



Reconstructing spatially extended brain sources via enforcing multiple transform sparseness

Min Zhu^a, Wenbo Zhang^{b,c}, Deanna L. Dickens^{b,c}, Lei Ding^{a,d,e,*}

^a School of Electrical and Computer Engineering, University of Oklahoma, Norman, OK, USA

^b Minnesota Epilepsy Group, P.A. at John Nasseff Neuroscience Institute, Allina Health, St Paul, MN, USA

^c Department of Neurology, University of Minnesota, Minneapolis, MN, USA

^d Center for Biomedical Engineering, University of Oklahoma, Norman, OK, USA

^e Laureate Institute for Brain Research, Tulsa, OK, USA

ARTICLE INFO

Article history:

Accepted 28 September 2013

Available online 5 October 2013

Keywords:

Sparse source imaging

L1-norm regularization

MEG

Extent

Transform sparseness

Multiple penalties

ABSTRACT

Accurate estimation of location and extent of neuronal sources from EEG/MEG remain challenging. In the present study, a new source imaging method, i.e. variation and wavelet based sparse source imaging (VW-SSI), is proposed to better estimate cortical source locations and extents. VW-SSI utilizes the L1-norm regularization method with the enforcement of transform sparseness in both variation and wavelet domains. The performance of the proposed method is assessed by both simulated and experimental MEG data, obtained from a language task and a motor task. Compared to L2-norm regularizations, VW-SSI demonstrates significantly improved capability in reconstructing multiple extended cortical sources with less spatial blurredness and less localization error. With the use of transform sparseness, VW-SSI overcomes the over-focused problem in classic SSI methods. With the use of two transformations, VW-SSI further indicates significantly better performance in estimating MEG source locations and extents than other SSI methods with single transformations. The present experimental results indicate that VW-SSI can successfully estimate neural sources (and their spatial coverage) located in close areas while responsible for different functions, i.e. temporal cortical sources for auditory and language processing, and sources on the pre-bank and post-bank of the central sulcus. Meantime, all other methods investigated in the present study fail to recover these phenomena. Precise estimation of cortical source locations and extents from EEG/MEG is of significance for applications in neuroscience and neurology.

© 2013 Elsevier Inc. All rights reserved.

Introduction

Electroencephalography (EEG) and magnetoencephalography (MEG) are promising noninvasive neuroimaging technologies to probe human brain activities with excellent temporal resolutions (milliseconds), as compared with other neuroimaging technologies, e.g. functional magnetic resonance imaging (fMRI). Meantime, the precise reconstruction of brain sources behind EEG/MEG still remains challenging, which are usually obtained by solving so-called EEG/MEG inverse problems (Baillet et al., 2001). Accurate estimation of EEG/MEG source locations and extents, however, is of significance in understanding human brain functions (Dhond et al., 2001; Hillyard, 1993) and addressing clinical needs (Brodbeck et al., 2011). EEG/MEG inverse solutions based on equivalent current dipole (ECD) source models provide localizations for single or a few focal brain activations (Stefan et al., 2003; Wood, 1982), when the number of dipoles is known or can be estimated (Wood, 1982). However, ECD solutions provide no estimation of source

extents since ECDs are only discrete source points, which can even potentially cause bias in locating spatially extended brain sources (Ou et al., 2009; Plummer et al., 2008).

Recognizing the limitation of ECD, distributed current density (DCD) models have been developed to model extended brain sources, in which the source space is defined as a set of distributed dipoles over a three dimensional (3D) brain volume (Ding and He, 2008; Pascual-Marqui et al., 1994) or a two dimensional (2D) cortical surface (cortical current density, i.e. CCD, models) (Dale and Sereno, 1993). It is thus theoretically possible to infer both source locations and extents in inverse solutions. Practically, the estimation accuracy still largely depends on how efficiently DCD-based inverse problems can be solved, which have infinite number of solutions to a given set of measurements (Baillet et al., 2001). Unique solutions are typically obtained by introducing anatomical and/or functional priors, via a procedure known as regularization (Vega-Hernández et al., 2008). The most common regularization approach is to have minimum overall energy (i.e. L2-norm) in inverse solutions, such as minimum norm estimate (MNE) (Hämäläinen and Ilmoniemi, 1994) and its variants, i.e. weighted MNE (wMNE) (Dale and Sereno, 1993) and low-resolution electromagnetic tomography (LORETA) (Pascual-Marqui et al., 1994). L2-norm regularizations

* Corresponding author at: University of Oklahoma, 110 W. Boyd St. DEH Room 150, Norman, OK 73019, USA. Fax: +1 405 325 7066.

E-mail address: leiding@ou.edu (L. Ding).

belong to quadratic regression problems (Hoerl and Kennard, 1970), which assume Gaussian source fields (Uutela et al., 1999) and produce over-smooth estimations (Pascual-Marqui et al., 1994; Vega-Hernández et al., 2008). The nature of smoothness makes them not suitable for estimating source extents from early sensory brain responses and focal epilepsy, in which sources are proved to be compact (Allison et al., 1989; Oishi et al., 2002).

Other regularizations utilizing non-quadratic regression schemes, such as minimum current (i.e. L1-norm), have demonstrated sparse solutions to address the smoothness problem (Ding and He, 2008; Tibshirani, 1996; Uutela et al., 1999). However, most of them, including selective minimum-norm method (Matsuura and Okabe, 1995), minimum current estimate (MCE) (Uutela et al., 1999), least absolute shrinkage selection operator (LASSO) (Tibshirani, 1996), and sparse source imaging (SSI) (Ding and He, 2008), produce over-focused solutions, which do not reflect accurate source extents either (Chang et al., 2010; Ding et al., 2011). The over-focality is caused by insufficient consideration of source extents in models due to the enforcement of sparseness in the original source domain, which leads to the idea of enforcing sparseness in transform domains (Chang et al., 2010; Ding, 2009), where signals (i.e. current densities) can be sparser or more compressible (Candès and Romberg, 2007). Variation transform, which computes the difference between neighbored elements, has been first reported to use transform sparseness (Ding, 2009). Inverse solutions of variation based L1-norm regularizations allow the reconstruction of extended sources, in which source extents can be inferred. Laplace transform has also been proposed to compress current densities using the second-order spatial derivative (Chang et al., 2010), which promotes smoothness in neighborhoods. However, since the minimization of both variation and second-order derivative does not limit the global energy of inverse solutions, L1-norm regularizations with these transforms must incorporate additional priors to constrain global energies. Wavelet transforms, efficient methods in compressing signals and/or images, have been developed for complex 2D surfaces with either regular subdivisions, e.g. spherical wavelets (Schröder and Sweldens, 1995), or irregular subdivisions, e.g. face-based wavelets (Liao et al., 2012; Valette and Prost, 2004), which can be applied to compress current densities on highly convoluted cortical surfaces, i.e. CCD models. The L1-norm regularizations based on wavelet transforms have thus been pursued (Chang et al., 2010; Liao et al., 2012). Liao et al. (2012) further indicates that the L1-norm regularizations using face-based wavelets have better accuracy in recovering sources than spherical wavelets (Chang et al., 2010). Furthermore, adaptive estimation of source extents of both focal and extended sources has been suggested possible by controlling the level of wavelet compression (Chang et al., 2010), which is, however, challenging without a priori knowledge about the size of sources.

In the present study, a new sparse source imaging method, i.e. variation and wavelet based SSI (VW-SSI), is proposed using multiple penalties in L1-norm regularization, i.e. enforcing sparseness in both variation and wavelet domains (with the face-based wavelet). It aims to address the limitation in the variation-based method since the wavelet penalty term constrains the global energy, and to stabilize the issue of selecting the wavelet compression level with the variation penalty term. With the hybrid sparseness constraints integrated, it is expected that SSI techniques can achieve better accuracy in estimating source locations and extents even in data with low signal-to-noise ratio (SNR). The performance of the proposed method was evaluated in simulated and experimental MEG data. In simulations, neural activations of various extents were randomly located. The VW-SSI inverse solutions were assessed using multiple metrics on the accuracy of both locations and extents, as compared with other SSIs and L2-norm methods. The sensitivities of VW-SSI to SNR, wavelet compression level, and hyper-parameter for multiple penalty terms were studied. Experimental MEG data collected from both language and motor tasks in an epilepsy patient were analyzed to evaluate and compare all methods in reconstructing distributed neural activations.

Material and methods

Sparse source imaging using L1-norm regularization

Giving the CCD source model and the conductive profile of the head, the relationship between MEG measurements $\vec{\phi}$ and source \vec{s} can be expressed as (Nunez, 1981):

$$\vec{\phi} = \mathbf{A} \vec{s} + \vec{n} \quad (1)$$

where \vec{n} denotes the noise and \mathbf{A} is the lead field. Since the number of measurements M is much smaller than the number of dipoles N , its inverse problem has no unique solution. L1-norm regularizations have been proposed to search for a unique solution by enforcing sparseness in either original source domain or transform domains (Chang et al., 2010; Ding, 2009; Ding and He, 2008; Liao et al., 2012; Tibshirani, 1996; Uutela et al., 1999), which can be universally expressed as a constrained optimization problem:

$$\min \|\mathbf{H} \vec{s}\|_1 \quad \text{subject to} \quad \|\vec{\phi} - \mathbf{A} \vec{s}\|_2 < \varepsilon \quad (2)$$

where the matrix \mathbf{H} is either an identity matrix or any matrix for a transform. The regularization parameter ε is estimated by the discrepancy principle (Morozov, 1966). Assuming Gaussian white noise with variance σ^2 , $(1/\sigma^2) \|\vec{n}\|_2^2$, where $\vec{n} = \vec{\phi} - \mathbf{A} \vec{s}$, can be treated as a χ^2 -distribution with degree of freedom as the number of sensors. To make the probability of $\|\vec{\phi} - \mathbf{A} \vec{s}\|_2 \geq \varepsilon$ small enough, ε is selected as the upper bound of the confidence interval $[0, \varepsilon]$ that integrates to 0.99 probabilities.

Using different transform matrix \mathbf{H} , various L1-norm regularizations can be formed: (1) SSI (Ding and He, 2008) where \mathbf{H} is the identity matrix; (2) variation based SSI (V-SSI) (Ding, 2009) using the variation operator \mathbf{V} as \mathbf{H} (Variation transform section); (3) wavelet based SSI (W-SSI) (Liao et al., 2012) using the wavelet transform matrix \mathbf{W}_m as \mathbf{H} (Face-based wavelet transform section), where the index m stands for the level of wavelet compression ($m = 1, 2, 3, 4$).

Variation transform

In the CCD model meshed with triangles (see Simulation protocol section for details), the variation transform is defined as (Ding, 2009):

$$\mathbf{V} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1N} \\ v_{21} & v_{22} & \cdots & v_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ v_{P1} & v_{P2} & \cdots & v_{PN} \end{bmatrix} \quad \begin{cases} v_{ij} = 1; v_{ik} = -1; & \text{if elements } j, k \text{ share the same edge } i \\ v_{ij} = 0; & \text{otherwise} \end{cases} \quad (3)$$

where P is the number of triangular edges. Each element in the variation vector, $\mathbf{V} \vec{s}$, is defined on each triangular edge indicating the change of values over neighbored triangles.

Face-based wavelet transform

To define wavelet transform on the cortical surface, a multi-resolution cortical model (Fig. 1) is firstly constructed by iteratively compressing the highly convoluted cortical structure to create a series of spaces for multi-resolution wavelet analysis (Valette and Prost, 2004). The compression procedure is accomplished by hierarchically merging multiple triangles on a finer level into one triangle on a coarser level. With the multi-resolution cortical model, the scaling function supported on a triangle at level m is designed as one on the triangle and zeros otherwise. Then, its decompositions to approximations (scaling coefficients) and details (wavelet coefficients) at the next coarse level $m+1$ can be obtained by directly applying analysis matrices \mathbf{A}^m and \mathbf{B}^m (Liao et al., 2012). See details on how to construct \mathbf{A}^m and \mathbf{B}^m in

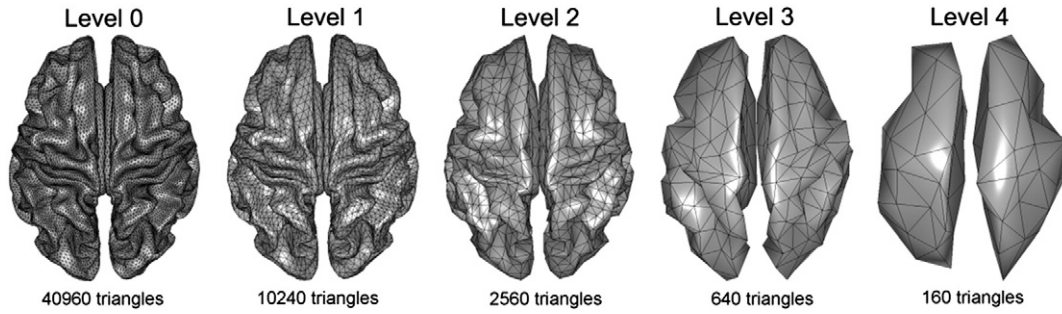


Fig. 1. An example of multi-resolution model with four levels of compression.

Appendix A. Wavelet transform matrix \mathbf{W}_m at level m is composed using analysis matrices (Liao et al., 2012):

$$\mathbf{W}_m = \begin{bmatrix} (\mathbf{G}^m)^{1/2} \mathbf{A}^{m-1} \mathbf{A}^{m-2} \dots \mathbf{A}^0 \\ \mathbf{B}^{m-1} \mathbf{A}^{m-2} \dots \mathbf{A}^0 \\ \vdots \\ \mathbf{B}^1 \mathbf{A}^0 \\ \mathbf{B}^0 \end{bmatrix} \quad (4)$$

where \mathbf{G}^m is the Gram–Schmidt matrix, i.e. a diagonal matrix with its diagonal elements being areas of triangles on the cortical mesh at level m .

Variation and wavelet based sparse source imaging (VW-SSI)

The proposed VW-SSI method integrates both variation and wavelet transforms in the L1-norm regularization and results in the optimization function as:

$$\min \|\mathbf{V} \bar{\mathbf{s}}\|_1 + \lambda \|\mathbf{W}_m \bar{\mathbf{s}}\|_1 \quad \text{subject to} \quad \|\bar{\boldsymbol{\phi}} - \mathbf{A} \bar{\mathbf{s}}\|_2 < \varepsilon \quad (5)$$

where $\lambda > 0$ is the hyper-parameter to balance variation and wavelet penalties.

By introducing intermediate variables, Eq. (5) is readily to be casted as a second-order cone program (SOCP) (Lobo et al., 1998):

$$\min_{\bar{\mathbf{s}}} \sum_{i=1}^P f_i + \lambda \sum_{k=1}^N g_k \quad \text{subject to} \quad \bar{\mathbf{n}} = \bar{\boldsymbol{\phi}} - \mathbf{A} \bar{\mathbf{s}}; \|\bar{\mathbf{n}}\|_2 \leq \varepsilon; \\ \left\| \sum_{j=1}^N v_{ij} s_j \right\|_2 \leq f_i \quad i = 1, 2, \dots, P; \quad \left\| \sum_{j=1}^N v_{kj} s_j \right\|_2 \leq g_k \quad k = 1, 2, \dots, N \quad (6)$$

where v_{ij} and w_{kl} are elements in \mathbf{V} and \mathbf{W} , respectively. Sums of elements f_i, g_k (non-negative) in vectors \mathbf{f}, \mathbf{g} serve as upper bounds for L1-norms of variation and wavelet coefficients. The primal-dual interior-point method (IPM) (Nesterov and Nemirovski, 1994) implemented in a Matlab package called Self-Dual-Minimization (SeDuMi) (Sturm, 1999) is used to solve the SOCP problem in Eq. (6), as well as optimization problems in other SSI methods as universally expressed in Eq. (2).

Simulation protocol

An averaged structural MRI from FreeSurfer (<http://surfer.nmr.mgh.harvard.edu>) was used to build the CCD model and head volume conductor model for simulations. The cortical surface was segmented at the interface of gray and white matters and tessellated into 40,960 triangles with each of $3.23 \pm 1.14 \text{ mm}^2$ (mean \pm SD) using FreeSurfer (Dale et al., 1999). Each triangular element was represented by a current dipole located at its center and the dipole orientation was perpendicular to the cortical surface. A boundary element (BE) model (He et al., 1999) was used as the realistic head volume conductor, consisting of the

scalp, skull and brain with conductivities of $0.33(\Omega \cdot \text{m})^{-1}$, $0.0165(\Omega \cdot \text{m})^{-1}$, and $0.33/\Omega \cdot \text{m}$ (the brain-to-skull conductivity ratio as 20) (Lai et al., 2005), respectively.

To perform wavelet based algorithms (i.e. W-SSI and VW-SSI), the original cortical mesh with 40,960 triangles was iteratively compressed into four coarse meshes with 10,240, 2560, 640, and 160 triangles, respectively, to form a multi-resolution model (Fig. 1). Cortical current densities ($\bar{\mathbf{s}}$) as functions defined on the finest level can thus be analyzed by applying the face-based wavelet transform.

To mimic brain activations from synchronized neuronal populations (Hämäläinen et al., 1993), cortical patch sources were formed using region growing technique by randomly selecting a seed triangle and iteratively adding its neighbors. The extent degree x was defined as an indicator of the size of patch source that includes x iterations of adding neighbored triangles. Patch sources were simulated with varied source extents (extent degrees 3, 5, 7, 9, 11 and their corresponding extents are $2.34 \pm 0.38 \text{ cm}^2$, $7.68 \pm 1.06 \text{ cm}^2$, $16.11 \pm 2.10 \text{ cm}^2$, $27.47 \pm 3.49 \text{ cm}^2$, $41.21 \pm 5.23 \text{ cm}^2$). Cortical current densities ($\bar{\mathbf{s}}$) were simulated with varied numbers of patch sources (i.e. 1, 2, 3, 5) and each cortical current density with the same number of sources was repeated 200 times, in which sources were randomly selected. The dipole amplitude at each triangle was simulated as multiplication of its triangular area and the dipole moment density of 100 pAm/mm^2 (Jerbi et al., 2004). MEG signals were computed by first solving the forward problem using boundary element method (BEM) (Mosher et al., 1999) to obtain \mathbf{A} based on an 148 magnetometers MEG system (Magnes 2500 WH 4-D Neuroimaging Inc., San Diego, CA) and then multiplying \mathbf{A} to $\bar{\mathbf{s}}$. MEG signals were contaminated with Gaussian white noise and SNR was scaled to different levels (i.e. 20, 15, 10, and 5 dB):

$$\text{SNR} = 20 \log_{10} \frac{\|\mathbf{A} \bar{\mathbf{s}}\|_2}{\sqrt{M} \sigma_{\text{noise}}} \quad (7)$$

where $M = 148$ was the number of sensors and σ_{noise} was the standard deviation of noise.

The sensitivity of VW-SSI to the hyper-parameter λ and the level of wavelet compression in both W-SSI and VW-SSI were investigated. The performance of VW-SSI was compared to other techniques using both L2-norm regularizations, i.e. MNE, wMNE, and cortical LORETA (cLORETA), and L1-norm regularizations, i.e. SSI, V-SSI, and W-SSI. The implementation of L2-norm regularizations adopted the similar constrained optimization problem in Eq. (2): $\min \|\mathbf{H} \bar{\mathbf{s}}\|_2$

subject to $\|\bar{\boldsymbol{\phi}} - \mathbf{A} \bar{\mathbf{s}}\|_2 < \varepsilon$, where \mathbf{H} is an identity matrix in MNE (Hämäläinen and Ilmoniemi, 1994), a depth-weighted matrix in wMNE (Dale and Sereno, 1993), and the multiplication of a depth-weighted matrix and a Laplacian matrix in cLORETA (Pascual-Marqui et al., 1994). The selection of regularization parameter ε is the same as in Eq. (2). Four metrics were used to evaluate the performances of all methods, i.e. the accuracy in recovering source locations and their spatial extents. The first

metric, area under the receiver operating characteristic (ROC) curve, i.e., AUC (Grova et al., 2006), measured the sensitivity and specificity of detection of simulated cortical sources, which evaluated the accuracy of source locations and extents. The metrics of distance of localization error (DLE) and spatial dispersion (SD) (Molins et al., 2008) were adopted to assess reconstruction accuracies of source locations and extents, separately. By parceling estimated sources to the closest true sources, DLE was calculated as the averaged displacement from maxima of estimated sources to true sources, while the SD metric measured the spatial blurredness of estimated sources as compared with the extent of true sources. Relative mean square error (RMSE) (Grova et al., 2006) computed total errors from source amplitudes, locations, and extents. Formulae of AUC, DLE, SD and RMSE can be found in Appendix B. The values of these metrics were statistically compared among conditions using the test on median in boxplots (McGill et al., 1978) with the significance level of $p < 0.05$, paired t -test with Bonferroni correction when necessary, and ANOVA.

Experiment protocol

Experimental MEG data for a language task and a motor task were collected from one patient with medically refractory epilepsy, who underwent presurgical evaluation at Minnesota Epilepsy Group, St.

Paul, MN. The study protocol was approved by the Institutional Review Boards of University of Oklahoma Health Science Center, OK and United Hospital, St. Paul, MN. The language task used a continuous recognition memory paradigm (Breier et al., 1998; Papanicolaou et al., 2004). It consisted of 90 abstract English nouns, with 30 of them as targets and the remaining 60 words as distractors. The target words were presented once before MEG recordings and the patient was instructed to remember them as many as possible. During MEG recording sessions, target words were mixed with distractors and presented randomly, and the patient was asked to lift index finger when a target word was detected. The motor task was designed as a self-paced finger tapping task to push a button with the index finger every 1 to 2 s. MEG data were acquired using a whole head 148-channel magnetometer array (Magnes 2500 WH 4-D Neuroimaging Inc., San Diego, CA) at the sampling rate of 254.31 Hz in the language task and 290.64 Hz in the motor task.

MEG epochs were extracted with a window of -150 ms to 1000 ms with respect to the onset of word stimuli in the language task and a window of -45 ms to 300 ms with respect to the onset of movement cues. Band-pass filters of 0.1 to 20 Hz in the language task and 2 to 20 Hz in the motor task were employed. The pre-stimulus data were used as intervals to perform baseline correction in both datasets. Independent component analysis (ICA) was conducted to reject artifacts related to, such

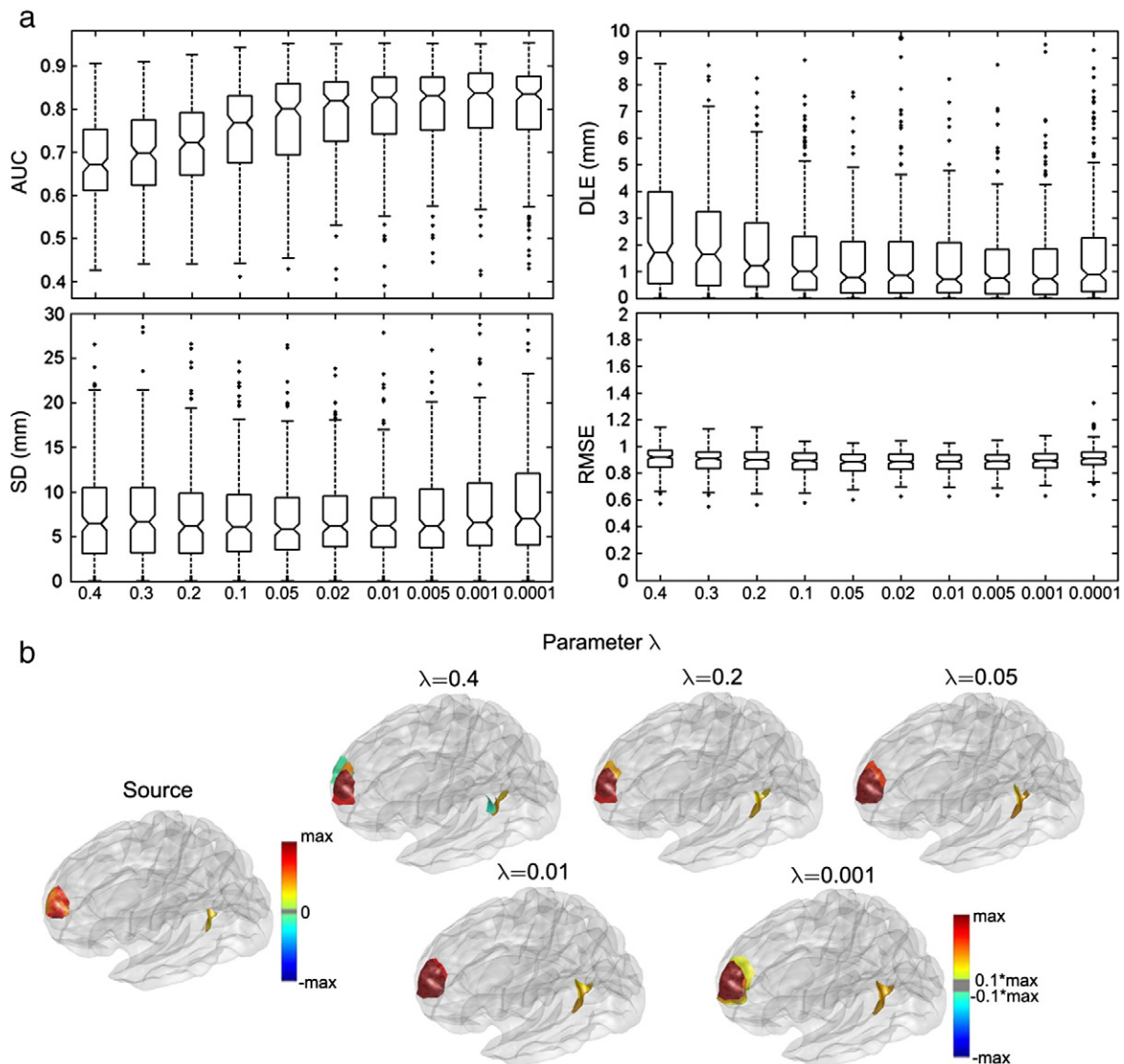


Fig. 2. Sensitivity of VW-SSI to the hyper-parameter λ . (a) Notched boxplots of four metrics at different values of λ . The notch indicates the 95% confidence interval on median. (b) An example of simulated source and reconstructions from VW-SSI with different λ values. The cortical surfaces were displayed with transparency.

as eye blinks, muscle movements, and heartbeats. Event-related fields (ERFs) were obtained by averaging remaining 120 epochs for the language task and 191 epochs for the motor task after artifact removal. Individual BE head models and CCD models were built from the patient's T1-weighted MRI data using the same protocol in simulation. Co-registration between CCD models, BE models, and MEG sensors was conducted by aligning fiducial points (left and right pre-auricular points and nasion) and further refined by fitting digitized head points to BE models.

Results

Influence of the hyper-parameter λ

It was found that L1-norms of wavelet coefficients are approximately five times as large as L1-norms of variations from 200 simulated cases. The sensitivity of VW-SSI on λ (using \mathbf{W}_1) was tested by selecting

the hyper-parameter at 10 different values (Fig. 2(a)) toward the preference of either wavelet constraint ($\lambda > 0.2$) or variation constraint ($\lambda < 0.2$). As shown in Fig. 2(a), the localization accuracy increases (i.e. DLE decreases) for λ values of 0.4 to 0.1, and then shows of no significant changes for λ values of 0.05 to 0.0001. It is observed that the median estimation accuracy of source extent of VW-SSI, indicated by AUC, significantly increases when placing more weights on the variation constraint (λ from 0.4 to 0.02). The accuracy is also significantly different between $\lambda = 0.4$ and 0.02 in a t -test ($p < 0.001$), and then shows no significant changes when $\lambda < 0.02$. When λ continuously decreases from 0.01 to 0.0001, the blurredness starts to degrade the performance of VW-SSI in estimating source extents, i.e. significantly increased SDs ($p < 0.01$). The total estimation errors, otherwise, suggest no significant changes, as indicated by RMSE. In Fig. 2(b), a selected example indicates that the estimated sources at $\lambda = 0.01$ have the most accurate location and extent estimations. It is observed that spurious sources show up surrounding the true sources at $\lambda = 0.4$ and blurred estimations appear

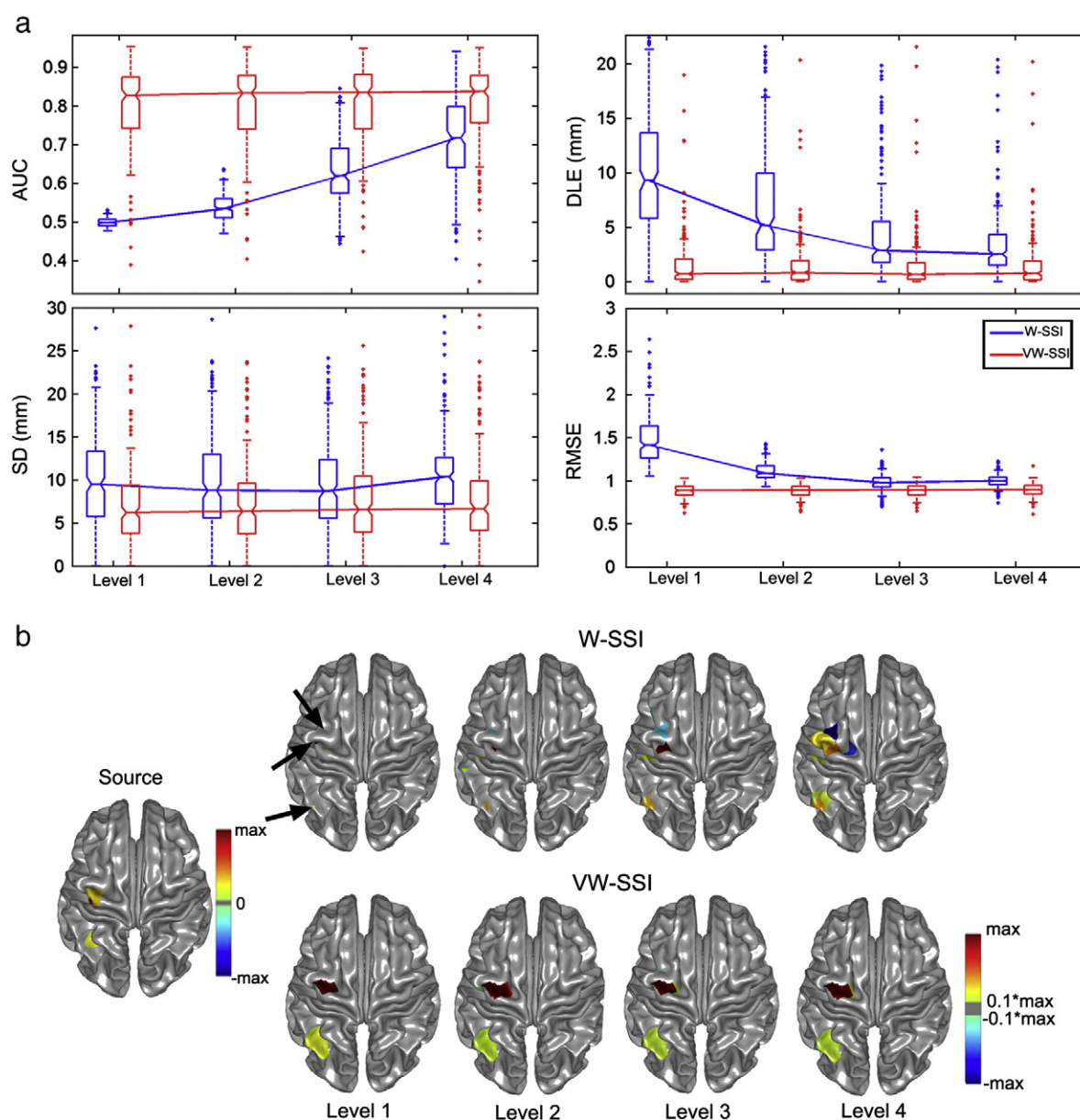


Fig. 3. (a) Notched boxplots of four metrics for W-SSI and VW-SSI with different wavelet compression levels. (b) An example of simulated sources and reconstructions from W-SSI and VW-SSI at the wavelet compression level from one to four.

when λ is between 0.01 and 0.001, collaborated with the increased SD values in Fig. 2(a). Taking all metrics into considerations, the hyper-parameter λ was set to 0.01 in the following analyses.

Influence of the wavelet compression level

The influence of the wavelet compression level was investigated in both VW-SSI and W-SSI. It is observed that the performance of W-SSI is more sensitive to the wavelet compression levels than that of VW-SSI, indicated by significantly increased AUCs (medians, $p < 0.05$; $p < 0.001$ in t -test), and significantly decreased DLEs (medians, $p < 0.05$; $p < 0.001$ in t -test) (Fig. 3(a)). Through the integration of variation and wavelet constraints, VW-SSI is much less influenced by the parameter of wavelet compression level and has much more stable and better performance than W-SSI in all four metrics (i.e. AUC, DLE, SD, and RMSE). The insensitivity of VW-SSI to the wavelet compression level is further visually inspected on estimated sources at four levels, which have similar spatial locations and extents (Fig. 3(b)). These reconstructions are also better than reconstructions from W-SSI, which

show over-focused sources with fewer levels of wavelet compressions and blurred sources in more levels of compressions (Fig. 3(b)), as indicated by the increased SD values (Fig. 3(a)). Since VW-SSI is insensitive to the compression level, the one-level wavelet compression was adopted in the following analyses.

Comparisons with L2-norm regularizations

In Fig. 4(a), with comparisons to classic L2-norm regularizations, i.e. MNE, wMNE, and cLORETA, the proposed VW-SSI method indicates significantly better performance (medians, $p < 0.05$; $p < 0.01$ in t -test) than all L2-norm regularizations in all metrics. The relatively high AUC values in MNE and wMNE (medians close to 0.8) are probably because the metric AUC favors blurred source estimations (Liao et al., 2012; Petrov, 2012) and estimated sources from L2-norm regularizations are spatially blurred, as suggested by high SD values (Fig. 4(a)). The median SD values of MNE, wMNE, and cLORETA are about 20 mm, 25 mm, and 40 mm, respectively, while the median SD value of VW-SSI is only about 5 mm. VW-SSI also reveals much less biases in source locations (the median

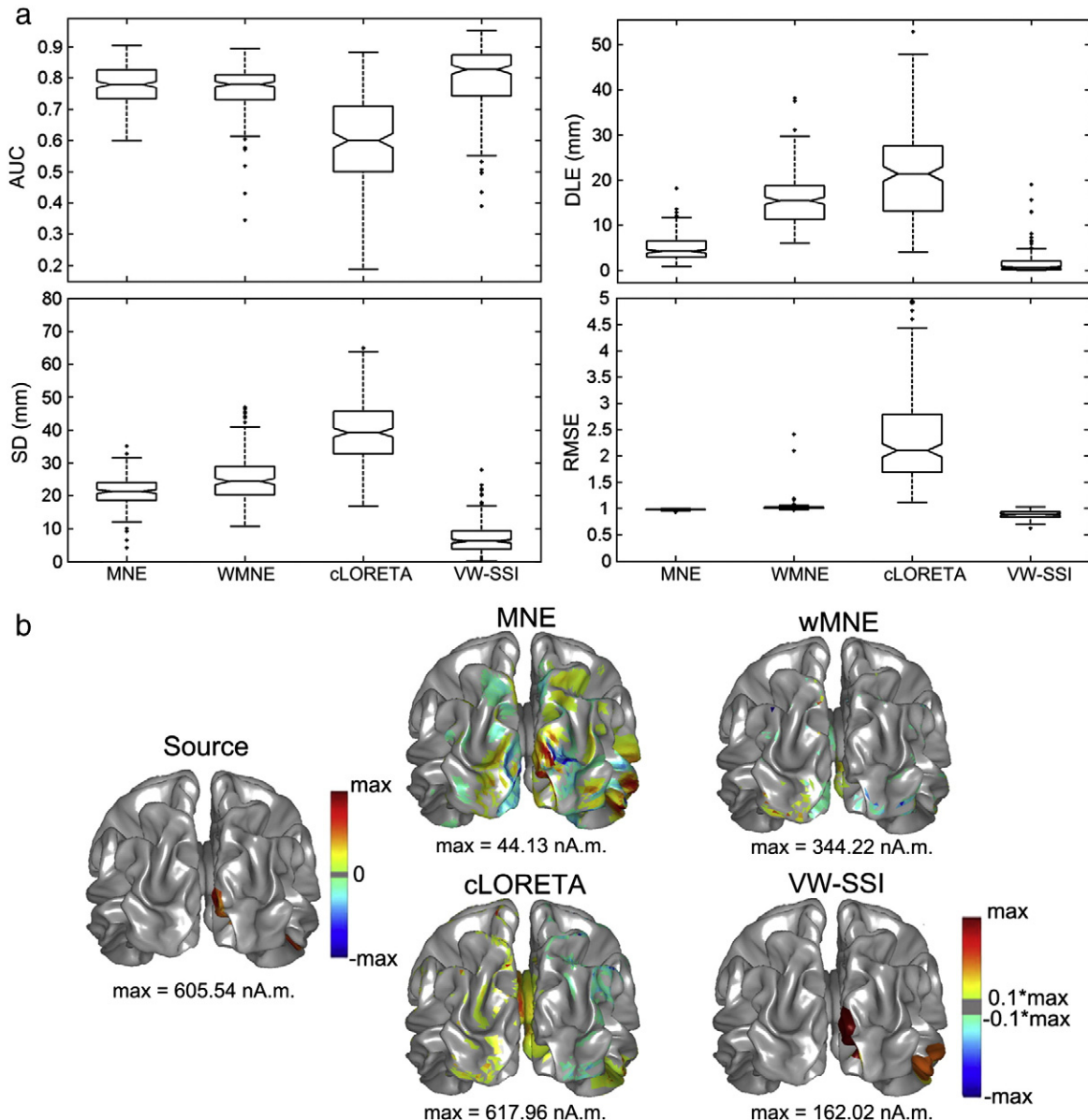


Fig. 4. (a) Notched boxplots of four metrics for MNE, wMNE, cLORETA and VW-SSI. (b) An example of simulated sources and reconstructions from compared methods.

value of DLE is 1 mm) than L2-norm regularizations (5 mm in MNE, 15 mm in wMNE, and 20 mm in cLORETA). While RMSE from VW-SSI is significantly lower than all L2-norm regularizations ($p < 0.001$), the source strengths are still underestimated (maximal values: 162 nA·m in estimated sources and 606 nA·m in simulated sources). The selected example (Fig. 4(b)) also indicates the better performance of VW-SSI than MNE, wMNE, and cLORETA. The sources estimated from MNE, wMNE and cLORETA are observed with problems of diffused distributions (in MNE), spurious sources surrounding true sources (in MNE and cLORETA) and even at areas far from true sources (in cLORETA), and missed sources (in wMNE and cLORETA).

Comparisons to other L1-norm regularizations

In Fig. 5, the performance of VW-SSI was also compared to those of other L1-norm methods, i.e. SSI, V-SSI and W-SSI with a level-four compression (denoted as W₄-SSI). VW-SSI shows significantly better performance than SSI and W₄-SSI in all metrics (medians, $p < 0.05$; $p < 0.01$ in t -test). It also has significantly lower SD ($p < 0.001$), DLE ($p < 0.05$), and RMSE values ($p < 0.05$) than V-SSI. It is noted that SSI, as compared

with other three L1-norm methods, has significantly lower AUC values, higher DLE and RMSE values (medians, $p < 0.05$; $p < 0.001$ in t -test). The selected example (Fig. 5(b)) also indicates the better performance of VW-SSI in estimating both source locations and extents than those of the other three methods. The sources estimated from SSI are only active on few elements (about 7 mm²), which are over-focused as compared with simulated ones (about 743 mm²). Spurious sources are observed in W₄-SSI surrounding true sources and, in V-SSI, spatially more extended sources are observed, which are all suppressed in VW-SSI.

Effect of source extent

The performance of VW-SSI was further investigated with varied source extents (two patch sources) as compared with those of MNE and V-SSI (Fig. 6(a)). All three methods indicate significantly degraded AUC values (medians, $p < 0.05$; $p < 0.01$ in t -test), when the size of sources increases. In contrast, DLE, SD, and RMSE of VW-SSI become slightly, but not significantly, better. It is observed that VW-SSI has significantly better performance (medians, $p < 0.05$; $p < 0.01$, t -test) than MNE in all four metrics and all source extents. VW-SSI also shows

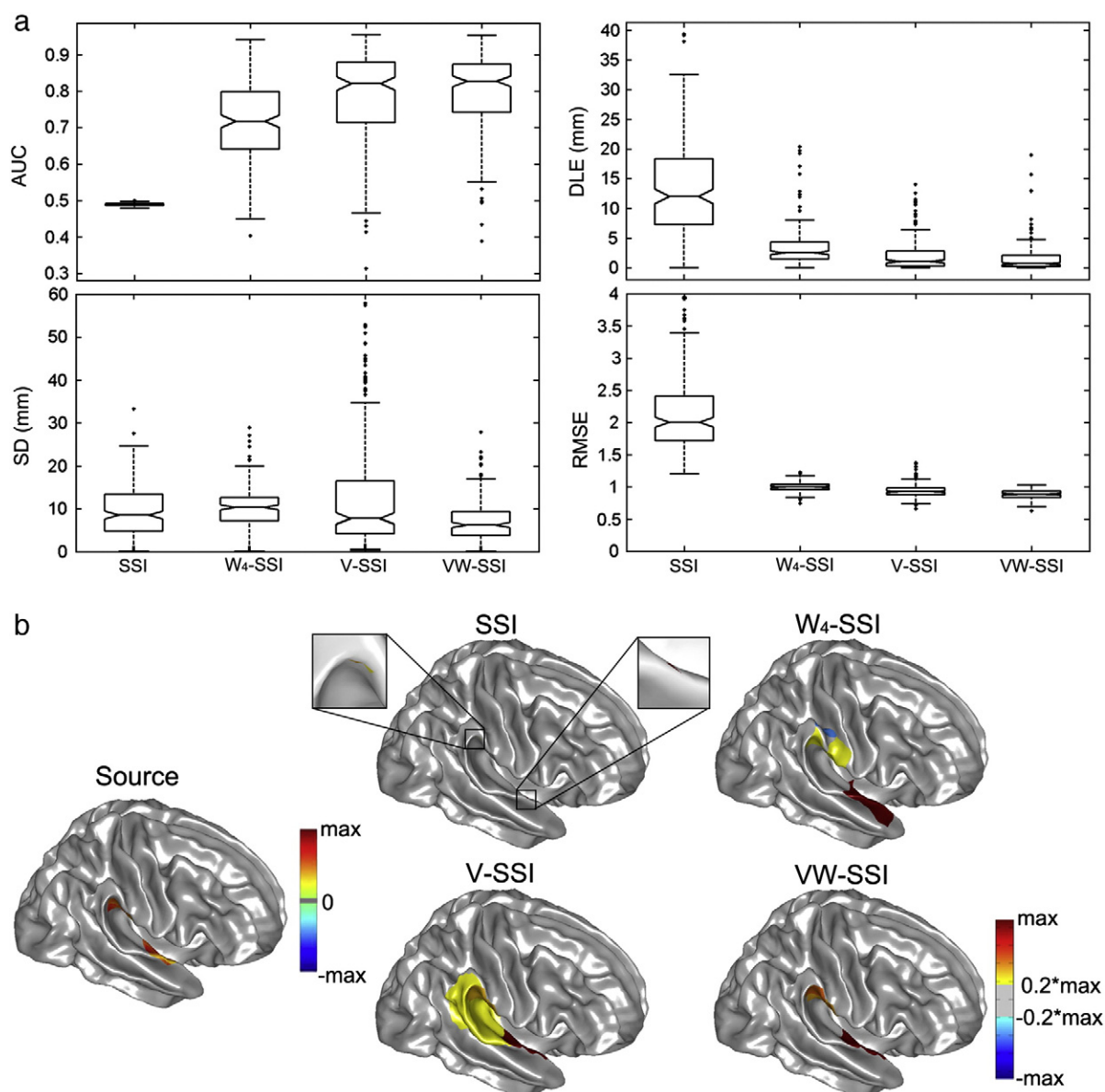


Fig. 5. (a) Notched boxplots of four metrics for SSI, W₄-SSI, V-SSI and VW-SSI. (b) An example of simulated sources and reconstructions from compared methods.

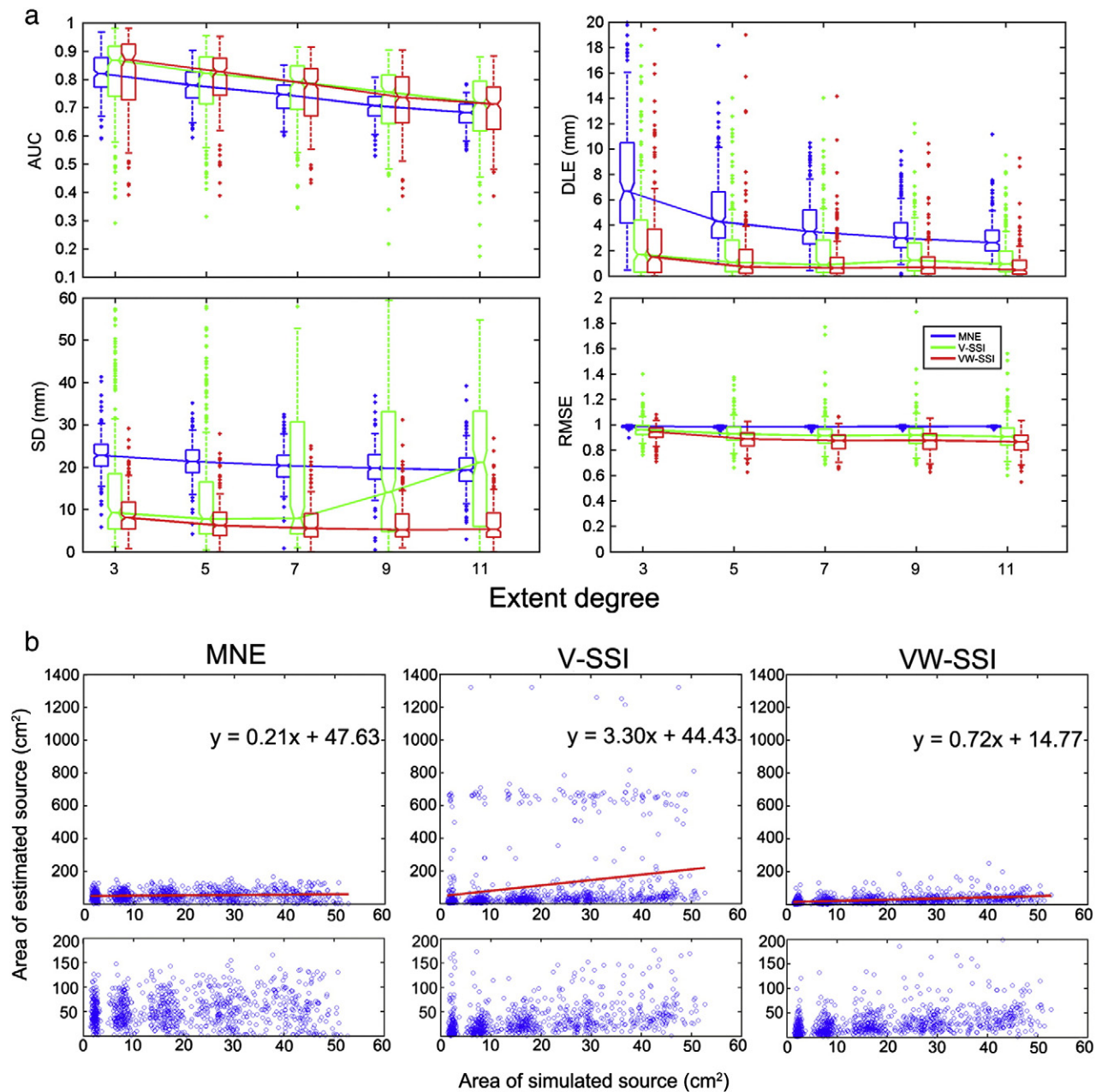
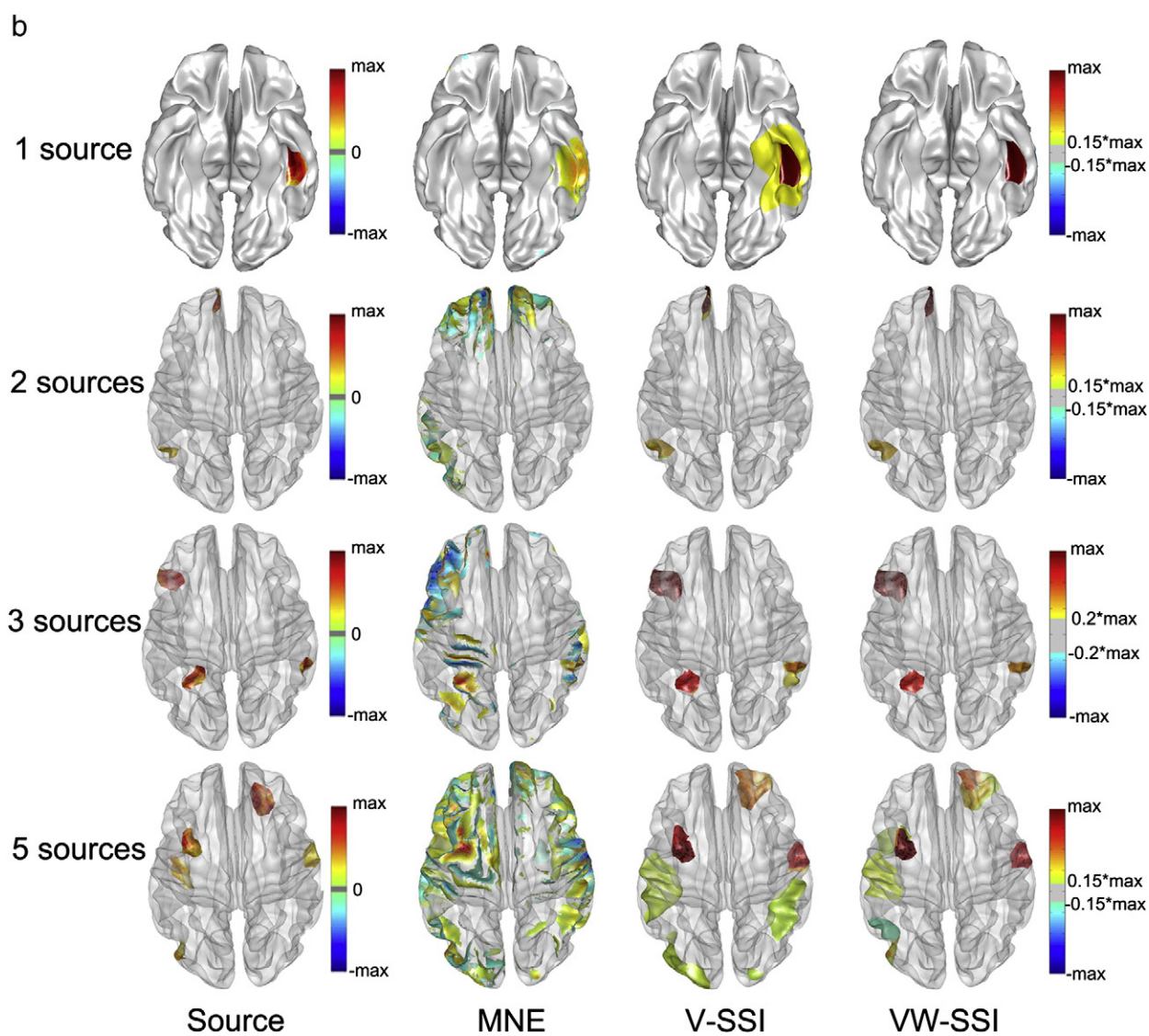
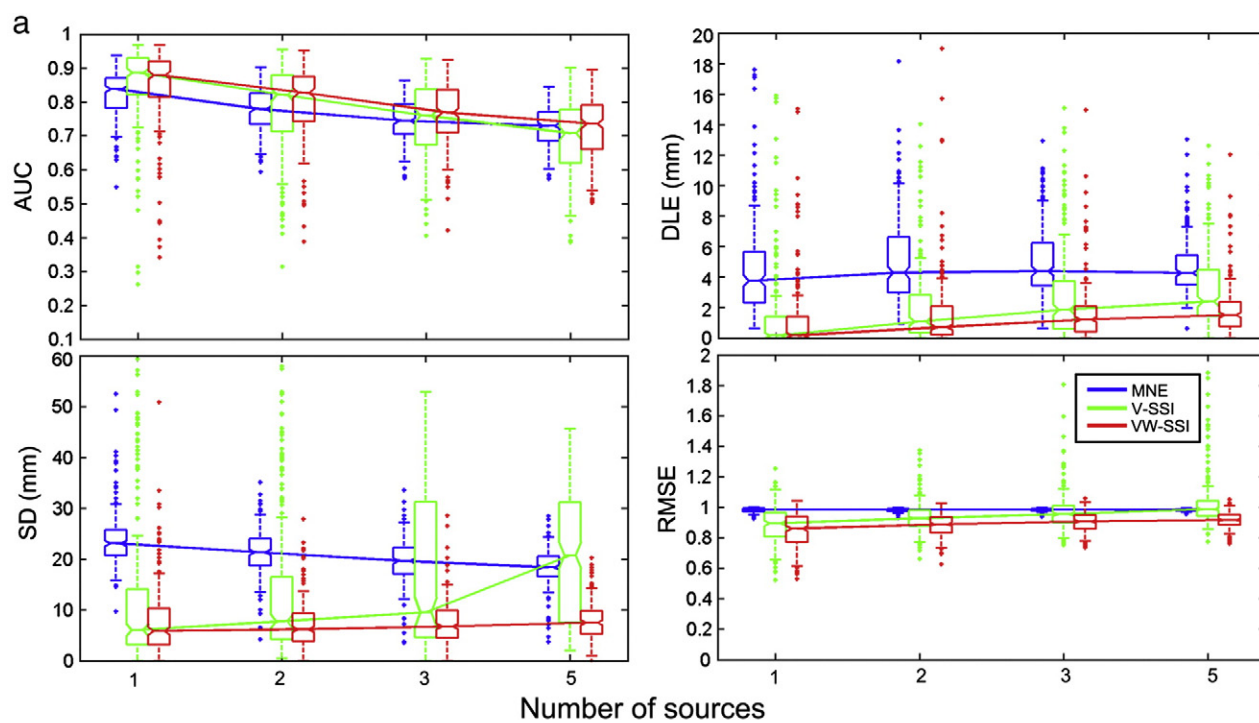


Fig. 6. (a) Notched boxplots of four metrics for MNE, V-SSI and VW-SSI at different source extents. (b) Linear regressions (red lines) on extent data of simulated sources (x-axis) and extent data of estimated sources (y-axis) from MNE, V-SSI, and VW-SSI. Each circle indicates the result from one of 1000 simulations for five different source extents. The underneath is the zoom-in of the lower parts of graphs with y-axis from 0 to 200 cm².

significantly better performance ($p < 0.01$, t -test) than V-SSI in DLE, SD, and RMSE. Each circle in Fig. 6(b) indicates the relation between true source extent and estimated source extent (thresholded at 20% of individual maxima to reject more background activity in MNE) from each simulation. Its linear regression results (Fig. 6(b)) indicate that VW-SSI has the better preservation of extent contrast in simulated sources with the slope of 0.72 (close to 1) than those of MNE (slope: 0.21), which underestimates extent contrast, and V-SSI (slope: 3.30), which overestimate extent contrast. It is noted that V-SSI has a better extent contrast pattern when estimated source extent below 200 cm². However, its solutions sometime suffer from the unconstrained global power problem, which leads to unreasonable source extent estimations (beyond 600 cm²). Furthermore, the relatively large intercept (~14 cm²) in VW-SSI suggests the existence of systematic bias, while it is only about one-third of those from MNE and V-SSI.

Effect of number of sources

VW-SSI indicates significant decreases in AUC ($p < 0.01$, t -test) and slight increases in other three metrics when the number of sources increases (extent degree 5) (Fig. 7(a)). VW-SSI indicates significantly better performance than MNE in all four metrics and all numbers of sources (t -test, $p < 0.01$) except for AUC with 5 sources. As compared to V-SSI, VW-SSI has better performance with multiple sources, indicated by significantly higher AUCs (t -test, $p < 0.01$) in simulations with 3 and 5 sources and significantly lower DLEs (t -test, $p < 0.01$) in simulations with 2, 3, and 5 sources. VW-SSI also indicates better performance than V-SSI with significantly lower SDs and RMSEs (t -test, $p < 0.01$) in all numbers of sources. Examples in Fig. 7(b) suggest that both VW-SSI and V-SSI recover more accurate source locations and extents in the conditions with 1, 2, and 3 source(s) than MNE. In the condition



with 5 sources, V-SSI exhibits worse performance than VW-SSI since spurious sources are observed in V-SSI.

Sensitivity to SNR

In Fig. 8, all four metrics at different SNR levels indicate that the performance of W_4 -SSI, V-SSI and VW-SSI significantly depends on noises (AUC: $p < 0.01$, ANOVA; DLE: $p < 0.005$, ANOVA). Among them, V-SSI indicates the highest sensitivity to SNR with a dramatic increase of SD values at both 10 dB and 5 dB, and significant decreases in both AUC and DLE as well, which is because the global energy is not limited in V-SSI (Ding, 2009) and becomes severe when SNR is low. While both W_4 -SSI and VW-SSI reveal similar degrading patterns in performance as SNR decreases, VW-SSI always suggest significantly better performances (medians, $p < 0.05$; $p < 0.001$ in t -test) than W_4 -SSI in all metrics at all SNR levels. In the examples in Fig. 8(b), spurious sources (black circles) observed in W_4 -SSI and blurred distributions observed in V-SSI at low SNRs (e.g. 5 dB) are sufficiently suppressed in VW-SSI.

Experimental results

In the language task, the estimated sources from VW-SSI (Fig. 9) locate on the right superior temporal cortex within the coverage of Brodmann areas (BAs) 41 and 42 at the latency of 94 ms, which suggests the early sensory response in the primary auditory cortices around 100 ms (Zouridakis et al., 1998). At the latency of 232 ms, the estimated sources from VW-SSI show up on the bilaterally medial temporal cortices, suggesting the late language process occurred between 200 ms and 300 ms post-stimulus, associated with the decoding of phonological and semantic components of words (Breier et al., 1998; Frye et al., 2009). The sources estimated with MNE and wMNE are scattered on multiple brain areas including the right temporal, right anterior parietal, and left temporal cortices similarly at both latencies. cLORETA produces more smoothed sources covering similar brain areas as MNE and wMNE. More importantly, their estimated sources at different timings (i.e. 94 ms and 232 ms) do not reveal observable distinctions among the early sensory response and the late language processing. The estimated sources from V-SSI indicate similar reconstructions on BAs 41, 42 at 94 ms, while the estimations are obviously suffered from unconstrained global powers and become worse at 232 ms. SSI and W-SSI produce more compact estimations on the temporal cortices at both times. However, both SSI and W-SSI indicate spurious extra-temporal activations and their temporal sources at 232 ms more toward the inferior part of the temporal lobe.

In the finger tapping task, the estimated sources from VW-SSI (Fig. 10) indicate outward currents on the anterior bank of the central sulcus (the primary motor cortex, M1) at 45 ms and outward currents on the posterior bank of the central sulcus (the primary somatosensory cortex, S1) at 151 ms, which is well consistent with the generators of motor field (0–100 ms) from M1 and post-movement motor field (>120 ms) from S1 (Gerloff et al., 1998; Pollok et al., 2003). The sources estimated from MNE, wMNE, cLORETA and V-SSI are diffused (V-SSI also suffers from unconstrained global powers) and not able to distinguish activations from M1 and S1. The sources estimated from SSI and W-SSI are confined on S1 at both 45 ms and 151 ms with currents changing from inward to outward.

Discussion

In the present study, a new L1-norm regularization, i.e. VW-SSI, was proposed to solve ill-posed EEG/MEG inverse problems, which enforced

sparseness in more than one transformed domain, i.e. variation and wavelet domains. Monte-Carlo simulations using MEG data were conducted with multiple extended sources that were randomly selected to mimic distributed brain activations. The proposed L1-norm regularization showed significantly better accuracies in estimating source locations and extents than L2-norm regularizations (i.e. MNE, wMNE, and cLORETA) and other L1-norm regularizations (i.e. SSI, V-SSI, and W-SSI). Using MEG data from two experimental tasks, it was demonstrated that VW-SSI could successfully reconstruct cortical sources revealing well-documented neural activation phenomena in human language and motor functions while all other L2-norm and L1-norm regularizations were not able to recover these phenomena.

With the ultimate goal of EEG/MEG inverse solutions to recover cortical sources with exact locations, extents, and amplitudes, most available solutions focus on locations, such as localizations of dipole sources (Stefan et al., 2003; Wood, 1982) and extended sources (Chowdhury et al., 2013). It has been demonstrated that brain sources from the same cortical areas with different extents generate different patterns in surface electromagnetic signals (Ding and Yuan, 2013). The goal of the proposed method is to achieve better estimation of source extents, beyond source localization, which is of significance for applications in neuroscience and neurology (Dhond et al., 2001; Hillyard, 1993; Petrov, 2012). As an example, the identification of eloquent brain areas of functional importance and/or epileptogenic zones and the decision on whether they are overlapped are essential for the presurgical planning of brain surgery in epilepsy (Brodbeck et al., 2011; Plummer et al., 2008), which all need critical extent information besides locations. While L2-norm regularizations based on distributed source models (Dale and Sereno, 1993; Hämäläinen and Ilmoniemi, 1994; Pascual-Marqui et al., 1994) have been used to recover distributed brain sources, their inverse solutions are spatially diffused as suggested by the present data (AUC and SD data in Fig. 4) and literatures (Grova et al., 2006; Vega-Hernández et al., 2008). Due to the blurredness, their estimated sources are easily mislocalized (DLE data in Figs. 4(a), 6(a), 7(a)). Due to the lack of considering spatial continuity (in MNE and wMNE), L2-norm inverse solutions suffer from the spatial discontinuity problem in source distributions (Figs. 4(b), 7(b)), which is difficult to separate multiple sources and estimate their extents (Fig. 7(b)). The L1-norm inverse solutions (i.e. SSI) obtained by enforcing sparseness in the original source domain (Tibshirani, 1996; Uutela et al., 1999), otherwise, suffer from the over-focality problem (Chang et al., 2010). V-SSI and W-SSI, enforcing sparseness in transform domains, indicate better performance in recovering source extents, as suggested by the present data (AUC and SD data in Fig. 5) and literatures (Chang et al., 2010; Ding, 2009; Liao et al., 2012). However, the performance of W-SSI highly depends on the compression scheme (e.g. spherical wavelets vs. face-based wavelets) (Liao et al., 2012) and the compression level as suggested in Fig. 3 and literatures (Chang et al., 2010). Due to the unconstrained global power, inverse solutions in V-SSI can potentially have uncontrolled global activity (Figs. 8(b), 9(b), 10(b)). The VW-SSI method takes the advantages of both V-SSI (its capability in recovering source boundaries and thus source extents) and W-SSI (its capability in compressing sources for better sparse reconstructions), which results in significant improvements in estimating source locations (e.g. better DLE) and extents (i.e. better AUC and SD) than L2-norm and other L1-norm regularization methods. It is also noted that VW-SSI has significantly better performance in preserving extent contrast among sources of different sizes than other L2-norm and L1-norm methods (Fig. 6(b)). However, a systematic bias (the intercept from the linear regression in Fig. 6(b)) in estimating source extents still exists in VW-SSI, which indicates the overestimation of source

Fig. 7. (a) Notched boxplots of four metrics for MNE, V-SSI and VW-SSI at different numbers of sources. (b) Examples of simulated sources and reconstructions from compared methods. The cortical surfaces were displayed with transparency when necessary to visualize all sources at the same visual angle.

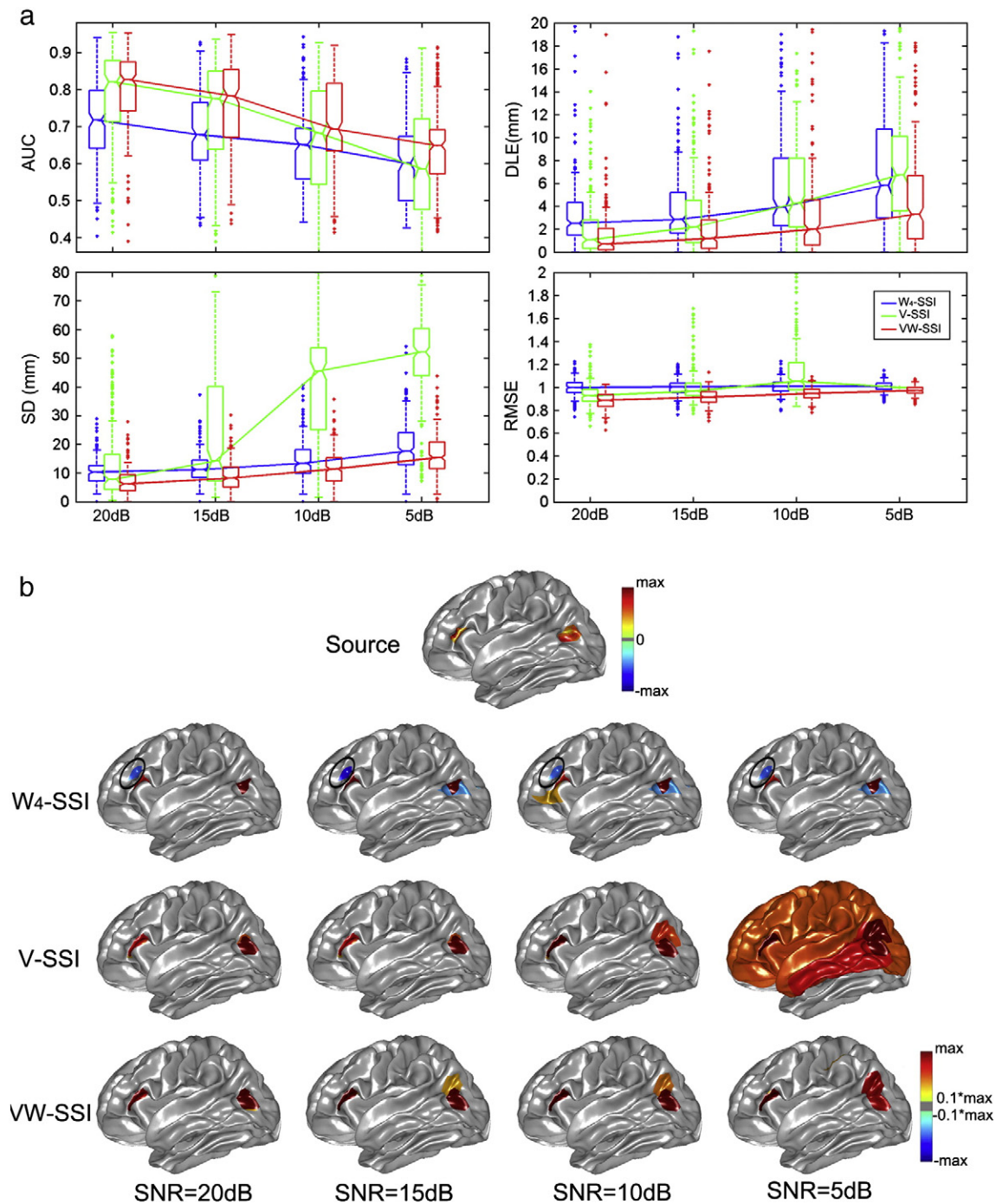


Fig. 8. (a) Notched boxplots of four metrics for W₄-SSI, V-SSI and VW-SSI at different SNR levels. (b) An example of simulated sources and estimated sources from compared methods at different SNR levels.

extents, as collaborated by underestimated source strengths (Fig. 4(b)). Such a constant bias can be further reduced if multimodal data (i.e. EEG and MEG) are integrated (Ding and Yuan, 2013).

The combination of variation and wavelet penalties in VW-SSI is also demonstrated with the robust performance on influential factors, such as wavelet compression level and noise. While the performance of W-SSI largely depends on the compression level, the performance of VW-SSI is not significantly influenced by this factor, which is achieved with the use of variation penalty (Fig. 3). The present results further indicate that VW-SSI is less sensitive to noise than V-SSI (Fig. 8).

In a language task, VW-SSI successfully recovers the dynamic neural activations (Fig. 9) from early auditory response (Zouridakis et al., 1998) to late language processing (Breier et al., 1998; Frye

et al., 2009). From the spatial locations and extents of these sources, neural activations due to auditory responses (over superior temporal cortex) and language processes (over medial temporal cortex) are well distinguished in VW-SSI inverse solutions. All other methods were not able to recover these neural activations, except W-SSI, which recovers several patterns as seen in VW-SSI with degradations. Due to the ill-posed nature of EEG/MEG inverse problems, it has been suggested challenging to recover cortical sources across a sulcus (Ou et al., 2009) since sources on the pre-bank and post-bank of a sulcus can generate similar surface EEG/MEG signals by just flipping source orientations. In Fig. 10, VW-SSI demonstrates the capability in distinguishing neural activations from the pre-bank and post-bank of the central sulcus in a motor task, which have been

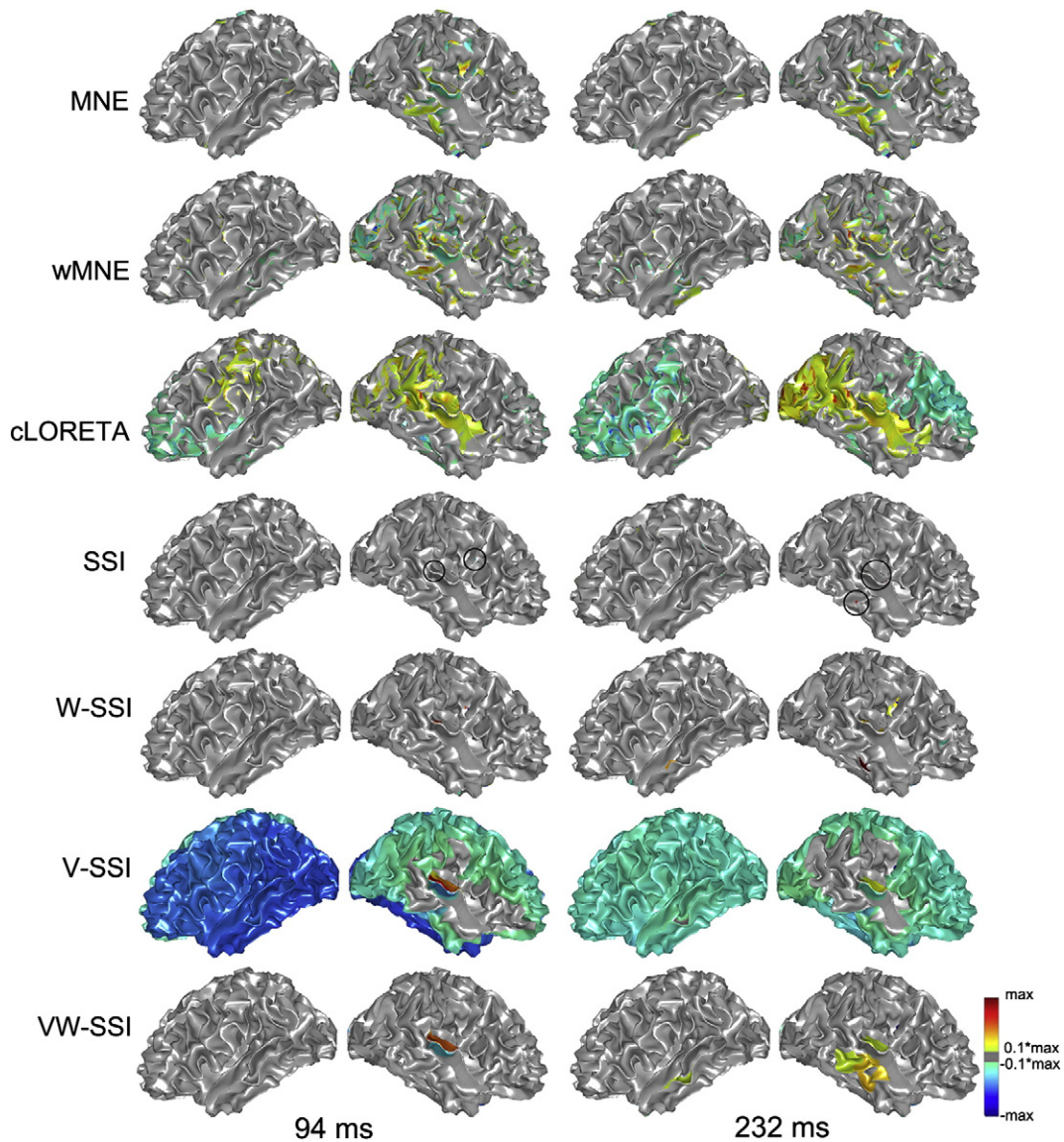


Fig. 9. Estimated sources of early auditory response and late language processing from MNE, wMNE, cLORETA, SSI, V-SSI, W-SSI and VW-SSI at 94 ms and 232 ms, respectively, after the onset of stimulus.

well studied using fMRI (Jäncke et al., 2000). Meanwhile, all other methods investigated fail to recover such sources.

One contribution of the present study is to evaluate the feasibility and performance of utilizing multiple transform sparseness in the L1-norm regularization. Multiple penalties have been used in regularization for years to select preference on solutions to either smoothness or

sparseness (Chang et al., 2010; Haufe et al., 2008; Hebiri and van de Geer, 2011; Tibshirani et al., 2005; Vega-Hernández et al., 2008; Zou and Hastie, 2005). Ridge fused (Vega-Hernández et al., 2008) combines two L2-norm penalties used in MNE and LORETA to achieve controlled blurredness in inverse solutions. Methods, such as, Elastic Net (Zou and Hastie, 2005), Elastic Net-L (Vega-Hernández et al., 2008), and

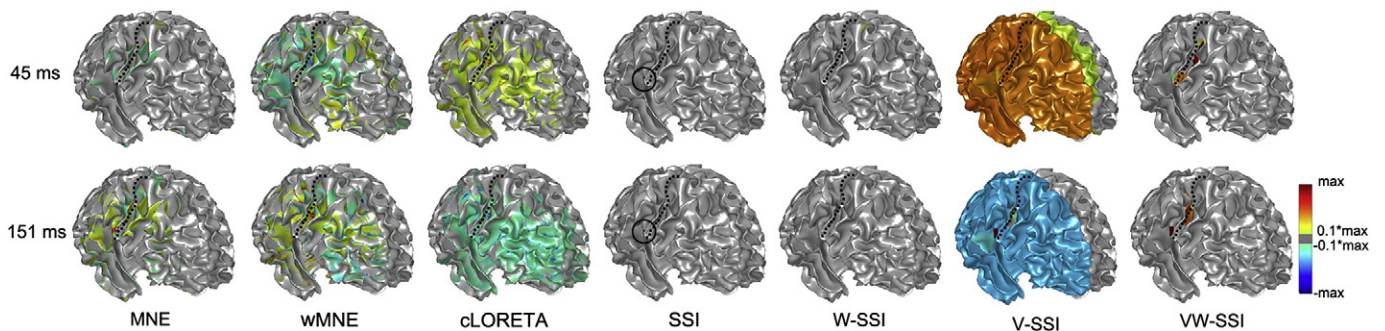


Fig. 10. Estimated sources of the motor task from MNE, wMNE, cLORETA, SSI, V-SSI, W-SSI and VW-SSI at 45 ms and 151 ms after the onset of stimulus. Black dashed line: the central sulcus.

smooth LASSO (Hebiri and van de Geer, 2011), have been developed to compensate the smoothness from the L2-norm penalty using the L1-norm penalty. Other methods, e.g. fused LASSO (Tibshirani et al., 2005), focal vector field reconstruction (Haufe et al., 2008), and compressive neuromagnetic tomography (Chang et al., 2010), apply sparseness in the original source domain as SSI and, at the same time, constrain local smoothness using either successive differences or spatial transforms. The present study, for the first time, enforces sparseness simultaneously in two transform domains (i.e. variation and wavelet domains). While the new method has demonstrated better accuracy in recovering source location and extent and improved performance in low SNRs as compared with other regularizations using single penalties, its advantages over regularizations using multiple penalties need to be investigated in the future.

It is important to note that one limitation of using multiple penalties in a regularization problem is how to decide the value of hyper-parameter. While some strategies have been developed for L2-norm regularizations (Vega-Hernández et al., 2008), there are few studies addressing such an issue in L1-norm regularizations. The present simulation results suggest that the proposed method is quite robust to the hyper-parameter in a wide range of values. The experimental results also indicate the stable performance of VW-SSI with the hyper-parameter value selected based on simulation results. Nevertheless, more studies, especially those about criteria on automatically selecting the parameter, will be needed, particularly when there are other complicating factors, e.g. wavelet compression level. It is also noted that the present study only provides experimental data to qualitatively evaluate the accuracy in recovering source extents. The quantitative evaluation would need other independent measurements, such as fMRI (Sharon et al., 2007) or electrocorticography (ECoG) (Ding et al., 2007).

Acknowledgments

This work was supported in part by OCAST HR09-125S, NSF CAREER ECCS-0955260, and DOT-FAA 10-G-008. Some of the computing for this project was performed at the OU Supercomputing Center for Education & Research (OSCER) at the University of Oklahoma (OU).

Appendix A

If T_i^m is denoted as a set of triangles on the mesh at the finer level m that merge to the j^{th} triangle at the coarser level $m+1$, each entry a_{ji}^m in the synthesis matrix \mathbf{A}^m is computed as the proportion of triangular area of intersection between T_j^{m+1} and T_i^m to the area of T_j^{m+1} :

$$a_{ji}^m = \frac{\text{area}(T_j^{m+1} \cap T_i^m)}{\text{area}(T_j^{m+1})}. \quad (\text{A.1})$$

The analysis matrix \mathbf{B}^m is designed to have the property of semi-orthogonality in basis functions (Bonneau, 1996) and obtained by solving the following set of equations (see details in Liao et al., 2012):

$$\begin{cases} \mathbf{A}^m (\mathbf{G}^m)^{-1} (\mathbf{B}^m)^T = \mathbf{0} \\ \mathbf{B}^m (\mathbf{G}^m)^{-1} (\mathbf{B}^m)^T = \mathbf{I} \end{cases} \quad (\text{A.2})$$

where \mathbf{G}^m is the Gram–Schmidt matrix, i.e. a diagonal matrix with its diagonal elements being the areas of triangles at level m .

Appendix B

The AUC metric is derived from the ROC analysis, which is a graphic plot of the true positive rate (TPR) against false positive rate (FPR) at various values of the threshold α :

$$\begin{aligned} \text{TPR}(\alpha) &= TP(\alpha) / (TP(\alpha) + FN(\alpha)); \quad \text{FPR}(\alpha) \\ &= FP(\alpha) / (TN(\alpha) + FP(\alpha)) \end{aligned} \quad (\text{B.1})$$

where $TP(\alpha)$ and $FP(\alpha)$ are the numbers of active elements in estimations inside or outside simulated source patches. $FN(\alpha)$ and $TN(\alpha)$ are the numbers of inactive elements inside or outside simulated source patches. To reduce bias in estimating AUC because of unbalanced numbers of active (~ 100) and inactive elements ($\sim 10^4$) in EEG/MEG inverse problems, a procedure (Grova et al., 2006), which randomly samples (50 times) inactive elements to the same number as active elements from both close fields and far fields of simulated sources, is used.

The SD metric (Chang et al., 2010; Molins et al., 2008) is defined as:

$$\text{SD} = \sqrt{\frac{\sum_{k=1}^N \sum_{i \in I_k} d_{ki}^2 \|\hat{s}_i\|_2^2}{\sum_{i=1}^n \|\hat{s}_i\|_2^2}}, I_k = \left\{ i | k = \arg \min_{k'} \{d_{k'i}\} \right\}, 1 \leq i \leq n, 1 \leq k \leq N, 1 \leq k' \leq N \quad (\text{B.2})$$

where \hat{s}_i is the amplitude of i th dipole in the estimated current density and d_{ki} defines the distance from the k th dipole inside the simulated current density (consisting of multiple simulated patch sources) to the i th dipole in the estimated current density. I_k is the index set of dipoles in the estimated current density, which is spatially closest to the k th dipole in the simulated current density. N and n denote the number of dipoles in the simulated current density and the number of dipoles in the estimated current density, respectively.

The DLE metric (Chang et al., 2010; Molins et al., 2008) is defined as the average distance between the location of each dipole (k th) in the simulated current density to the maxima of its closest dipoles (from I_k) in the estimated current density:

$$\begin{aligned} \text{DLE} &= \frac{1}{N_J} \sum_{k \in J} \text{DLE}_k, \quad J = \{k | I_k \neq \emptyset\}, \quad \text{DLE}_k \\ &= \left\{ d_{ki} | i = \arg \max_{i'} (\|\hat{s}_{i'}\|_2), i' \in I_k \right\} \end{aligned} \quad (\text{B.3})$$

where N_J is the number of elements in set J . The definitions of i , k , I_k and \hat{s}_i are the same as in Eq. (B.1).

The relative mean square error (RMSE) (Grova et al., 2006) is defined as:

$$\text{RMSE} = \frac{\sum_{i=1}^N \|\hat{s}_i - s_i\|_2^2}{\sum_{i=1}^N \|s_i\|_2^2} \quad (\text{B.4})$$

where s_i is the amplitude of i th dipole in simulated sources.

References

- Allison, T., McCarthy, G., Wood, C., Darcey, T., Spencer, D., Williamson, P., 1989. Human cortical potentials evoked by stimulation of the median nerve. II. Cyto-architectonic areas generating short-latency activity. *J. Neurophysiol.* 62, 694–710.
- Baillet, S., Mosher, J., Leahy, R., 2001. Electromagnetic brain mapping. *IEEE Signal Proc. Mag.* 18, 14–30.
- Bonneau, G.-P., 1996. Multiresolution analysis with non-nested spaces. *Proc. Dagstuhl J. Geom. Mod.* 1996, 43–48.
- Breier, J.I., Simos, P.G., Zouridakis, G., Papanicolaou, A.C., 1998. Relative timing of neuronal activity in distinct temporal lobe areas during a recognition memory task for words. *J. Clin. Exp. Neuropsychol.* 20, 782–790.
- Brodbeck, V., Spinelli, L., Lascano, A., Wissmeier, M., Vargas, M., Vulliemoz, S., Pollo, C., et al., 2011. Electroencephalographic source imaging: a prospective study of 152 operated epileptic patients. *Brain* 134, 2887–2897.
- Candès, E., Romberg, J., 2007. Sparsity and incoherence in compressive sampling. *Inverse Probl.* 23, 969–985.
- Chang, W., Nummenmaa, A., Hsieh, J., Lin, F., 2010. Spatially sparse source cluster modeling by compressive neuromagnetic tomography. *Neuroimage* 53, 146–160.
- Chowdhury, R.A., Lina, J.M., Kobayashi, E., Grova, C., 2013. MEG source localization of spatially extended generators of epileptic activity: comparing entropic and hierarchical Bayesian approaches. *PLoS One* 8, e55969.

- Dale, A.M., Sereno, M.I., 1993. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. *J. Cogn. Neurosci.* 5, 162–176.
- Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis. *Neuroimage* 9, 179–194.
- Dhond, R.P., Buckner, R.L., Dale, A.M., Marinkovic, K., Halgren, E., 2001. Spatiotemporal maps of brain activity underlying word generation and their modification during repetition priming. *J. Neurosci.* 21, 3564–3571.
- Ding, L., 2009. Reconstructing cortical current density by exploring sparseness in the transform domain. *Phys. Med. Biol.* 54, 2683–2697.
- Ding, L., He, B., 2008. Sparse source imaging in electroencephalography with accurate field modeling. *Hum. Brain Mapp.* 29, 1053–1067.
- Ding, L., Yuan, H., 2013. Simultaneous EEG and MEG source reconstruction in sparse electromagnetic source imaging. *Hum. Brain Mapp.* 34, 775–795.
- Ding, L., Wilke, C., Xu, B., Xu, X., van Drongelen, W., Kohrman, M., He, B., 2007. EEG source imaging: correlate source locations and extents with ECoG and surgical resections in epilepsy patients. *J. Clin. Neurophysiol.* 24, 130–136.
- Ding, L., Ni, Y., Sweeney, J., He, B., 2011. Sparse cortical current density imaging in motor potentials induced by finger movement. *J. Neural Eng.* 8, 036008.
- Frye, R.E., Rezaie, R., Papanicolaou, A.C., 2009. Functional neuroimaging of language using magnetoencephalography. *Phys. Life Rev.* 6, 1–10.
- Gerloff, C., Uenishi, N., Nagamine, T., Kunieda, T., Hallett, M., Shibasaki, H., 1998. Cortical activation during fast repetitive finger movements in humans: steady-state movement-related magnetic fields and their cortical generators. *Electroencephalogr. Clin. Neurophysiol.* 109, 444–453.
- Grova, C., Daunizeau, J., Lina, J.M., Bénar, C.G., Benali, H., Gotman, J., 2006. Evaluation of EEG localization methods using realistic simulations of interictal spikes. *Neuroimage* 29, 734–753.
- Hämäläinen, M.S., Ilmoniemi, R.J., 1994. Interpreting measured magnetic fields of the brain: minimum norm estimates. *Med. Biol. Eng. Comput.* 32, 35–42.
- Hämäläinen, M.S., Hari, R., Ilmoniemi, R.J., Knuutila, J., Lounasmaa, O.V., 1993. Magnetoencephalography — theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev. Mod. Phys.* 65, 413–497.
- Haufe, S., Nikulin, V., Ziehe, A., Müller, K., Nolte, G., 2008. Combining sparsity and rotational invariance in EEG/MEG source reconstruction. *Neuroimage* 42, 726–738.
- He, B., Wang, Y., Wu, D., 1999. Estimating cortical potentials from scalp EEG's in a realistically shaped inhomogeneous head model by means of the boundary element method. *IEEE Trans. Biomed. Eng.* 46, 1264–1268.
- Hebiri, M., van de Geer, S., 2011. The smooth-lasso and other $\ell_1 + \ell_2$ -penalized methods. *Electron J. Stat.* 5, 1184–1226.
- Hillyard, S., 1993. Electrical and magnetic brain recordings: contributions to cognitive neuroscience. *Curr. Opin. Neurobiol.* 3, 217–224.
- Hoerl, A.E., Kennard, R.W., 1970. Ridge regression: biased estimation for nonorthogonal problems. *Technometrics* 12, 55–67.
- Jäncke, L., Loose, R., Lutz, K., Specht, K., Shah, N.J., 2000. Cortical activations during paced finger-tapping applying visual and auditory pacing stimuli. *Cogn. Brain Res.* 10, 51–66.
- Jerbi, K., Baillet, S., Mosher, J.C., Nolte, G., Garnero, L., Leahy, R.M., 2004. Localization of realistic cortical activity in MEG using current multipoles. *Neuroimage* 22, 779–793.
- Lai, Y., van Drongelen, W., Ding, L., Hecox, K.E., Towle, V.L., Frim, D.M., He, B., 2005. Estimation of *in vivo* human brain-to-skull conductivity ratio from simultaneous extra- and intra-cranial electrical potential recordings. *Clin. Neurophysiol.* 116, 456–465.
- Liao, K., Zhu, M., Ding, L., Valette, S., Zhang, W., Dickens, D., 2012. Sparse representation of cortical current density maps using wavelets. *Phys. Med. Biol.* 57, 6881–6901.
- Lobo, M.S., Vandenberghe, L., Boyd, S., Lebret, H., 1998. Applications of second-order cone programming. *Linear Algebra Appl.* 284, 193–228.
- Matsuura, K., Okabe, Y., 1995. Selective minimum-norm solution of the biomagnetic inverse problem. *IEEE Trans. Biomed. Eng.* 42, 608–615.
- McGill, R., Tukey, J.W., Larsen, W.A., 1978. Variations of box plots. *Am. Stat.* 32, 12–16.
- Molins, A., Stufflebeam, S.M., Brown, E.N., Hämäläinen, M.S., 2008. Quantification of the benefit from integrating MEG and EEG data in minimum ℓ_2 -norm estimation. *Neuroimage* 42, 1069–1077.
- Morozov, A.V., 1966. On the solution of functional equations by the method of regularization. *Sovrem. Math. Dokl.* 7, 414–417.
- Mosher, J.C., Leahy, R.M., Lewis, P.S., 1999. EEG and MEG: forward solutions for inverse methods. *IEEE Trans. Biomed. Eng.* 46, 245–259.
- Nesterov, Y., Nemirovski, A., 1994. Interior-point polynomial algorithms in convex programming.
- Nunez, P.L., 1981. *Electric Field of the Brain*. Oxford University Press, London.
- Oishi, M., Otsubo, H., Kameyama, S., Morota, N., Masuda, H., Kitayama, M., Tanaka, R., 2002. Epileptic spikes: magnetoencephalography versus simultaneous electrocorticography. *Epilepsia* 43, 1390–1395.
- Ou, W., Hämäläinen, M.S., Golland, P., 2009. A distributed spatio-temporal EEG/MEG inverse solver. *Neuroimage* 44, 932–946.
- Papanicolaou, A.C., Simos, P.G., Castillo, E.M., Breier, J.I., Sarkari, S., Patariaia, E., Billingsley, R.L., et al., 2004. Magnetoencephalography: a noninvasive alternative to the Wada procedure. *J. Neurosurg.* 100, 867–876.
- Pascual-Marqui, R.D., Michel, C.M., Lehmann, D., 1994. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *Int. J. Psychophysiol.* 18, 49–65.
- Petrov, Y., 2012. Harmony: EEG/MEG linear inverse source reconstruction in the anatomical basis of spherical harmonics. *PLoS One* 7, e44439.
- Plummer, C., Harvey, A., Cook, M., 2008. EEG source localization in focal epilepsy: where are we now? *Epilepsia* 49, 201–218.
- Pollok, B., Müller, K., Aschersleben, G., Schmitz, F., Schnitzler, A., Prinz, W., 2003. Cortical activations associated with auditorily paced finger tapping. *Neuroreport* 14, 247–250.
- Schröder, P., Sweldens, W., 1995. Spherical wavelets: efficiently representing functions on the sphere. *Conf. Proc. Comput. Graph. Interact. Technol.* 161–172.
- Sharon, D., Hämäläinen, M., Tootell, R., Halgren, E., Belliveau, J., 2007. The advantage of combining MEG and EEG: comparison to fMRI in focally-stimulated visual cortex. *Neuroimage* 36, 1225–1235.
- Stefan, H., Hummel, C., Scheler, G., Genow, A., Druschky, K., Titz, C., Kaltenhäuser, 2003. Magnetic brain source imaging of focal epileptic activity: a synopsis of 455 cases. *Brain* 126, 2396–2405.
- Sturm, J.F., 1999. Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones. *Optim. Methods Softw.* 11, 625–653.
- Tibshirani, R., 1996. Regression shrinkage and selection via the Lasso. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* 58, 267–288.
- Tibshirani, R., Saunders, M., Rosset, S., Zhu, J., Knight, K., 2005. Sparsity and smoothness via the fused lasso. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* 67, 91–108.
- Uutela, K., Hämäläinen, M.S., Somersalo, E., 1999. Visualization of magnetoencephalographic data using minimum current estimates. *Neuroimage* 35, 168.
- Valette, S., Prost, R., 2004. Wavelet-based progressive compression scheme for triangle meshes: wavemesh. *IEEE Trans. Vis. Comput. Graph.* 10, 123–129.
- Vega-Hernández, M., Martínez-Montes, E., Sanches-Bornot, J.M., Lage-Castellanos, A., Valdes-Sosa, P.A., 2008. Penalized least squares methods for solving the EEG inverse problem. *Stat. Sin.* 18, 1535–1551.
- Wood, C.C., 1982. Application of dipole localization methods to source identification of human evoked potentials. *Ann. N. Y. Acad. Sci.* 388, 139–155.
- Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* 67, 301–320.
- Zouridakis, G., Simos, P.G., Papanicolaou, A.C., 1998. Multiple bilaterally asymmetric cortical sources account for the auditory N1m component. *Brain Topogr.* 10, 183–189.