spatiotemporal MEG covariance matrix modeled as a sum of Kronecker products. NeuroImage 2005;27:402–415. [PubMed: 16019231]

Brooks DH, Ahmad GF, MacLeod RS, Maratos GM. Inverse electrocardiography by simultaneous imposition of multiple constraints. IEEE Trans Biomed Engr 1997;44:374–385.

Bottereau B, Jerbi K, Baillet S. Multiresolution imaging of MEG cortical sources using an explicit piecewise model. NeuroImage 2007;38:439–451. [PubMed: 17889564]

Dale AM, Sereno M. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. J Cogn Neurosci 1993;5:162–176.

Dale AM, Fischl B, Sereno M. Cortical surface-based analysis: I. segmentation and surface reconstruction. NeuroImage 1999;9:179–194. [PubMed: 9931268]

Ding L, He B. Sparse source imaging in electroencephalography with accurate field modeling. Human Brain Mapping. 2007in press

Donoho DL, Elad M. Maximal sparsity representation via ℓ_1 minimization. PNAS 2003;100:2197–2202. [PubMed: 16576749]

Efron B, Tibshirani R. Empirical bayes methods and false discovery rates for microarrays. Genet Epidemiol 2002;23:70–86. [PubMed: 12112249]

Fischl B, Sereno M, Dale AM. Cortical surface-based analysis: II. inflation, flattening, and a surface-based coordinate system. NeuroImage 1999;9:195–207. [PubMed: 9931269]

Forss N, Hari R, Salmelin R, Ahonen A, Hämäläinen M, Kajola M, Knuutila J, Simola J. Activation of the human posterior parietal cortex by median nerve stimulation. Exper Brain Res 1994;99:309–15. [PubMed: 7925811]

Galka A, Yamashita O, Ozaki T, Biscay R, Valdes-Sosa P. A solution to the dynamical inverse problem of EEG generation using spatiotemporal Kalman filtering. NeuroImage 2004;23:435–453. [PubMed: 15488394]

Genovese CR, Lazar NA, Nichols T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. NeuroImage 2002;15:870–878. [PubMed: 11906227]

Geva AB. Bioelectric sources estimation using spatio-temporal matching pursuit. Applied Sig Proc 1998;5:195–208.

Gorodnitsky IF, Rao BD. Sparse signal reconstruction from limited data using FOCUSS: a re-weighted minimum norm algorithm. IEEE Trans Sig Proc 1997;45:600–616.

Hadamard J. Sur les problems aux derivees partielles et leur signification physique. Princeton University Bulletin 1902:49–52.

Hämäläinen MS, Ilmoniemi R. Interpreting measured magnetic fields of the brain: Estimates of current distributions. Techical Report TKK-F-A559. 1984

Hämäläinen MS, Sarvas J. Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. IEEE Trans Biomed Engr 1989;36:165–171.

Hämäläinen MS, Hari R, Ilmoniemi R, Knuutila J, Lounasmaa OV. Magnetoencephalography - theory, instrumentation, and applications to noninvasive studies of the working human brain. Reviews of Modern Physics 1993;65:413–497.

Hämäläinen, MS. MNE software user's guide. NMR Center, Mass General Hospital, Harvard University; 2005.

Hari R, Forss N. Magnetoencephalography in the study of human somatosensory cortical processing. Philos Trans R Soc Lond B 1999;354:1145–1154. [PubMed: 10466142]

Huang M, Aine CJ, Supek S, Best E, Ranken D, Flynn ER. Multistart downhill simplex method for spatio-temporal source localization in magnetoencephalography. Electroencephalogr Clin Neurophysiol 1998;108:32–44. [PubMed: 9474060]

Huang MX, Dale AM, Song T, Halgren E, Harrington DL, Podgomy I, Canive JM, Lewis S, Lee RR. Vector-based spatial-temporal minimum L1-norm solution for MEG. NeuroImage 2006;31:1025–1037. [PubMed: 16542857]

Huizenga HM, de Munck JC, Waldorp LJ, Grasman RPPP. Spatiotemporal EEG/MEG source analysis based on a parametric noise covariance model. IEEE Tran Biomed Engr 2002;49:533–539.

Kaipio J, Somersalo E. Statistical and Computational Inverse Problems. Springer. 2004

Jerbi K, Baillet S, Mosher JC, Nolte G, Gamero L, Leahy RM. Localization of realistic cortical activity in MEG using current multipoles. NeuroImage 2004;22:779–793. [PubMed: 15193607]

Kass J, Nelson R, Sur M, Lin C-S, Merzenich M. Multiple representations of the body within the primary somatosensory cortex of primates. Science 1979;204:521–523. [PubMed: 107591]

Knake S, et al. The value of multichannel MEG and EEG in the presurgical evaluation of 70 epilepsy patients. Epilepsy Res 2006;69:80–86. [PubMed: 16516443]

Lin F-H, Belliveau JW, Dale AM, Hämäläinen MS. Distributed current estimates using cortical orientation constraints. Human Brain Mapping 2006;27:1–13. [PubMed: 16082624]

Lamus C, Long CJ, Hämäläinen MS, Brown EN, Purdon PL. Parameter estimation and dynamic source localization for the magnetoencephalography (MEG) inverse problem. ISBI. 2007

Limpiti T, Van Veen BD, Wakai RT. Cortical patch basis model for spatially extended neural activity. IEEE Trans Biomed Engr 2006;53:1740–1754.

Liu AK, Belliveau JW, Dale AM. Spatiotemporal imaging of human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations. PNAS 1998;95:8945–50. [PubMed: 9671784]

Malioutov M, Çetin M, Willsky AS. A sparse signal reconstruction perspective for source localization with sensor arrays. IEEE Trans Sig Proc 2005;53:3010–3022.

Matsuura, K.; Okabe, Y. Advances in Biomagnetism Res: Biomag96. Springer; New York: 1999. Multiple current-dipole distribution reconstruction reconstructed by modified selective minimum-norm method; p. 290-293.

Mosher JC, Lewis PS, Leahy RM. Multiple dipole modeling and localization from spatio-temporal MEG data. IEEE Trans Biomed Engr 1992;39:541–557.

Nelder JA, Mead R. A Simplex Method for Function Minimization. Computer J 1965;7:308–313.

Nesterov, Y.; Nemirovski, A. Studies in Applied Mathematics. SIAM; Philadelphia, PA: 1994. Interior-point polynomial methods in convex programming.

Oostendorp TF, Van Oosterom A. Source parameter estimation in inhomogeneous volume conductors of arbitrary shape. IEEE Trans Biomed Engr 1989;36:382–391.

Pantazis D, Nichols TE, Baillet S, Leahy RM. A comparison of random field theory and permutation methods for the statistical analysis of MEG data. NeuroImage 2005;25:383–394. [PubMed: 15784416]

Pascual-Marqui RD, Michel CM, Lehmann D. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. International J of Psychophysiology 1994;18:49–65.

Rao BD, Kreutz-Delgado K. An affine scaling methodology for best basis selection. IEEE Trans Sig Proc 1999:47, 187–200.

Rowe M, Turman A, Murray G, Zhang H. Parallel organization of somatosensory cortical areas I and II for tactile processing. Clin Exp Pharmacol Physiol 1996;23:931–938. [PubMed: 8911737]

Scherg M, Von Cramon D. Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model. Electroencephalogr Clin Neurophysiol 1985;62:32–44. [PubMed: 2578376]

Schmitt U, Louis AK, Darvas F, Buchner H, Fuchs M. Numerical aspects of spatio-temporal current density reconstruction from EEG-/MEG-data. IEEE Trans Med Imag 2001;20:314–324.

Sturm, JF.; Romanko, O.; Polik, I.; Terlaky, T. Self-Dual-Minimization (SeDuMi): solver for optimization problems over symmetric cones. 2001. http://sedumi.mcmaster.ca/

Trujillo-Barreto N, Aubert-Vazquez E, Penny W. Bayesian M/EEG source reconstruction with spatio-temporal priors. NeuroImage 2007;39:318–335. [PubMed: 17904869]

Tuomisto T, Hari R, Katila T, Poutanen T, Varpula T. Studies of auditory evoked magnetic and electric responses: Modality specificity and modelling. Nuovo Cimento 1983;D 2:471–483.

Uutela K, Hämäläinen MS, Salmelin R. Global optimization in the localization of neuromagnetic sources. IEEE Trans Biomed Engr 1998;45:716–723.

Uutela K, Hämäläinen MS, Somersalo E. Visualization of magnetoencephalographic data using minimum current estimates. NeuroImage 1999;10:173–180. [PubMed: 10417249]

Wang JZ, Williamson SJ, Kaufman L. Magnetic source imaging based on the minimum-norm least-squares inverse. Brain Topography 1993;5:365–371. [PubMed: 8357709]

Weerd JPC, Kap JI. Spectro-temporal representations and time-varying spectra of evoked potentials. Biological Cybernetics 1981;41:101–107. [PubMed: 7248340]

Wikström H, Huttunen J, Korvenoja A, Virtanen J, Salonen O, Aronen H, Ilmoniemi RJ. Effects of interstimulus interval on somatosensory evoked magnetic fields (SEFs): a hypothesis concerning SEF generation at the primary sensorimotor cortex. Electroencephalogr Clin Neurophysiol 1996;100:479–487. [PubMed: 8980411]

Wood CC. Application of dipole localization methods to source identification of human evoked potentials. Ann N Y Acad Sci 1982;388:139–155. [PubMed: 6953865]

Zhang Y, Ghodrati A, Brooks DH. An analytical comparison of three spatio-temporal regularization methods for dynamic linear inverse. problems in a common statistical framework. Inverse Problems 2005;21:357–382.

Source modeling software (xfit). Elekta-Neuromag Oy; Helsinki, Finland:

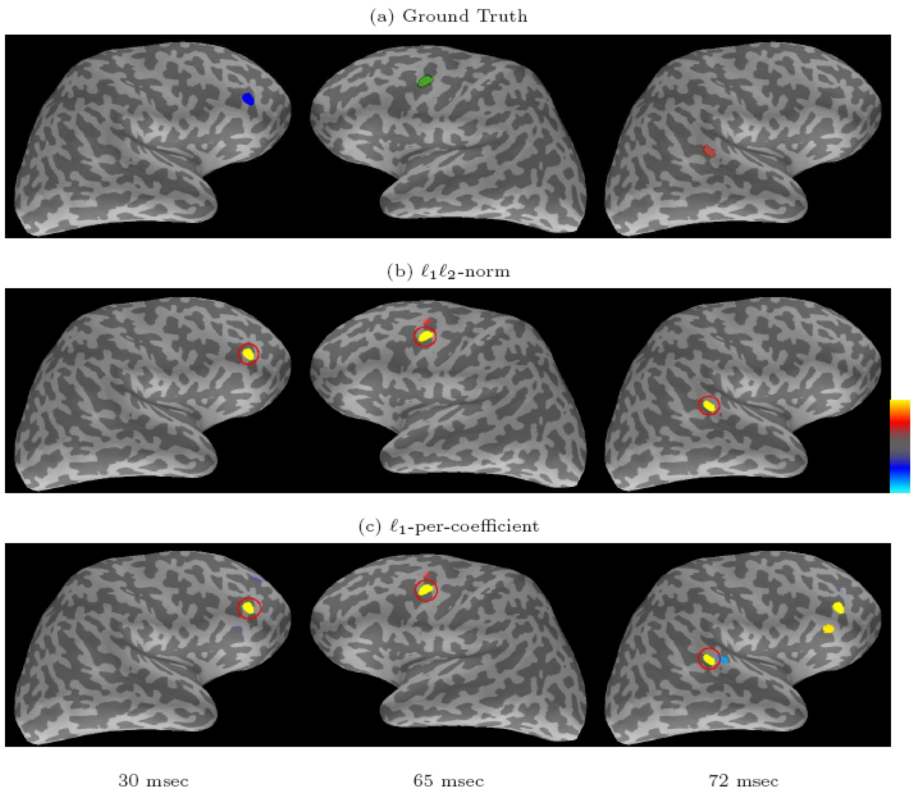


Fig. 1.

Activation maps at different time frames. (a) Ground truth activation maps at peak response time for three sources. (b) The spatial maps estimated using the $\ell_1\ell_2$-norm solver. (c) The spatial maps estimated using the ℓ_1-per-coefficient solver. The color codes in (a) do not indicate current directions. Hot/cold colors in (c,d) correspond to outward/inward current flow. The most active areas in the solutions are highlighted, and their estimated time courses are shown in Fig. 2a-c.

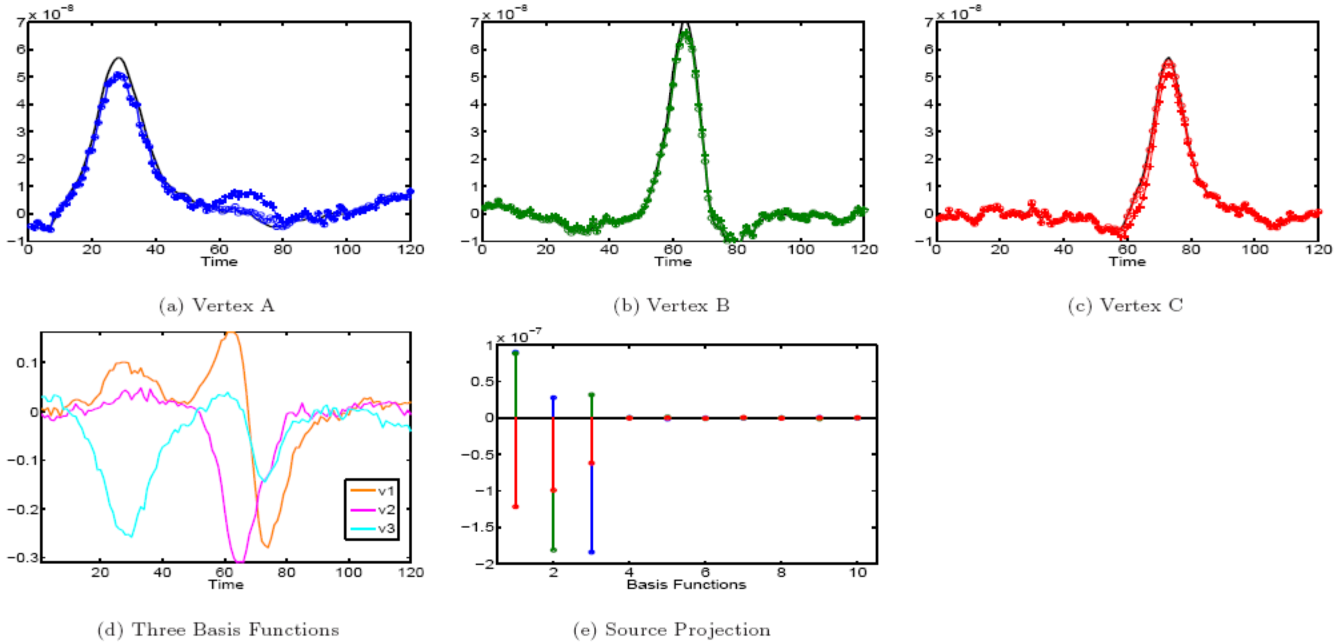


Fig. 2.

Reconstructed source signals and the three basis functions. The top row illustrates the simulated (black solid curves) and the reconstructed source time courses ($\ell_1\ell_2$-norm marked as 'o' and ℓ_1-per-coefficient marked as '+') for the three active vertices; the bottom row presents the selected basis functions (d) and the projection coefficients of the simulated source time courses onto the top ten basis functions (e).

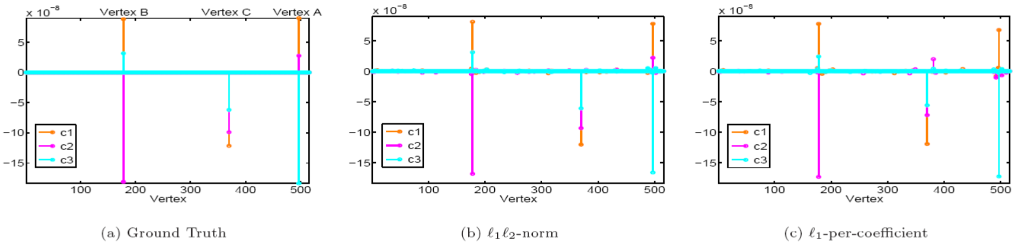


Fig. 3.

Reconstructed coefficients $\tilde{\mathbf{S}}$. (a) Projection coefficients of the simulated data onto the three basis functions, corresponding to the top three singular components. (b) Reconstructed coefficients from $\ell_1\ell_2$-norm. (c) Reconstructed coefficients from ℓ_1-per-coefficient.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

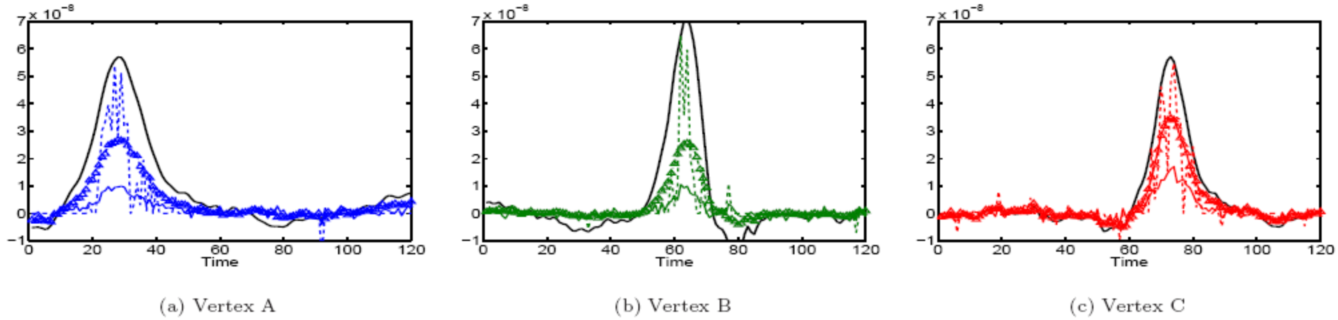


Fig. 4.
Reconstructed source signals from MNE, MCE, and VESTAL. Each panel illustrates the simulated (black solid curves) and the reconstructed source time courses: MNE (blue solid), MCE (dashed), and VESTAL (marked as ‘Δ’).

NIH-PA Author Manuscript
NIH-PA Author Manuscript

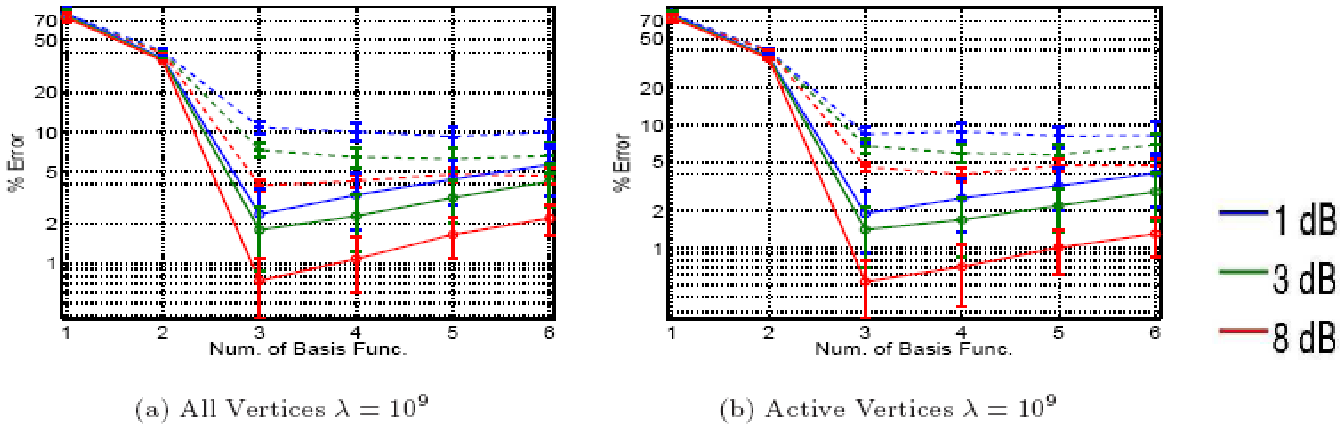
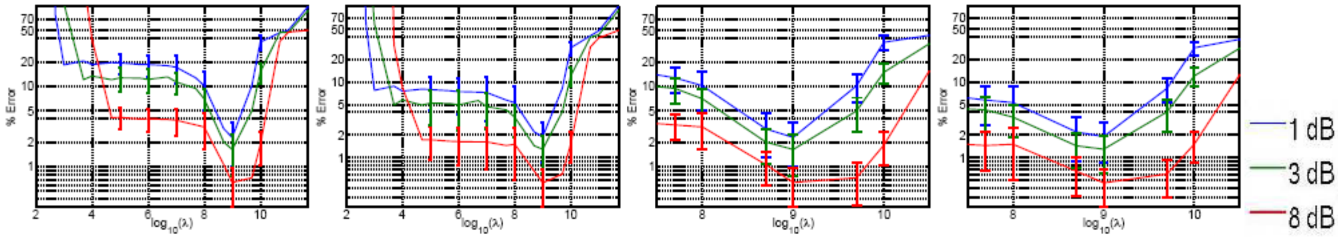


Fig. 5.

Relative MSE. This figure presents the percentage relative MSE of the reconstruction results using the $\ell_1\ell_2$-norm (solid curves) and the ℓ_1-per-coefficient (dashed curves) for all vertices and the three active vertices under three different SNR settings. Note that the error bars close to the bottom of the figures appear large due to the logarithmic scale.



(a) All Vertices (b) Active Vertices (c) All Vertices (Zoomed-in) (d) Active Vertices (Zoomed-in)

Fig. 6.

Relative MSE vs. regularization strength. This figure presents the relative MSE of the reconstruction results from $\ell_1\ell_2$-norm for all vertices and for the three active vertices under three SNR settings as the regularization strength, λ, varies between 10^2 and 10^{12}. Note that the error bars close to the bottom of the figures appear large due to the logarithmic scale. (c,d) are the corresponding zoomed-in versions.

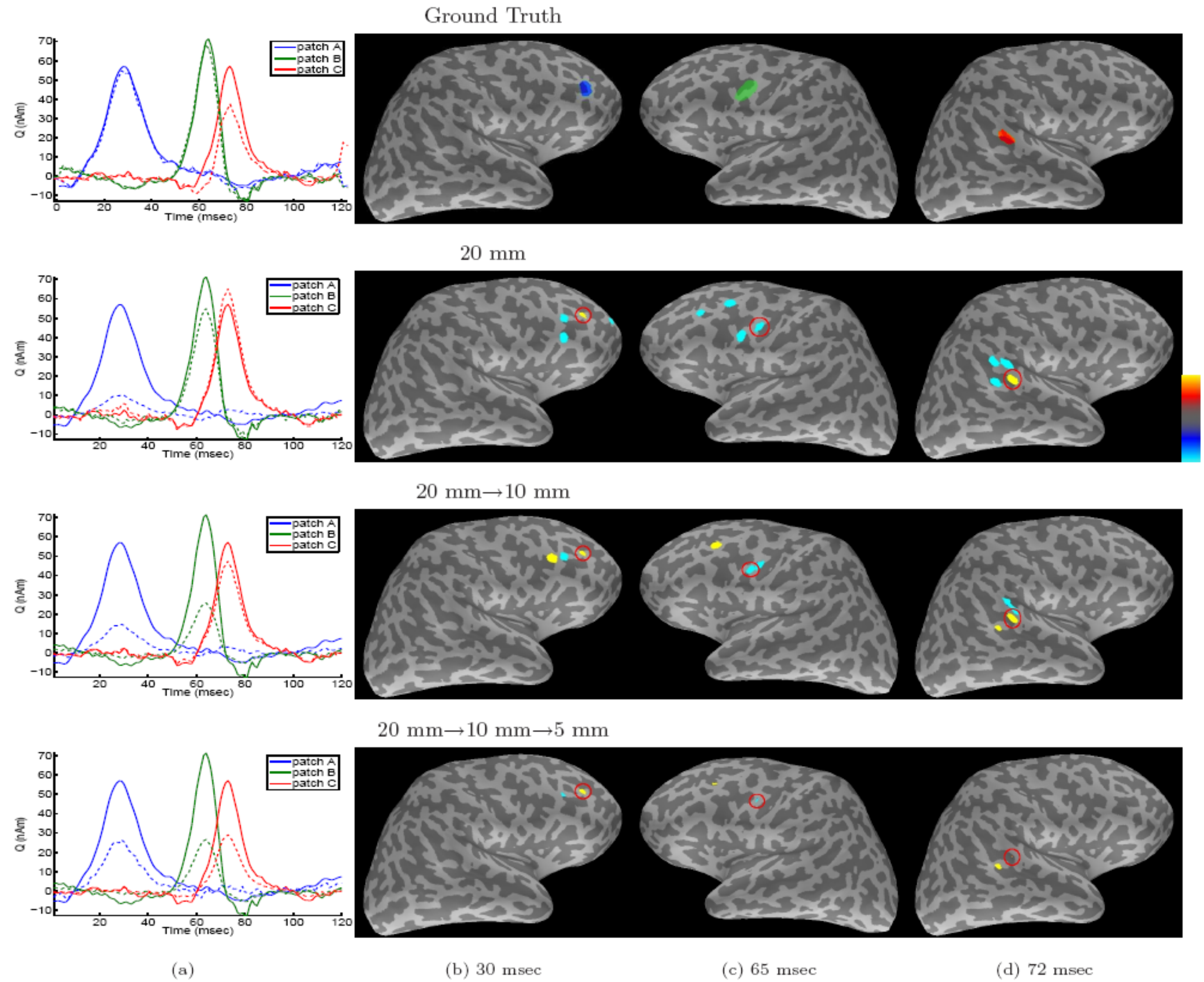


Fig. 7.

Reconstructions obtained using the $\ell_1\ell_2$-norm solver with different multi-resolution schemes. The top row presents the three simulated activation patches (color is used to label patches and does not indicate current directions). The detected results (row two to four) are shown in hot or cold colors corresponding to current flowing outward and inward, respectively. The time courses of the highlighted areas are shown in column (a). Solid curves in (a) are the simulated time courses. The dashed curves in row one are the dipole fitting results. The dashed curves in row two to four are the reconstructed time courses from $\ell_1\ell_2$-norm, which are scaled by a factor of four for illustration purposes.

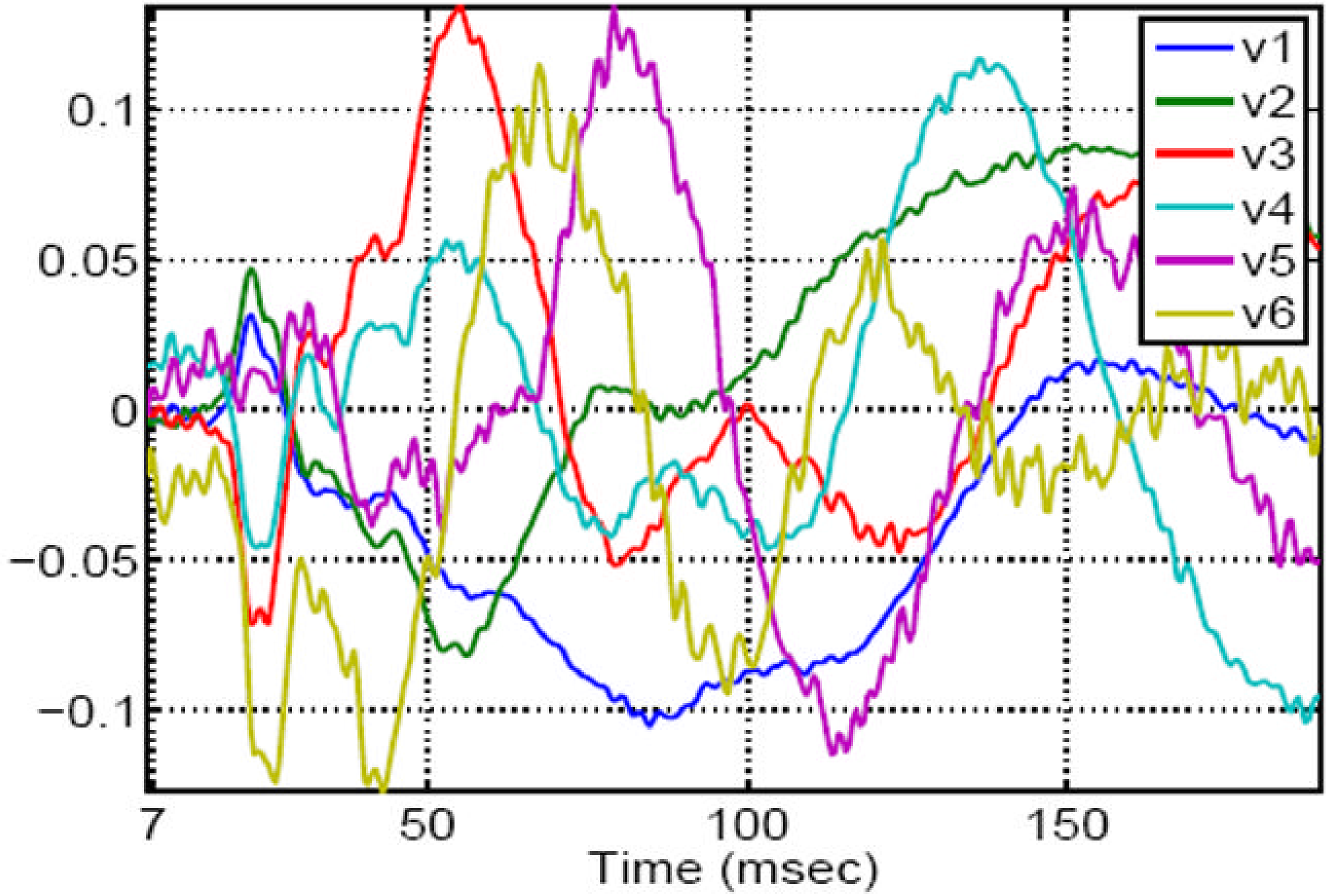


Fig. 8.
Six selected basis functions. The basis functions were obtained from SVD of the MEG measurements between 6 msec and 200 msec after stimulus onset.

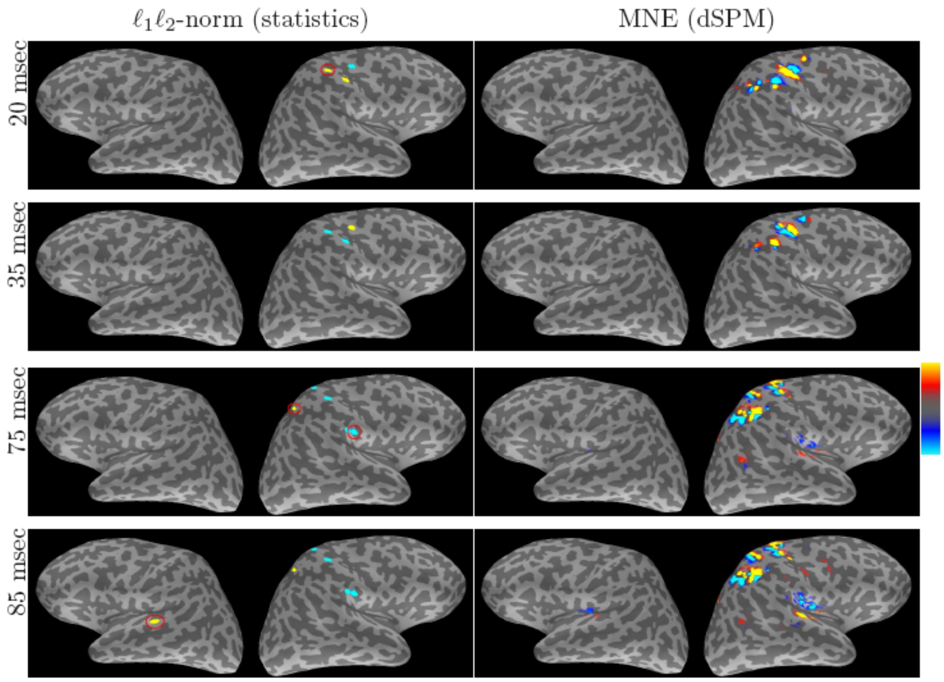


Fig. 9.
Significance statistics of the $\ell_1\ell_2$-norm solver and MNE for the median-nerve experiment. Hot/cold color corresponds to outward/inward current flow. The most active areas in the $\ell_1\ell_2$-norm solutions are highlined, and their estimated time courses are shown in Fig. 12.

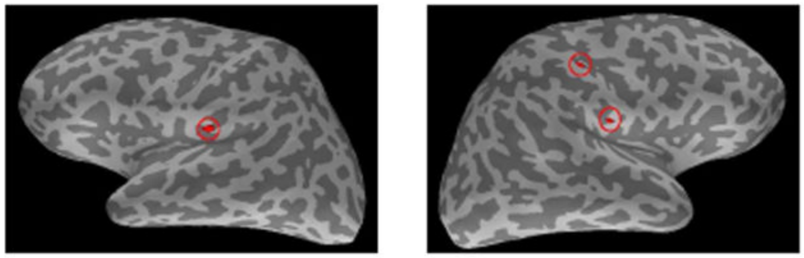


Fig. 10.
Dipole fitting results with three sources.

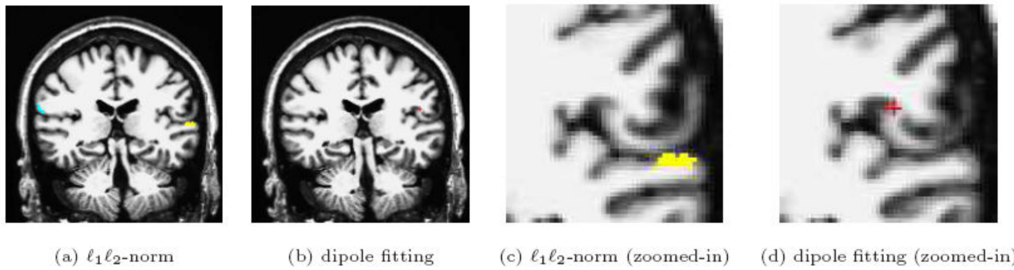


Fig. 11.

Coronal slices for the detected iSII activations from $\ell_1\ell_2$-norm and from dipole fitting. Subfigures (c,d) are the corresponding zoomed-in versions.

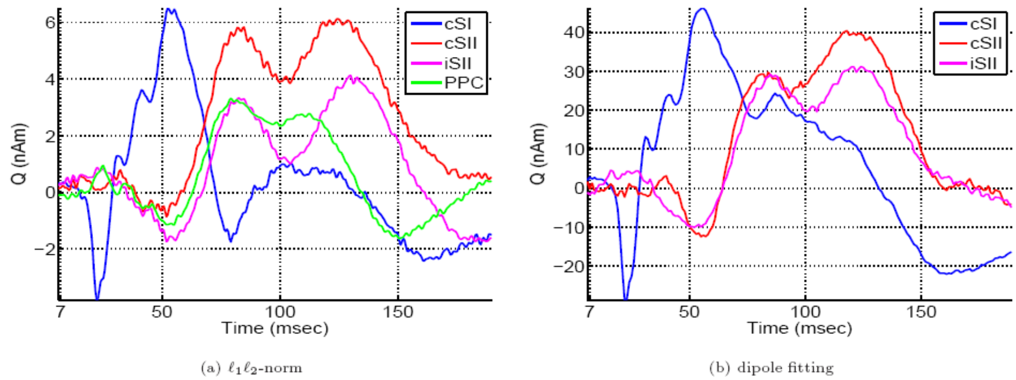


Fig. 12.

Reconstructed time courses obtained from the $\ell_1\ell_2$-norm solver and dipole fitting. The corresponding activation maps are reported in Fig. 9 and Fig. 10.

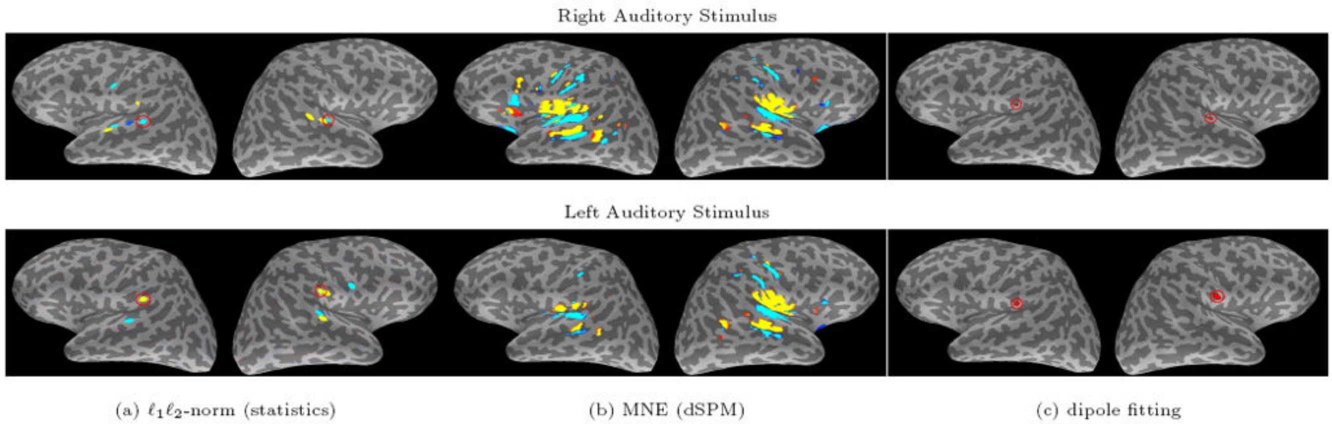


Fig. 13.

Significance statistics of the $\ell_1\ell_2$-norm solver and MNE, and dipole fitting at 110 msec after right (top) and left (bottom) auditory stimulus onset. Hot/cold color corresponds to outward/ inward current flow. The most active areas are highlighted, and their estimated time courses are shown in Fig. 14.

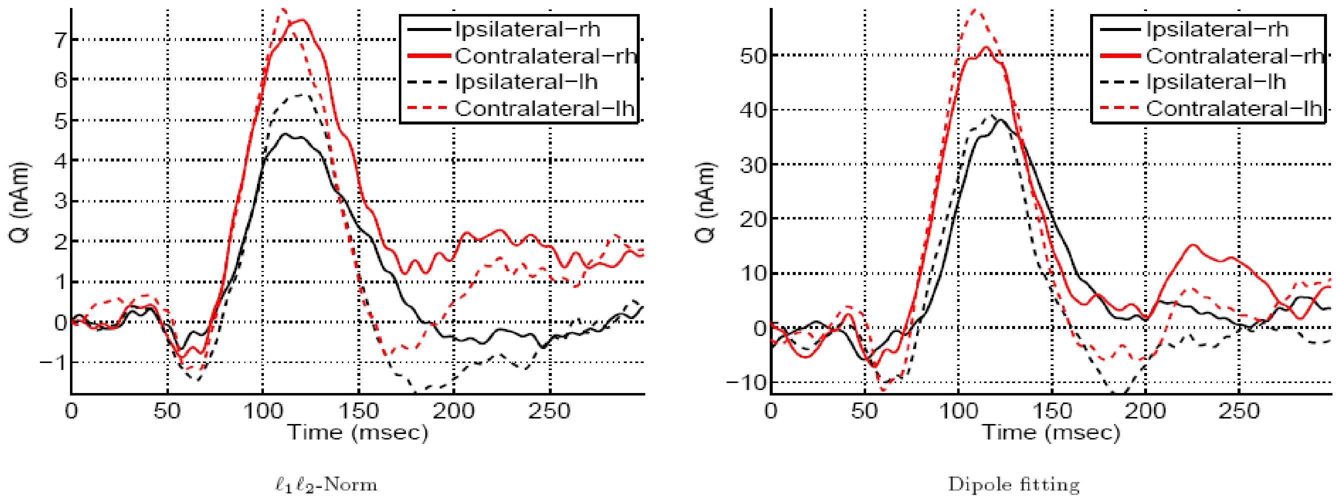


Fig. 14.
Reconstructed time courses from $\ell_1\ell_2$-norm and dipole fitting.

NIH-PA Author Manuscript
NIH-PA Author Manuscript