Spatio-Temporal Analysis in Functional Brain Imaging

by

Wanmei Ou

B.E., Electrical Engineering, the City College of New York, 2003

S.M., Electrical Engineering and Computer Science,
Massachusetts Institute of Technology, 2005

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in Electrical Engineering and Computer Science
at the Massachusetts Institute of Technology

February, 2010

Signature of Author: _____

Department of Electrical Engineering and Computer Science
January 15, 2010

Certified by: _____

Polina Golland, Associate Professor
of Electrical Engineering and Computer Science
Thesis Supervisor

Certified by: _____

Matti S. Hämäläinen, Associate Professor
of Radiology at Harvard Medical School
Thesis Supervisor

Accepted by: _____

Terry P. Orlando, Professor
of Electrical Engineering and Computer Science
Chair, Committee for Graduate Students

Spatio-Temporal Analysis in Functional Brain Imaging

by Wanmei Ou

To be submitted to the Department of
Electrical Engineering and Computer Science
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Electrical Engineering and Computer Science

Abstract

Localizing sources of activity from electroencephalography (EEG) and magnetoencephalography (MEG) measurements involves solving an ill-posed inverse problem, where infinitely many source distribution patterns can give rise to identical measurements. This thesis aims to improve the accuracy of source localization by incorporating spatio-temporal models into the reconstruction procedure.

First, we introduce a novel method for current source estimation, which we call the $\ell_1\ell_2$-norm source estimator. The underlying model captures the sparseness of the active areas in space while encouraging smooth temporal dynamics. We compute the current source estimates efficiently by solving a second-order cone programming problem. By considering all time points simultaneously, we achieve accurate and stable results as confirmed by the experiments using simulated and human MEG data.

Although the $\ell_1\ell_2$-norm estimator enables accurate source estimation, it still faces challenges when the current sources are close to each other in space. To alleviate problems caused by the limited spatial resolution of EEG/MEG measurements, we introduce a new method to incorporate information from functional magnetic resonance imaging (fMRI) into the estimation algorithm. Whereas EEG/MEG record neural activity, fMRI reflects hemodynamic activity in the brain with high spatial resolution. We examine empirically the neurovascular coupling in simultaneously recorded MEG and diffuse optical imaging (DOI) data, which also reflects hemodynamic activity and is compatible with MEG recordings. Our results suggest that the neural activity and hemodynamic responses are aligned in space. However, the relationship between the temporal dynamics of the two types of signals is non-linear and varies from region to

region.

Based on these findings, we develop the fMRI-informed regional EEG/MEG source estimator (FIRE). This method is based on a generative model that encourages similar spatial patterns but allows for differences in time courses across imaging modalities. Our experiments with both Monte Carlo simulation and human fMRI-EEG/MEG data demonstrate that FIRE significantly reduces ambiguities in source localization and accurately captures the timing of activation in adjacent functional regions.

Thesis Supervisors: Polina Golland, Associate Professor
of Electrical Engineering and Computer Science at MIT

Matti S. Hämäläinen, Associate Professor
of Radiology at Harvard Medical School

Acknowledgments

Upon entering graduate school I was told repeatedly that finding an advisor is like finding a date: the best match is based on personality. Looking back over the past few years, I really enjoyed the journey accompanied by my two "dates", my thesis advisors Polina Golland and Matti Hämäläinen. Polina not only enabled me to tackle difficult research problems, but more importantly she accentuated my presentation skills, open-mindedness and optimism. Matti has always been available to help out with technical details including the 3 a.m. call before my thesis defense. From him, I also obtained an organized style: from file-naming to project organization.

I would like to thank my thesis committee George Verghese and Sandy Wells for their constructive feedback. I still remember George's personal greeting during my MIT open-house visit; his personable hospitality is one of the main factors in choosing MIT for my graduate study. Sandy is also a long-term collaborator. Through his connections, I gained tremendous resources from hospitals coast to coast.

This thesis would not be possible without the intellectual contributions from Aapo Nummenmaa, Mike Siracusa, Bruce Fischl, Thomas Yeo, Jyrki Ahveninen, Maria Angela Franceschini, and Tommi Raij. I would like to thank Aapo for playing the devil's advocate. Without his advice our complex algorithms would not have been easily perceived by the reviewers. Mike taught me algorithms in machine learning and information theory; Bruce and Thomas helped me to better understand brain anatomy. This thesis required a lot of evaluations with experimental subject data. I would like to thank Jyrki, Mari, and Tommi in providing the necessary data as well as sharing with me their deep understanding of neuroscience.

I would like to thank many people in the medical vision group: Archana Venkataraman, Bjoern Menze, Danial Lashkari, Georg Langs, Mert Sabuncu, Michal Depa, Ramesh Sridharan, Tammy Riklin-Raviv for simulated discussions in group meetings and for proof reading paper drafts. I would also like to thank vision group alumni Eric Cosman, Anna Custo, Gerald Dalley, Lauren O'Donnell, Kilian Pohl, Kinh Tieu, Kevin Wilson, and Lilla Zollei for teaching newbies like me how to steer around in gradu-

ate school. Gerald is especially kind for sharing many in-depth discussions on career planning.

Pursuing a Ph.D. degree would never have crossed my mind had I not known Harold Falk, Catherine Falk and Thao Nguyen from my undergraduate study at the City College of New York and part-time research work at the New York Blood Center. Their encouragement and guidance has led me to expand my career development opportunities.

My stay at MIT would not have been as interesting without a group of friends. I feel grateful for the continuing friendship with Ying Chen, Xuemin Chi, Aiyan Lu, and Fang Zhao. I would like to specially thank my friend Biz Bose for his honest and objective advice to various difficult situations that I faced and for teaching me a lot of fun activities from tennis to skiing.

Finally, I would like to thank my parents for their support over the years. I would like to specially thank my boyfriend Jeff Hughes for his love and unrelenting support.

Contents

List of Figures

List of Tables

Notation

Symbol	Definition
General Notation	
$\mathbf{E}$	Electric field
$\mathbf{B}$	Magnetic field
$\mathbf{J}$	Current density
$\mathbf{j}_n$	Current density at location n
V	Electric potential
$\mathbf{Y}$	EEG/MEG measurements
$\mathbf{y}(t)$	EEG/MEG measurements at time t
$\mathbf{A}$	EEG/MEG lead-field matrix
$\mathbf{F}$	fMRI measurements
$\mathbf{f}_n$	fMRI signal at location n
ϵ	Permittivity
μ	Magnetic permeability
σ	Electrical conductivity
N	Number of locations (i.e., vertices or voxels) in the brain
M	Number of EEG electrodes and/or MEG sensors
$\|\mathbf{x}\|_{\ell_2}$	ℓ_2-norm of vector $\mathbf{x}$
$\|\mathbf{x}\|_{\ell_1}$	ℓ_1-norm of vector $\mathbf{x}$
$\|\mathbf{X}\|_F$	Frobenius norm of matrix $\mathbf{X}$
FIRE	
z_n	Latent activation amplitude at location n
$\mathbf{u}_k$	Neural waveform in brain region k
$\mathbf{v}_k$	Vascular waveform in brain region k

Chapter 1

Introduction

[Ahlfors & Simpson, 2004] Electroencephalography (EEG) and magnetoencephalography (MEG) are widely used for functional brain mapping. With appropriate source estimation algorithms one can locate the activated regions and estimate their dynamics. The non-invasive nature of EEG and MEG makes these methods particularly suitable for neuroscience research [Chafetz & Cadilhac, 1954, Hari & Forss, 1999, Engel *et al.*, 2001, Jaaskelainen *et al.*, 2004, Ahveninen *et al.*, 2006] and clinical studies, such as surgical planning for epilepsy patients [Kershman *et al.*, 1951, Knake *et al.*, 2006]. There are two main types of EEG and/or MEG (EEG/MEG) studies: signal analysis in the sensor space [Lachaux *et al.*, 1999, Engel *et al.*, 2001, Grill-Spector *et al.*, 2006] and spatio-temporal analysis in the source space [Hari & Forss, 1999, Lin *et al.*, 2004, Jaaskelainen *et al.*, 2004, Ahveninen *et al.*, 2006]. In this thesis, we focus on the spatio-temporal modeling in the source space, which entails mapping of the data from the sensor space to the source space. This procedure is referred to as source localization and requires an inverse solver. Fig. 1.1 illustrates the relationship between the EEG/MEG measurements and the underlying neural activity; the inverse problem is the inference of neural activity from the measurements.

■ 1.1 Challenges in Current Source Estimation

The principal difficulty in EEG/MEG source localization stems from the ill-posed nature of the electromagnetic inverse problem: infinitely many source distributions can give rise to identical measurements [Hadamard, 1902, Hämäläinen *et al.*, 1993]. Additional assumptions, such as minimum energy [Hämäläinen & Ilmoniemi,1984, Wang *et al.*, 1993] or minimum current [Uutela *et al.*, 1999], must be incorporated into the estimation procedure to obtain a unique estimate [Baillet *et al.*, 2001]. The corresponding estimation

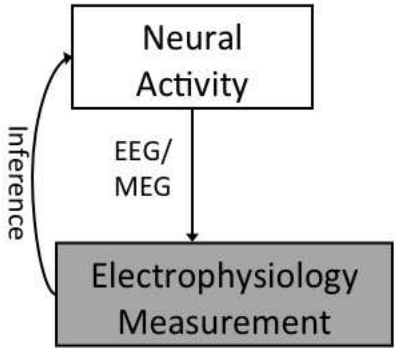


Figure 1.1. Relationship between neural activity and EEG/MEG measurements.

methods belong to the class of algorithms that model the sources as a spatial distribution, in contrast to the dipole fitting approach where the EEG/MEG data is explained by a small number of current dipole sources [Wood, 1982, Scherg & Von Cramon, 1985, Mosher *et al.*, 1992]. In this thesis, we focus on the distributed approach.

The widely used minimum norm estimate (MNE) [Hämäläinen & Ilmoniemi,1984, Dale & Sereno, 1993, Wang *et al.*, 1993] recovers the source distribution with minimum overall energy (or minimum ℓ_2-norm) that induces signals at the sensors consistent with the measurements. Although the ℓ_2-norm method leads to an efficient linear estimator, the MNE solutions are often too diffuse. Hence, MNE is not appropriate for localization of focal epilepsy and early sensory activations, which have been shown to be highly localized in intra-cranial experiments [Barth *et al.*, 1982, Allison *et al.*, 1989]. To overcome this disadvantage, the minimum current estimate (MCE) utilizes the sparse property of the ℓ_1-norm to achieve more focal solutions[Uutela *et al.*, 1999].

■ 1.2 Sparse Source Estimation

Similar to MNE, MCE is applied to each time point in the data separately. One of the drawbacks of MCE is its sensitivity to noise: the estimated activations tend 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. Hence, in Chapter 3 of this thesis, we propose a novel solver that takes advantage of the smooth properties of the underlying source signal and estimates the current source distribution for all time points simultaneously. The corresponding

solver is called the $\ell_1\ell_2$-norm estimator.

The $\ell_1\ell_2$-norm estimator imposes ℓ_1-norm regularization in the spatial domain and ℓ_2-norm regularization in the temporal domain. Similar to MCE, the $\ell_1\ell_2$-norm estimator uses ℓ_1-norm regularizer to encourage spatial sparsity. For the temporal domain, we assume that the source signals are linear combinations of *multiple* temporal basis functions, and estimate 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. The resulting model can be formulated as a second-order cone programming (SOCP) problem and solved efficiently using the interior-point method [Alizadeh & Goldfarb, 2001]. Experimental comparisons in Section 3.3 reveal that this joint spatio-temporal model implicitly increases the signal-to-noise ratio (SNR) and achieves more accurate estimates than the traditional methods.

■ 1.3 Neurovascular Coupling

Due to the limited spatial resolution of EEG/MEG measurements, relying solely on assumptions about the spatial patterns of activation, such as ℓ_1- and ℓ_2-norm, does not allow us to distinguish activation sources that are close in space. Specific prior knowledge about activated locations can be obtained from other imaging modalities. Among these, functional magnetic resonance imaging (fMRI) provides the most relevant information due to its good spatial resolution. In addition, its non-invasive nature makes this imaging technique suitable for a wide range of applications.

The main concern in integrating fMRI information into the estimation of the neural activity is that fMRI measures hemodynamic activity, which only indirectly reflects the neural activity measured by EEG/MEG. The flow chart in Fig. 1.2 illustrates the source activities for EEG/MEG and fMRI measurements, as well as the relationship between these two types of activities. The time courses of the neural and the vascular activities differ substantially, and their exact spatio-temporal relationship is yet to be fully characterized [Logothetis & Wandell, 2004, Sharon *et al.*, 2007]. In addition to the differences in their physiological origins, EEG/MEG and fMRI have different sensitivity characteristics. For example, a brief transient neural activity may be difficult to detect in fMRI, but it is often easily observable by EEG/MEG. On the other hand, while a sustained weak neural activity may lead to a relatively strong fMRI signal, it might have a poor SNR in EEG/MEG. Hence, in Chapter 4 we examine the neurovascular

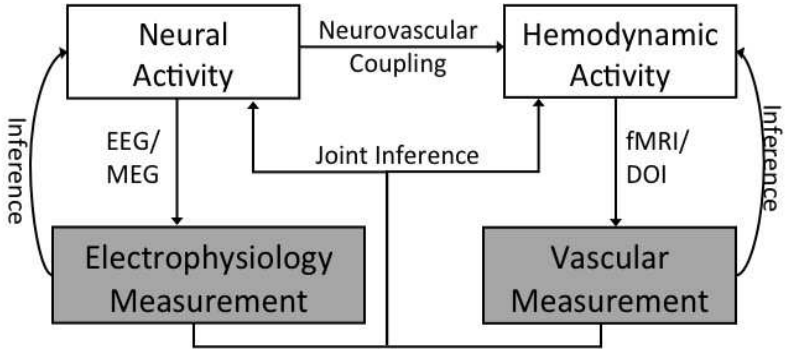


Figure 1.2. Source activities for EEG/MEG and fMRI measurements, as well as the relationship between neural and hemodynamic activities.

coupling relation.

In this work, we develop a simultaneous MEG-diffuse optical image (DOI) recording technique to examine the neurovascular relationship using somatosensory stimulus trains with varying train duration. Similar to fMRI, DOI measures hemodynamic activity. Despite its lower spatial resolution, DOI provides a higher temporal resolution and good SNR for weak stimuli, such as the median-nerve stimulation, compared to fMRI. Furthermore, MEG and DOI are compatible with each other with a few technical precautions. These two modalities allow for simultaneous measurement of neural and vascular responses.

We choose median-nerve electrical stimulation because the corresponding somatosensory responses have been extensively studied in humans with EEG and MEG [Brenner *et al.*, 1978, Goff *et al.*, 1978, Kaufman *et al.*, 1981, Hari *et al.*, 1984, Tiihonen *et al.*, 1989, Allison, 1992, Mauguiere *et al.*, 1997, Hari & Forss, 1999]. The active regions include the contralateral primary sensory cortex (cSI), the bilateral secondary sensory cortex (SII), and other associated areas such as the posterior parietal cortex (PPC). With DOI we recover both oxy- and deoxy-hemoglobin (HbO and HbR) evoked responses in cSI for different stimulus train durations. From MEG data, we estimate the sites of neural activity and extract the cSI current dipole time course. By co-registering the two measurements with respect to each subject's structural magnetic resonance imaging (MRI), we verify that there is spatial co-localization between neural and hemodynamic activities. Our results suggest that neural activity and hemodynamic responses are generally aligned in space.

The MEG response to median nerve stimulation typically contains several peaks. They are induced by excitatory post-synaptic potentials (EPSPs) or inhibitory post-synaptic potentials (IPSPs) [Wikström *et al.*, 1996]. We evaluate which peak best predicts the hemodynamic responses as measured by DOI. By limiting the consideration to specific peaks of the neural responses, we aim to isolate the part of neuronal activity that causes the local excitatory hemodynamic response. Our experimental results show that the habituation effect in the hemodynamic response is stronger than that in the earliest cortical neural responses. Furthermore, using a linear convolution model to predict hemodynamic responses, we find that including late neural components (≥ 30 msec) improves the prediction. This finding suggests that in addition to the initial evoked-response deflections related to the thalamic afferent input, later cortical activity needs to be taken into account to predict the hemodynamic response.

■ 1.4 Joint Estimation

Based on the findings reported in Chapter 4, we propose a novel method, the fMRI-informed regional estimation (FIRE), to improve the accuracy of the EEG/MEG source estimate, which is referred to as "joint inference" in Fig. 1.2. We present the method in Chapter 5. Fig. 1.3 illustrates the model assumptions of FIRE. The regions indicated by different colors are chosen based on the subject-specific cortical parcellation [Fischl *et al.*, 2002]. Our neurovascular coupling study implies that neural and hemodynamic activities are aligned in space and that the relationship between the dynamics of these two types of activities is largely unknown. Therefore, we only model the similarity of spatial patterns in the two processes in FIRE. Furthermore, we expect the shape of the activation time courses to vary across brain regions, especially for the neural activity. To account for this fact, FIRE treats the temporal dynamics in different brain regions independently. In other words, there is no constraint imposing similarity of the activation time courses across regions. We assume the shape of the activation time courses to be constant within a brain region, modulated by a set of location-specific latent variables. Handling the temporal dynamics of the two types of activities separately while exploiting their common spatial pattern preserves the temporal resolution of EEG/MEG and helps to achieve accurate source localization.

To fit the model to the data, we use the coordinate descent method, alternating between the estimation of current sources and that of other model parameters. This itera-

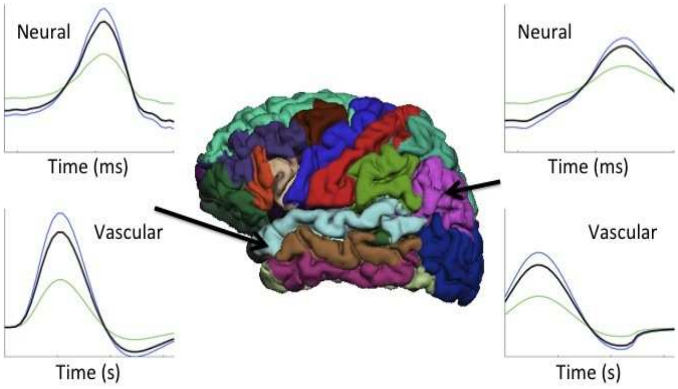


Figure 1.3. Graphical illustration of the model assumptions in FIRE. The anatomical regions of the left hemisphere are depicted in the middle of the figure. Region-specific neural waveforms (top two panels) and the vascular waveforms (bottom two panels) for two separate regions are shown in black. Location-specific current time courses and fMRI time courses for two locations in each of the two highlighted regions are shown in blue and green. The current time courses and the fMRI time courses are scaled versions of the corresponding region-specific waveforms.

tive update scheme is similar to the re-weighted MNE methods, such as the FOCal Underdetermined System Solver (FOCUSS) [Gorodnitsky & Rao, 1997]. In contrast to the re-weighted MNE, our method determines the weights jointly from both the estimated neural activity and the vascular activity measured by fMRI. Moreover, the estimates at different time points influence each other. The computation of the weights is related to problems arising in continuous Gaussian mixture modeling, which can be efficiently optimized using the Expectation-Maximization (EM) algorithm [Dempster *et al.*, 1977].

We employ a Monte Carlo evaluation procedure to compare FIRE to several other joint EEG/MEG-fMRI algorithms. Our results show that FIRE provides the best trade-off in estimation accuracy between the spatial and the temporal accuracy. Analysis using human EEG/MEG-fMRI data reveals that FIRE significantly reduces the ambiguities in source localization present in the MNE estimates, and that it accurately captures activation timing in adjacent functional regions.

■ 1.5 Contributions of This Thesis

To summarize, this thesis introduces three main contributions to improve functional brain mapping accuracy:

1. Improvement in robustness of MCE through a temporal model that encourages

smooth underlying source signals – $\ell_1\ell_2$-norm estimator.

2. An empirical study of neurovascular coupling relationship. Neural activity and hemodynamic response are generally aligned in space, but their temporal relationship is non-linear and varies across brain regions.

3. A new method for source estimation, FIRE, that combines spatial and temporal advantages in EEG/MEG and fMRI data to distinguish sources close in space.

In the following, we first review background material, including various functional imaging measurements and their corresponding analysis methods, as well as the neurovascular coupling relationship, in Chapter 2. We then introduce the $\ell_1\ell_2$-norm estimator to improve robustness of inverse solutions in Chapter 3. In Chapter 4, we present our empirical findings on the neurovascular coupling. Based on these findings, we develop the FIRE estimator in Chapter 5. We summarize and propose future research directions in Chapter 6.

Chapter 2

Background

In contrast to anatomical images, which provide a static view of the brain, functional brain imaging records the temporal dynamics of brain activity. In this section, we discuss various functional brain imaging techniques, followed by a literature review on data analysis methods for localizing activation areas in the brain. We then present recent findings on neurovascular coupling and review joint analysis methods that integrate functional data across multiple imaging modalities.

■ 2.1 Functional Brain Imaging Methods

Each functional imaging method measures either the neural activity directly or the associated changes such as hemodynamics and metabolism. While electroencephalography (EEG), intra-cranial EEG, and magnetoencephalography (MEG) measures neural activity, functional magnetic resonance imaging (fMRI) and near infrared spectroscopy (NIRS)/diffuse optical imaging (DOI) reflect hemodynamic activity. Positron emission tomography (PET) measures metabolism of neurons.

These functional imaging methods can be also categorized as invasive or non-invasive. Intra-cranial EEG requires an open-skull surgery where the electrodes directly measure voltage fluctuation at a specific location on the cortex. Therefore, only few locations can be tested during the surgery. For PET, a small dose of radioactive tracer isotope is needed to measure changes of glucose metabolism in cells. In contrast, EEG, MEG, fMRI and DOI are completely non-invasive. Consequently, they are widely used in both neuroscience research and clinical practice, such as localization of functional differences in schizophrenic and Alzheimer's patients [Mayberg *et al.*, 2000, Maestú *et al.*, 2001] and surgical planning in epilepsy patients [Knake *et al.*, 2006]. They all measure activation in the whole brain, in contrast to specific locations probed by

inter-cranial methods. However, this characteristic also makes source localization difficult, especially for EEG, MEG and DOI. Here, we mainly focus on the source localization problem for EEG and MEG.

■ 2.2 EEG/MEG Measurements

For a linear isotropic material, Faraday's Law and Ampere's Law in Maxwell's equations describe the relationship among magnetic field $\mathbf{B}$, current density $\mathbf{J}$, and electric field intensity $\mathbf{E}$ as:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2.1}$$

$$\nabla \times \mathbf{B} = \mu \left(\mathbf{J} + \epsilon \frac{\partial \mathbf{E}}{\partial t} \right), \tag{2.2}$$

where μ and ϵ are the permeability and permittivity of a material, respectively.

The brain is a passive non-magnetic medium. In the typical frequency range of neural activity of 1 kHz, the electric and magnetic fields of the brain can be accounted for by the quasi-static approximation of Maxwell's equations – that is, magnetic induction and capacitive effect are negligible [Plonsey 1969, Nunez, 1981]. In other words, $\frac{\partial \mathbf{B}}{\partial t}$ and $\frac{\partial \mathbf{E}}{\partial t}$ are negligible in Eq. (2.1)-(2.2). For the quasi-static approximation to be valid, we need to satisfy two conditions: (a) the time-derivative of the electric field must be much smaller than the ohmic current: $|\epsilon \partial \mathbf{E}/\partial t| << |\sigma \mathbf{E}|$ and (b) $\partial \mathbf{B}/\partial t$ must be small. Parameter σ denotes the electrical conductivity of the medium.

To show the validity of condition (a), we let

$$\mathbf{E} = \mathbf{E}_0 e^{i2\pi f t}, \tag{2.3}$$

where $i = \sqrt{-1}$ and f is the neural activity frequency. Given that brain tissue conductivity is $\sigma = 0.3\,\Omega^{-1}\mathrm{m}^{-1}$, permittivity $\epsilon = 10^5 \epsilon_0$, and $f = 100\,\mathrm{Hz}$,

$$|\epsilon \partial \mathbf{E}/\partial t| = 2\pi f \epsilon |\mathbf{E}_0| = 5.6 \times 10^{-4} |\mathbf{E}_0| \ll |\sigma \mathbf{E}| = 0.3 |\mathbf{E}_0|. \tag{2.4}$$

Parameter ϵ_0 denotes the permittivity of free space. Furthermore, condition (b) is true since the wavelength of the electromagnetic propagation is

$$|2\pi f \mu_0 \sigma (1 + i 2\pi f \epsilon/\sigma)|^{-1/2} \approx 65\,\mathrm{m}, \tag{2.5}$$

much longer than the diameter of the head. We can derive the wavelength formula by solving differential equation $\nabla \times \nabla \times \mathbf{E} = -\mu_0 \frac{\partial}{\partial t} \left(\sigma \mathbf{E} + \epsilon \frac{\partial \mathbf{E}}{\partial t} \right)$, which is obtained from Eq. (2.1) and Eq. (2.2) with $\mu = \mu_0$.

Therefore, under the quasi-static approximation, $\nabla \times \mathbf{E} = 0$. That means the electric field can be expressed as the gradient of a scalar potential:

$$\mathbf{E} = -\nabla V. \tag{2.6}$$

Moreover, with this approximation, the forward problem for magnetic field outside of the head $\mathbf{B}(\mathbf{r})$ due to currents in the brain $\mathbf{J}(\mathbf{r}')$ can be expressed as

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int \frac{\mathbf{J}(\mathbf{r}') \times \mathbf{d}}{|\mathbf{d}|^3} dv', \tag{2.7}$$

where $\mathbf{r}$ is a point where the field is computed, and $\mathbf{r}'$ is a point inside the brain volume v'. We define vector $\mathbf{d} = \mathbf{r} - \mathbf{r}'$ to simplify the notation. Substituting identities $\mathbf{d}/|\mathbf{d}|^3 = -\nabla(1/|\mathbf{d}|) = \nabla'(1/|\mathbf{d}|)$ and $\mathbf{J} \times \nabla'(1/|\mathbf{d}|) = (\nabla' \times \mathbf{J})/|\mathbf{d}| - \nabla' \times (\mathbf{J}/|\mathbf{d}|)$ into Eq. (2.7), we get

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \left[\int \frac{\nabla' \times \mathbf{J}(\mathbf{r}')}{|\mathbf{d}|} dv' - \int \nabla' \times (\mathbf{J}/|\mathbf{d}|) dv' \right]. \tag{2.8}$$

In the above, ∇ indicates the curl with respect to the observation coordinates, $\nabla = \mathbf{e}_x \partial/\partial x + \mathbf{e}_y \partial/\partial y + \mathbf{e}_z \partial/\partial z$; ∇' indicates the curl with respect to the source coordinates $\nabla' = \mathbf{e}_x \partial/\partial x' + \mathbf{e}_y \partial/\partial y' + \mathbf{e}_z \partial/\partial z'$.

If we transform the second integral of Eq. (2.8) into a surface integral, we can see that the current density approaches zero sufficiently fast when $\mathbf{r}'$ goes to infinity. Therefore,

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int \frac{\nabla' \times \mathbf{J}(\mathbf{r}')}{|\mathbf{d}|} dv'. \tag{2.9}$$

The current density $\mathbf{J}(\mathbf{r}')$ can be divided into the primary current $\mathbf{J}^p(\mathbf{r}')$ and the volume current $\mathbf{J}^v(\mathbf{r}')$. Neural activity gives rise to primary current mainly inside or in the vicinity of a cell, whereas the volume current flows passively everywhere in the medium [Hämäläinen *et al.*, 1993]:

$$\mathbf{J}(\mathbf{r}') = \mathbf{J}^p(\mathbf{r}') + \mathbf{J}^v(\mathbf{r}') = \mathbf{J}^p(\mathbf{r}') + \sigma \mathbf{E}(\mathbf{r}') = \mathbf{J}^p(\mathbf{r}') - \sigma \nabla' V(\mathbf{r}'). \tag{2.10}$$

Using the identity $\nabla' \times (\sigma \nabla' V) = \nabla' \sigma \times \nabla' V$, we substitute Eq. (2.10) into Eq. (2.9):

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \left[\int \frac{\nabla' \times \mathbf{J}^p(\mathbf{r}')}{|\mathbf{d}|} dv' - \int \frac{\nabla' \sigma \times \nabla' V}{|\mathbf{d}|} dv' \right]. \tag{2.11}$$

The first term is the magnetic field induced by the primary current $\mathbf{J}^p$, and the second term is induced by the volume current $\mathbf{J}^v$. If we are in an infinite homogeneous conductor, i.e., $\nabla' \sigma = 0$, the magnetic field is solely contributed by $\mathbf{J}^p$. However, generally,

both $\mathbf{J}^p$ and $\mathbf{J}^v$ determine the magnetic field. If we take the divergence of Eq. (2.10) and re-arrange the terms, we get

$$\nabla' \cdot \mathbf{J}^p(\mathbf{r}') = \nabla' \cdot \mathbf{J}(\mathbf{r}') + \nabla' \cdot (\sigma(\mathbf{r}')\nabla' V(\mathbf{r}')). \tag{2.12}$$

Using the quasi-static approximation and the fact that the divergence of a curl is zero, i.e., $\nabla' \cdot (\nabla' \times \mathbf{B}) = 0$ in Eq. (2.2), we get $\nabla' \cdot \mathbf{J}(\mathbf{r}') = 0$. Therefore, Eq. (2.12) can be reduced to

$$\nabla' \cdot \mathbf{J}^p(\mathbf{r}') = \nabla' \cdot (\sigma(\mathbf{r}')\nabla' V(\mathbf{r}')). \tag{2.13}$$

With proper boundary conditions, we can solve for Eq. (2.13) numerically. We then substitute the solution into Eq. (2.11) to compute $\mathbf{B}(\mathbf{r})$.

To simplify notation, we omit the superscript ^{p} in the rest of the thesis. From the above we see that there is a simple linear relationship between the electric/magnetic recordings and the strengths of the currents, allowing us to express the forward solution in a simple matrix form:

$$\mathbf{y}(t) = \mathbf{A}\mathbf{J}(t) + \mathbf{e}(t), \tag{2.14}$$

where $\mathbf{y}(t)$ is the vector of instantaneous electric and/or magnetic recordings at time t, $\mathbf{A}$ is the so-called gain matrix (with each column specifying the electric and/or magnetic forward model for a given dipole component). Throughout this thesis, the forward matrix $\mathbf{A}$ is of size $M \times N$, where M and N denote the number of electrodes/sensors and the number locations (i.e., vertices or voxels) in the brain, respectively. $\mathbf{J}$ is a vector of current dipole component strengths, and $\mathbf{e}$ is a vector specifying the noise at each electrode/sensor.

Historically, the forward matrix $\mathbf{A}$ has been calculated assuming an idealized head shape with multiple concentric spheres of different conductivities [Cuffin & Cohen, 1977]. With the advancements in high-resolution 3-D magnetic resonance imaging (MRI) scans and numerical techniques, we can now customize the forward models to the individual subject's anatomy via either finite element method (FEM) or boundary element method (BEM) [Hämäläinen & Sarvas, 1989, Oostendorp & Van Oosterom, 1989, Dale & Sereno, 1993].

■ 2.3 EEG/MEG Source Estimation

Localizing active 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 electrically silent, or magnetically silent, or both. Moreover, solutions might not depend continuously on the data without regularization, which means small errors in measurements might cause errors of arbitrary magnitude in the estimated sources. These two characteristics make the inverse problem ill-posed in the sense of Hadamard [Hadamard, 1902].

■ 2.3.1 Dipole Fitting and Distributed Source Estimates

Inverse solvers for EEG/MEG source estimation can be classified into two broad categories: discrete parametric solvers, also known as dipole fitting, and distributed inverse solvers. The 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 for the number of current dipoles. The desired solution is computed by introducing 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 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 of sources (see, e.g. [Wood, 1982, Hari & Forss, 1999]). Although it is possible to obtain an initial guess for the number of dipoles based on a singular-value decomposition (SVD) of the recordings [Huang *et al.*, 1998], this method is sensitive to 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 extent [Jerbi *et al.*, 2004].

■ 2.3.2 Distributed Estimations

Not restricted to a fixed number of dipoles, the distributed solvers estimate the source amplitudes of all possible source locations. In this subsection, we first present the basics of point-wise estimators that estimate the source current distribution at each time point separately, and discuss their limitations. We then summarize the recent development

in spatio-temporal estimators.

Point-Wise Solvers 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. It is computed for each time point separately. For time t, the solution $\mathbf{J}^{\text{MNE}}$ is expressed as

$$\begin{aligned} \mathbf{J}^{\text{MNE}} &= \arg\min_{\mathbf{J}} \|\mathbf{J}\|_2^2 \\ \text{s.t.} \quad \mathbf{y}(t) &= \mathbf{A}\mathbf{J}, \end{aligned} \tag{2.15}$$

where $\|\mathbf{J}\|_2^2$ is the square of the ℓ_2-norm of vector $\mathbf{J}$.

We use the method of Lagrange multipliers to solve Eq. (2.15):

$$\mathbf{J}^{\text{MNE}} = \arg\min_{\mathbf{J}} \|\mathbf{y}(t) - \mathbf{A}\mathbf{J}\|_2^2 + \lambda^2 \|\mathbf{J}\|_2^2, \tag{2.16}$$

where λ defines the relative importance of the data term $\|\mathbf{y}(t) - \mathbf{A}\mathbf{J}\|_2^2$ and the regularization term $\|\mathbf{J}\|_2^2$. Since the data term depends on the amplitude of noise in the measurements, λ^2 can be viewed as the signal-to-noise ratio. Taking the derivative of Eq. (2.16) with respect to $\mathbf{J}$ and setting it to zero, we get

$$\mathbf{J}^{\text{MNE}} = \left(\mathbf{A}^{\text{T}}\mathbf{A} + \lambda^2 \mathbf{I}_N\right)^{-1} \mathbf{A}^{\text{T}} \mathbf{y}(t) \tag{2.17}$$

$$= \mathbf{A}^{\text{T}} \left(\mathbf{A}\mathbf{A}^{\text{T}} + \lambda^2 \mathbf{I}_M\right)^{-1} \mathbf{y}(t) \tag{2.18}$$

based on the following set of identities:

$$\begin{aligned} \left(\mathbf{A}^{\text{T}}\mathbf{A} + \lambda^2 \mathbf{I}_N\right)^{-1} \mathbf{A}^{\text{T}} &= \left(\mathbf{A}^{\text{T}}\mathbf{A} + \lambda^2 \mathbf{I}_N\right)^{-1} \mathbf{A}^{\text{T}} \left(\mathbf{A}\mathbf{A}^{\text{T}} + \lambda^2 \mathbf{I}_M\right) \left(\mathbf{A}\mathbf{A}^{\text{T}} + \lambda^2 \mathbf{I}_M\right)^{-1} \\ &= \left(\mathbf{A}^{\text{T}}\mathbf{A} + \lambda^2 \mathbf{I}_N\right)^{-1} \left(\mathbf{A}^{\text{T}}\mathbf{A}\mathbf{A}^{\text{T}} + \lambda^2 \mathbf{A}^{\text{T}}\mathbf{I}_M\right) \left(\mathbf{A}\mathbf{A}^{\text{T}} + \lambda^2 \mathbf{I}_M\right)^{-1} \\ &= \left(\mathbf{A}^{\text{T}}\mathbf{A} + \lambda^2 \mathbf{I}_N\right)^{-1} \left(\mathbf{A}^{\text{T}}\mathbf{A} + \lambda^2 \mathbf{I}_N\right) \mathbf{A}^{\text{T}} \left(\mathbf{A}\mathbf{A}^{\text{T}} + \lambda^2 \mathbf{I}_M\right)^{-1} \\ &= \mathbf{A}^{\text{T}} \left(\mathbf{A}\mathbf{A}^{\text{T}} + \lambda^2 \mathbf{I}_M\right)^{-1}, \end{aligned}$$

where $\mathbf{I}_M$ and $\mathbf{I}_N$ are the $M \times M$ and $N \times N$ identity matrices, respectively. The above derivations assume that the data has been whitened in the preprocessing step; the noise covariance matrix is therefore an identity matrix. A more general derivation replaces the white noise correlation matrix $\mathbf{I}_M$ with the noise covariance matrix $\mathbf{C}$. Due to the size of the forward matrix $\mathbf{A}$, Eq. (2.18) is preferred since it leads to more economical computation. Operator $\mathbf{A}^{\text{T}} \left(\mathbf{A}\mathbf{A}^{\text{T}} + \lambda^2 \mathbf{I}_M\right)^{-1}$ is usually referred to as the MNE inverse

operator. It is linear and independent of the data. In practice, this operator is precomputed and applied to data at each time point separately.

Although the ℓ_2-norm estimation leads to an efficient linear inverse operator, the MNE solutions are often too diffuse in space. In particular, MNE is not appropriate for localization of early sensory activations and for focal epilepsy, which have been shown to be focal in intra-cranial experiments [Barth *et al.*, 1982, Allison *et al.*, 1989].

To overcome this problem, 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 $p \leq 1$ in previous work [Rao & Kreutz-Delgado, 1999]. Other regularizers based on a norm penalty can provide bias towards sparsity. Among them, the minimum current estimate (MCE) is the most popular [Uutela *et al.*, 1999]. In contrast to the minimum energy assumption in MNE, MCE assumes that the optimal current estimate is the one that produces the minimum current or minimum ℓ_1-norm:

$$\begin{aligned} \mathbf{J}^{\mathrm{MCE}} &= \arg\min_{\mathbf{J}} \|\mathbf{J}\|_1 \\ \text{s.t.} \quad \mathbf{y}(t) &= \mathbf{A}\mathbf{J}, \end{aligned} \tag{2.19}$$

where $\|\mathbf{J}\|_1$ is the sum of absolute value of the elements in vector $\mathbf{J}$. Although the objective function in Eq. (2.19) is convex, it is not differentiable at zero. Hence, a gradient-based method is not directly applicable. One can solve the above optimization problem using the simplex method and the interior point method [Bertsimas & Tsitsiklis, 2008].

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 estimated time courses at a particular location exhibit "spiky" discontinuities when viewed over time. To avoid this problem one commonly averages time courses across adjacent sites, at the expense of spatial resolution. In Chapter 3, we introduce a method that addresses this issue by incorporating a model for temporal properties of the signals into MCE.

Spatio-Temporal Solvers Two alternative approaches utilize temporal constraints to improve source estimation 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.*, 2005a, Lamus *et al.*, 2007], a regularizer is explicitly incorporated into the cost function to model the smoothness of the current source distributions across 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.*, 2005a, Lamus *et al.*, 2007] propose a state-space model with smooth transitions between states that represent the current source distributions at consecutive time points. 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 these methods address the problem of sensitivity to noise, their implementation requires a substantial amount of computation, except for a limited number of low-pass filters; see [Zhang *et al.*, 2005a] for a comprehensive comparison of 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 to the signal subspace defined by a set of temporal basis functions [Huang *et al.*, 2006]. In other words, it separates the spatial and the temporal models into a two-step estimation procedure. Although VESTAL significantly reduces the noisy sensitivity in MCE, its estimates have limited accuracy due to errors propagated from the first estimation step to the second one. 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.

■ 2.4 Hemodynamics

It is known that neural activity causes an increase in blood flow to the activated regions to provide extra oxygen and glucose for the active brain cells. As large amounts of freshly oxygenated blood pours into an activation region, it causes reduction of the deoxygenated hemoglobin (HbR) to oxygenated hemoglobin (HbO) ratio [Jezzard *et al.*, 2002].

To measure the hemodynamics, fMRI [Ogawa *et al.*, 1990, Kwong *et al.*, 1992] and DOI [Arridge *et al.*, 1999, Boas *et al.*, 2001] detect the changes in the paramagnetic and optical properties of blood, respectively, as the ratio of HbR-to-HbO changes. Compared to neural activity, the change of HbR-to-HbO ratio is a much slower process, and temporal sampling rate of every second is usually sufficient to capture this change.

■ 2.4.1 fMRI

fMRI signals are acquired using a standard MRI system with suitable pulse sequences, such as echo planar imaging (EPI) [Jezzard *et al.*, 2002]. fMRI signals reflect differences in the paramagnetic properties of oxygenated and de-oxygenated blood as a consequence of activation. Iron atoms are bound to oxygen in HbO, but not in HbR. Since iron atoms cause distortions in the magnetic field, but oxygenated iron atoms do not, HbO and HbR have different paramagnetic properties. A change in the HbO-to-HbR ratio causes a small change in the effective magnetic field in the active region, and thus a small variation in the fMRI signal. This type of image is called Blood-Oxygen-Level-Dependent (BOLD) fMRI.

It is important to remember that fMRI is a measure of hemodynamic activity and not a direct measure of neural activity. In contrast to EEG/MEG, fMRI produces a 4-D image, 3-D in space and 1-D in time. In other words, there is a time course associated with each location in the brain. The standard spatial resolution is about $2 \times 2 \times 3\,\mathrm{mm}^3$. Using the EPI technique, standard fMRI achieves a temporal sampling on the order of a second.

Block Design and Event-Related Design In today's neural studies, block design (boxcar) and event-related design are the most common experimental techniques. In a block design experiment, stimuli are usually presented at a constant pace over a fixed period of time, alternating with a fixed-length rest period in which no stimuli are presented. The experimental protocol is fixed regardless of subject responses. Because a block is treated as a single unit for analysis, samples within a block are considered as a single condition.

While block design has strong detection power, event-related design has the ability to estimate the shape of the hemodynamic response. Built on rapid imaging, event-related design experiments maximize the opportunity to analyze the data by randomizing and mixing trials or by using differences in speed or accuracy of subject responses. Event-related design does not assume that signals obtained under all physiological tasks share

the same HRF. Instead, it models the HRFs corresponding to various physiological tasks separately. Hence, signals are analyzed by task category. Event-related design trades improved specificity for reduced sensitivity. During an event-related design experiment, stimuli are presented according to a pre-designed protocol, whose timing might change based on the subject's response to the previous stimulus, or on the accuracy of the subject's responses. The event-related design is also called a single trial design because each trial is statistically independent of the other trials. To be statistically independent, different trial types must be intermixed so that it is impossible to predict the next trial type from the previous one.

Although the scheduling of stimulus presentation is different in the two types of experimental protocols, they are both based on a linear assumption of the responses. The total response to multiple stimuli is modeled as the sum of the responses to each individual stimulus. Hence, fMRI data acquired using either protocol can be analyzed via the general linear model (GLM).

fMRI Analysis Given an fMRI scan $[\mathbf{f}_1, ..., \mathbf{f}_N]$, where the vector $\mathbf{f}_n$ denotes the fMRI time course at location n with $T_{\mathbf{F}}$ time samples, the GLM detector produces an activation map by thresholding the statistic value corresponding to a hypothesis-testing procedure. The GLM detector assumes that the observed time course is a linear combination of the protocol-dependent component $\mathbf{B}$ and the protocol-independent component $\mathbf{C}$, such as the cardiopulmonary contributions to the fMRI signal. Activation corresponds to the presence of the protocol-dependent component $\mathbf{B}\boldsymbol{\beta}_n$, leading to a binary hypothesis test:

$$\begin{aligned} H_0: &\quad \mathbf{f}_n = \mathbf{C}\boldsymbol{\alpha}_n + \mathbf{e}_n \\ H_1: &\quad \mathbf{f}_n = \mathbf{C}\boldsymbol{\alpha}_n + \mathbf{B}\boldsymbol{\beta}_n + \mathbf{e}_n \end{aligned} \tag{2.20}$$

for $n = 1, ..., N$, and $\mathbf{e}_n \sim \mathcal{N}(\mathbf{0}, \sigma_n^2\mathbf{I})$. The matrix $\mathbf{D} = [\mathbf{C}\ \mathbf{B}]$ is referred to as the design matrix. The least squares estimates of the activation response $\boldsymbol{\beta}_n$ and the protocol-independent factor $\boldsymbol{\alpha}_n$

$$\begin{bmatrix} \widehat{\boldsymbol{\alpha}}_n \\ \widehat{\boldsymbol{\beta}}_n \end{bmatrix} = (\mathbf{D}^{\mathrm{T}}\mathbf{D})^{-1}\mathbf{D}^{\mathrm{T}}\mathbf{f}_n, \tag{2.21}$$

and the covariance of the estimates $\hat{\Sigma}_{\boldsymbol{\beta}_n}$ are used to form the corresponding F-statistic for rejecting the null hypothesis, $F_n = \frac{1}{N_{\boldsymbol{\beta}}}\widehat{\boldsymbol{\beta}}_n^{\mathrm{T}}\hat{\Sigma}_{\boldsymbol{\beta}_n}^{-1}\hat{\boldsymbol{\beta}}_n$ with $(N_{\boldsymbol{\beta}},\ T_{\mathbf{F}} - \mathrm{rank}(\mathbf{C}))$ degrees of

freedom, where $N_{\boldsymbol{\beta}}$ is the number of regression coefficients in $\boldsymbol{\beta}_n$. Instead of testing the entire estimate $\widehat{\boldsymbol{\beta}}_n$, it is common to apply a contrast matrix to construct the test statistic based on the neuroscientific question of interest. A user-specified threshold is then applied to the statistic at each location to create the activation map. A detailed discussion on the GLM framework can be found in [Friston *et al.*, 1995, Worsley *et al.*, 2002].

■ 2.4.2 DOI

The basic principle of DOI is to illuminate the tissue with an array of near-infrared light sources and to measure the light leaving the tissue with an array of detectors. The recorded signals at each sensor reflect the light reaching that detector from all sources. A model of light propagation in tissue is parameterized in terms of the unknown scattering and absorption as a function of position in the tissue. Then, using the model together with the ensemble of signals over all the sources, one can estimate the scattering and absorption parameters from the data. The estimation method shares similarities with the EEG/MEG source estimation algorithms.

DOI at depths greater than a centimeter is made possible by the relationship of the absorption spectra of three primary absorbers in tissue, namely water, HbO, and HbR, at near-infrared wavelengths between 700 and 1000 nm. At frequencies higher than this range, the absorption by water increases rapidly. Moreover, within this wavelength range, the spectra of HbO and HbR are distinct enough to offer the possibility of performing spectroscopy: illuminating with several wavelengths and recovering separate concentrations of both types of molecules. The physics of light transport in tissue has been explained in a number of recent review articles [Arridge *et al.*, 1999, Boas *et al.*, 2001, Dunsby & French, 2003].

In contrast to fMRI, DOI can separately measure concentrations of HbO and HbR, as well as blood volume with good temporal resolution. However, due to the highly scattered light propagation, we only detect a blurry image of the underlying structure. Compared to fMRI, DOI provides higher temporal resolution and better SNR for weak stimuli such as the median-nerve stimulus. Furthermore, it is inexpensive and portable. Recent work in [Custo & Wells, 2006] proposed a joint DOI and fMRI analysis to exploit the complementary strengths of these two modalities.

2.5 Neurovascular Coupling

As mentioned earlier, neural activity triggers a hemodynamic response. A large amount of fresh blood flows into active regions to supply oxygen and glucose. However, the exact relationship between the neural activity and hemodynamics is still largely unknown, and it is expected to be non-linear and to vary across brain regions.

To date, significant efforts have been devoted to study the relationship between the neural and vascular evoked responses in animals, measuring the two types of signals simultaneously with invasive techniques (see i.e., [Ngai *et al.*, 1999, Jones *et al.*, 2001, Norup Nielsen & Lauritzen, 2001, Devor *et al.*, 2003, Sheth *et al.*, 2003, Logothetis & Wandell, 2004, Ureshi *et al.*, 2004, Martindale *et al.*, 2005, Royl *et al.*, 2006]). Non-invasive studies of the neurovascular coupling are more challenging in humans because of the difficulty in simultaneous acquisition of fMRI and EEG data. While progress in simultaneous EEG-fMRI acquisition has been recently made [Negishi *et al.*, 2004, Ritter & Villringer, 2006, Schmid *et al.*, 2006, Vasios *et al.*, 2006, Mantini *et al.*, 2007, Riera *et al.*, 2007], the SNR of the combined measurements is still not optimal. The use of near infrared spectroscopy (NIRS)/DOI, instead of fMRI to measure the vascular response, in combination with EEG/MEG achieves excellent SNR since there is no interference between the optical and electrical measurements. NIRS/DOI has recently been introduced to study neurovascular coupling in humans in combination with EEG [Obrig *et al.*, 2002, Butti *et al.*, 2006, Rovati *et al.*, 2007, Koch *et al.*, 2008, Herrmann *et al.*, 2008] and DC-MEG [Mackert *et al.*, 2008].

Despite the technical difficulties, neurovascular coupling studies by Logothetis and Wandell have demonstrated similarity in spatial patterns of these two types of activations [Logothetis & Wandell, 2004]. However, the time courses of the neural and the vascular activities differ substantially, and their exact relationship is yet to be fully characterized.

To avoid the technical difficulties, some studies focus on either the relationship between the stimulus train duration and the neural responses or on the relationship between the stimulus train duration and the hemodynamic responses [Gruber *et al.*, 2002, Grill-Spector *et al.*, 2006]. Several fMRI studies have shown a strong non-linear dependency of the vascular response on the variations in stimulus train duration [Vazquez & Noll, 1998, Birn *et al.*, 2001, Birn & Bandettini, 2005]. In particular, it has been shown that while the hemodynamic responses to long stimulus trains behave in an approximately linear

fashion, the responses to shorter stimulus trains (<2 s) are larger than those predicted by a linear model. The non-linearity of the BOLD signal could arise from the habituation effect in the neuronal response, the habituation effect in the hemodynamics, or both.

■ 2.6 Combined EEG/MEG and fMRI Source Estimation

As discussed in Section 2.2, each EEG electrode and MEG sensor reflects activity from all current sources in the brain. Furthermore, there is no unique estimate of the current source from EEG/MEG measurements. In contrast, fMRI has spatial resolution on the order of one millimeter. Leveraging on spatial models of neurovascular coupling, several research groups proposed to incorporate the hemodynamic responses estimated from fMRI data into EEG/MEG source estimation [Liu *et al.*, 1998, Ahlfors & Simpson, 2004, Sato *et al.*, 2004b, Deneux & Faugeras, 2006, Daunizeau *et al.*, 2007]. We categorize them into two groups: point-wise estimation and spatio-temporal estimation.

Point-Wise Solvers The most straightforward way to incorporate fMRI information into EEG/MEG source estimation is through fMRI-weighted minimum-norm estimate (fMNE) [Liu *et al.*, 1998, Ahlfors & Simpson, 2004]. This method uses a thresholded statistical parametric map (SPM) from fMRI GLM analysis to construct weights for the standard MNE, leading to significant improvements when the SPM is accurate. It is formulated as

$$\begin{aligned} \mathbf{J}^{\text{fMNE}} &= \arg\min_{\mathbf{J}} \|\mathbf{RJ}\|_2^2 \\ \text{s.t.} \quad \mathbf{y}(t) &= \mathbf{AJ}, \end{aligned} \tag{2.22}$$

where the $M \times M$ diagonal weight matrix $\mathbf{R}$ identifies locations with hemodynamic activity. We can employ the approach used by the MNE estimator (Eq. (2.15)) to solve for the fMNE estimate, obtaining

$$\mathbf{J}^{\text{fMNE}} = \mathbf{RA}^{\text{T}} \left(\mathbf{ARA}^{\text{T}} + \lambda^2 \mathbf{I}_M\right)^{-1} \mathbf{y}(t). \tag{2.23}$$

Following [Liu *et al.*, 1998], the fMNE weighting parameters are conventionally set to 1 and 0.1 for active and inactive fMRI locations, respectively. The weights depend on arbitrary choices of the threshold and of the weighting parameters. Moreover, these weights are assumed to be identical for all time points in the EEG/MEG source estimation, causing excessive bias in the estimated source time courses.

Sato *et al.* [Sato *et al.*, 2004b] combined the automatic relevance determination (ARD) framework [Tipping, 2001] and fMNE to achieve more focal estimates. In this method, which we will subsequently refer to as fARD, the parameters of a hyper-prior are set based on the thresholded SPM. Similar to fMNE, fARD computes the source estimates for each time point separately and depends on the arbitrary choice of thresholds applied to the SPMs. fARD can be viewed as a "soft" variant of fMNE from the modeling perspective, and its inference procedure often leads to spatially sparse solutions, as shown in [Wipf & Rao, 2004]. The main limitation of fARD is that the estimates may be temporally unstable, often reflected in "spiky" estimated time courses.

Spatio-Temporal Solvers Since the relationship between the dynamics of the evoked neural and the evoked vascular signals is largely unknown, Daunizeau *et al.* only model the similarity of spatial patterns in the two processes. They symmetrically infer brain activities visible in either EEG or fMRI data [Daunizeau *et al.*, 2007]. The confidence of the estimated brain activation is reduced when there are discrepancies between the EEG and the fMRI measurements. In [Deneux & Faugeras, 2006], the temporal relationship between neural and vascular dynamics is approximated using the Balloon model [Buxton *et al.*, 1998, Friston *et al.*, 2000] and the current sources are estimated using the Kalman filter. Due to the complexity of this model, the estimation is limited to a coarse source space, effectively underutilizing the high spatial resolution provided by fMRI measurements.

■ 2.7 Summary

The physics of EEG/MEG implies ill-posed nature of the corresponding source estimation problem. In this chapter, we presented the limitations of several existing source estimation methods, including the diffuse character of the ℓ_2-norm solution and the noise sensitivity of the point-wise MCE (ℓ_1-norm). In the next chapter, we discuss a modified ℓ_1-norm method that incorporates temporal dynamics of the source signals into the estimation procedure to help mitigate this instability.

Moreover, recent findings in neurovascular coupling provide promising directions for leveraging good spatial resolution fMRI signals to improve EEG/MEG source estimation. In Chapters 4 and 5, we discuss our recent study of neurovascular coupling and introduce a new method for joint EEG/MEG-fMRI analysis.

Chapter 3

The $\ell_1\ell_2$-norm Inverse Solver

As discussed in Chapter 2, the minimum current estimator (MCE) (minimum ℓ_1-norm) is the most popular distributed method to model focal activation patterns. However, it is sensitive to noise present in the data, leading to "spiky" estimates. In this work, we aim to improve upon the robustness of MCE.

Similar to MCE, we employ the ℓ_1-norm regularizer to encourage spatial sparsity. We reduce MCE's sensitivity to noise by incorporating information about 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 also suggested in farfield narrowband sensor array applications [Malioutov *et al.*, 2005] to model the diffuse temporal structure of the source signals.

In contrast to the vector-based spatio-temporal minimum ℓ_1-norm solver (VESTAL) [Huang *et al.*, 2006], 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. Performing source estimation 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 human MEG data. The work presented in this chapter was published in [Ou *et al.*, 2009a].

The remainder of this chapter is organized as follows. Section 3.1 describes the $\ell_1\ell_2$-

norm inverse solver. Section 3.2 briefly addresses implementation issues. Section 3.3 presents experimental results using simulated and human MEG data, followed by a discussion and summary.

■ 3.1 Method

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

$$\mathbf{y}(t) = \mathbf{A}\mathbf{J}(t) + \mathbf{e}(t) \tag{3.1}$$

for all time instants as:

$$\mathbf{Y} = \mathbf{A}\mathbf{J} + \mathbf{\Sigma} \tag{3.2}$$

where $\mathbf{Y} = [\mathbf{y}(1), \mathbf{y}(2), \cdots, \mathbf{y}(T_\mathbf{J})]$ is an $M \times T_\mathbf{J}$ matrix that contains EEG/MEG measurements for all $T_\mathbf{J}$ temporal samples, and $\mathbf{J}$ is an $N \times T_\mathbf{J}$ matrix that represents the source signals. Here, we assume that noise $\mathbf{\Sigma}$ is temporally white, i.e., $E\left[\mathbf{\Sigma}^\mathrm{T}\mathbf{\Sigma}\right] = \mathbf{I}$. Time-dependent noise models as those suggested in [Huizenga *et al.*, 2002, Bijma *et al.*, 2005] can be in principle incorporated into the estimation procedure as well.

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, $\Psi = [\vec{\psi}_1, \vec{\psi}_2, \cdots, \vec{\psi}_K]$, that collectively capture the temporal properties of the source signals. $\vec{\psi}_k$, $T_\mathbf{J} \times 1$, denotes the k^{th} basis function. In Section 3.1.2, we will discuss how to obtain the basis appropriate for the source estimation. We assume that the basis functions are orthonormal; if they are not, minor modifications of the remaining derivations are needed, as addressed in Section 3.4.

Projecting both the sensor recordings and the source signals onto the basis functions, the new variables $\widetilde{\mathbf{Y}} = \mathbf{Y}\Psi$ and $\widetilde{\mathbf{J}} = \mathbf{J}\Psi$ are the corresponding projection coefficients. $\widetilde{\mathbf{Y}}$ and $\widetilde{\mathbf{J}}$ are of size $M \times K$ and $N \times K$, respectively. The (n, k) element of $\widetilde{\mathbf{J}}$, $\widetilde{j}_{nk}$, indicates the k^{th} coefficient for the source signal at location n. We can rewrite the original data model in Eq. (3.2) in the transformed domain:

$$\widetilde{\mathbf{Y}} = \mathbf{A}\widetilde{\mathbf{J}} + \widetilde{\mathbf{\Sigma}} \tag{3.3}$$

where $\widetilde{\mathbf{\Sigma}} = \mathbf{\Sigma}\Psi$. We use $\widetilde{\mathbf{e}}_k$ to denote the k^{th} column of $\widetilde{\mathbf{\Sigma}}$. The temporal independence assumption on $\mathbf{\Sigma}$ and orthonormality of Ψ imply that $\widetilde{\mathbf{e}}_k$ and $\widetilde{\mathbf{e}}_{k'}$ are independent for $k \neq k'$ and that $\widetilde{\mathbf{e}}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. Eq. (3.3) 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 Ψ, $\sqrt{\sum_{k=1}^{K} \widetilde{j}_{nk}^2}$, as an indicator of the activation status at 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 estimated 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 source distribution is spatially sparse. 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.

The ℓ_1-norm regularizer in the spatial domain and the ℓ_2-norm regularizer in the temporal domain translate into an integrated spatio-temporal $\ell_1\ell_2$-norm regularizer

$$|\widetilde{\mathbf{J}}|_{\ell_1}^{\ell_2} = \sum_{n=1}^{N} \sqrt{\sum_{k=1}^{K} \widetilde{j}_{nk}^2}. \tag{3.4}$$

Combining Eq. (3.4) with the standard data fit term, the estimation problem thus be-

comes:

$$\widetilde{\mathbf{J}}^* = \arg\min_{\widetilde{\mathbf{J}}} \|\widetilde{\mathbf{Y}} - \mathbf{A}\widetilde{\mathbf{J}}\|_F^2 + \lambda|\widetilde{\mathbf{J}}|_{\ell_1}^{\ell_2} \tag{3.5}$$

$$= \arg\min_{\widetilde{\mathbf{J}}} \sum_{k=1}^{K} \|\widetilde{\mathbf{y}}_k - \mathbf{A}\widetilde{\mathbf{j}}_k\|_{\ell_2}^2 + \lambda|\widetilde{\mathbf{J}}|_{\ell_1}^{\ell_2} \tag{3.6}$$

where $\widetilde{\mathbf{j}}_k$ and $\widetilde{\mathbf{y}}_k$ are the k^{th} column vectors in $\widetilde{\mathbf{J}}$ and $\widetilde{\mathbf{Y}}$. $\|\cdot\|_F$ and $\|\cdot\|_{\ell_2}$ (i.e., $\|\mathbf{x}\|_{\ell_2} = \sqrt{\mathbf{x}^{\mathrm{T}}\mathbf{x}}$) 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 3.3.1. After we obtain the optimal coefficients $\widetilde{\mathbf{J}}^*$, the estimated source signals are linear combinations of the temporal basis functions:

$$\mathbf{J}^* = \widetilde{\mathbf{J}}^*\Psi^{\mathrm{T}}. \tag{3.7}$$

In this work, we formulate the inverse problem as a regularized optimization. It also has an equivalent Bayesian interpretation. The first term in Eq.(3.5) 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.

■ 3.1.1 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. (3.6) since the $\ell_1\ell_2$-norm penalty term is not differentiable at zero. However, Eq. (3.6) can be reduced to the SOCP problem by converting the original unconstrained optimization problem to a constrained one:

$$< \widetilde{\mathbf{J}}^*, q^*, s^*, \mathbf{w}^*, \mathbf{r}^* > = \arg\min_{<\widetilde{\mathbf{J}},q,s,\mathbf{w},\mathbf{r}>} (q + \lambda s) \tag{3.8}$$

$$\text{s.t.} \quad \|\widetilde{\mathbf{y}}_k - \mathbf{A}\widetilde{\mathbf{j}}_k\|_{\ell_2}^2 \le w_k \quad \forall\, k = 1, \cdots, K \tag{3.9}$$

$$\sum_{k=1}^{K} w_k \le q \tag{3.10}$$

$$\sqrt{\sum_{k=1}^{K} \widetilde{j}_{nk}^2} \le r_n \quad \forall\, n = 1, \cdots, N \tag{3.11}$$

$$\sum_{n=1}^{N} r_n \le s. \tag{3.12}$$

New variables, q, s, $\{w_k\}_{k=1}^K$, 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 ψ_k. q is an upper bound on all w_k's. r_n is an upper bound on the activation strength for location n. s is an upper bound on the ℓ_1-norm of the activation strength of all N locations. At the minimum, the values of the constraints in Eq. (3.9)-(3.12) are satisfied with equality; otherwise, the objective function can be further reduced.

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

$$\mathcal{Q}_D := \left\{ \mathbf{x} = (x_0,\ \bar{\mathbf{x}}) \in \mathcal{R}^D : x_0 \geq \|\bar{\mathbf{x}}\|_{\ell_2} \right\} \tag{3.13}$$

where x_0 and $\bar{\mathbf{x}}$ denote the first element and the remaining elements of vector $\mathbf{x}$, respectively. We can see that Eq. (3.11) 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. (3.9), 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. (3.13) 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. Recent work in [Gramfort, 2009] proposed an alternative optimalization procedure, proximal iterations, to solve for SOCP problems.

3.1.2 Temporal Basis Selection

The formulation of our inverse solver is independent of the selected basis Ψ, 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 capture the time-varying frequency content and significant differences in source signals between different regions with a small number of basis functions. 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 in different experimental tasks and subject-to-subject variations. Since according to Eq. (3.2) the sensor signals are linear combinations of the source signals, the temporal patterns of the source signals are 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

$$\mathbf{Y} = \Phi_{\mathbf{Y}} \mathbf{\Lambda}_{\mathbf{Y}} \Psi_{\mathbf{Y}}^{\mathrm{T}} \tag{3.14}$$

where each column in $\Phi_{\mathbf{Y}} = [\phi_1, \phi_2, \cdots, \phi_M]$ denotes the electromagnetic field pattern; each column in $\Psi_{\mathbf{Y}} = [\vec{\psi}_1, \vec{\psi}_2, \cdots, \vec{\psi}_M]$ denotes the temporal pattern. $\mathbf{\Lambda}_{\mathbf{Y}}$ is a diagonal matrix of the singular values in 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 $\Psi = [\vec{\psi}_1, \vec{\psi}_2, \cdots, \vec{\psi}_K]$, but not in the subspace spanned by $\Psi^{\perp} = [\vec{\psi}_{K+1}, \vec{\psi}_{K+2}, \cdots, \vec{\psi}_M]$. Performing source estimation in the signal subspace helps to stabilize the estimated 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. We examine the performance of the proposed solver with varying K in Section 3.3.1 and discuss alternative approaches to basis function selection in Section 3.4.

3.1.3 $\ell_1\ell_2$-norm Estimator 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 parameterization 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. (3.4), $\widetilde{j}_{nk}$ is replaced by a three-dimensional vector, denoting the current dipole moments in the three coordinates $\widetilde{j}_{x_{nk}}$, $\widetilde{j}_{y_{nk}}$, and $\widetilde{j}_{z_{nk}}$. The optimization problem in Eq. (3.8)-(3.12) remains the same except that Eq. (3.11) is replaced by a constraint on the three coordinates:

$$\sqrt{\sum_{k=1}^{K}\left(\widetilde{j}^2_{x_{nk}}+\widetilde{j}^2_{y_{nk}}+\widetilde{j}^2_{z_{nk}}\right)} \leq r_n \qquad \forall\, n = 1:N \tag{3.15}$$

In the original problem, each cone specified in Eq. (3.11) lies in a $K+1$-dimensional subspace; in the free-orientation case, the corresponding cone is extended to a $3K+1$-dimensional subspace. Since the feasible region is an intersection of hyper-cones and hyper-planes, the new formulation is still consistent with the SOCP structure.

■ 3.1.4 Statistical Significance Testing

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 labels. Then we apply the inverse solver to the average data. This procedure preserves the noise covariance structure.

All the results presented in this chapter 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 ipsilateral 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.

■ 3.2 Implementations

■ 3.2.1 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.*, 1999b, 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 localiza-

tion 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 3.1.3.

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 thesis 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 $\mathbf{A}$.

■ 3.2.2 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.

■ 3.2.3 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 appropri-

ate 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 source estimation 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.

■ 3.2.4 Computation Requirements

Compared with the MCE, which solves $T_{\mathbf{J}}$ linear programming problems with N variables each, our solver performs a single SOCP optimization over NK variables. As described in Eq. (3.8) 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.

■ 3.3 Results

Due to a lack of ground truth in human 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 human MEG data from somatosensory and auditory studies.

3.3.1 Simulation Studies

To simulate MEG measurements, we chose active vertices A, B, and C (Fig. 3.1, 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 thesis, 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. 3.2a-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 human 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 $10 \log_{10} \frac{\|\mathbf{AJ}\|_F^2}{MT_{\mathbf{J}}\sigma^2}$, 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. 3.1b 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 3.3.1 and 3.3.1. Curves marked with 'o' in Fig. 3.2a-c correspond to the source signals estimated by the method. The resulting spatial maps and source time courses match well with the ground truth.

$\ell_1\ell_2$-norm vs. ℓ_1-per-coefficient

To further explore the behavior of the $\ell_1\ell_2$-norm regularizer, we compared its estimates 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

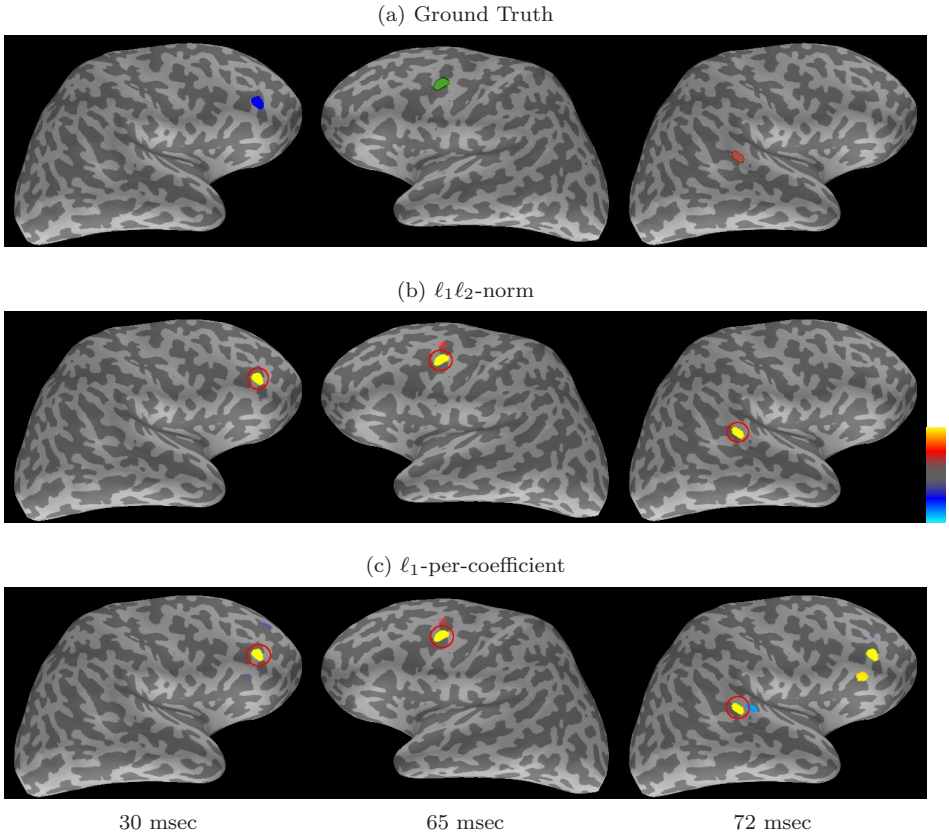
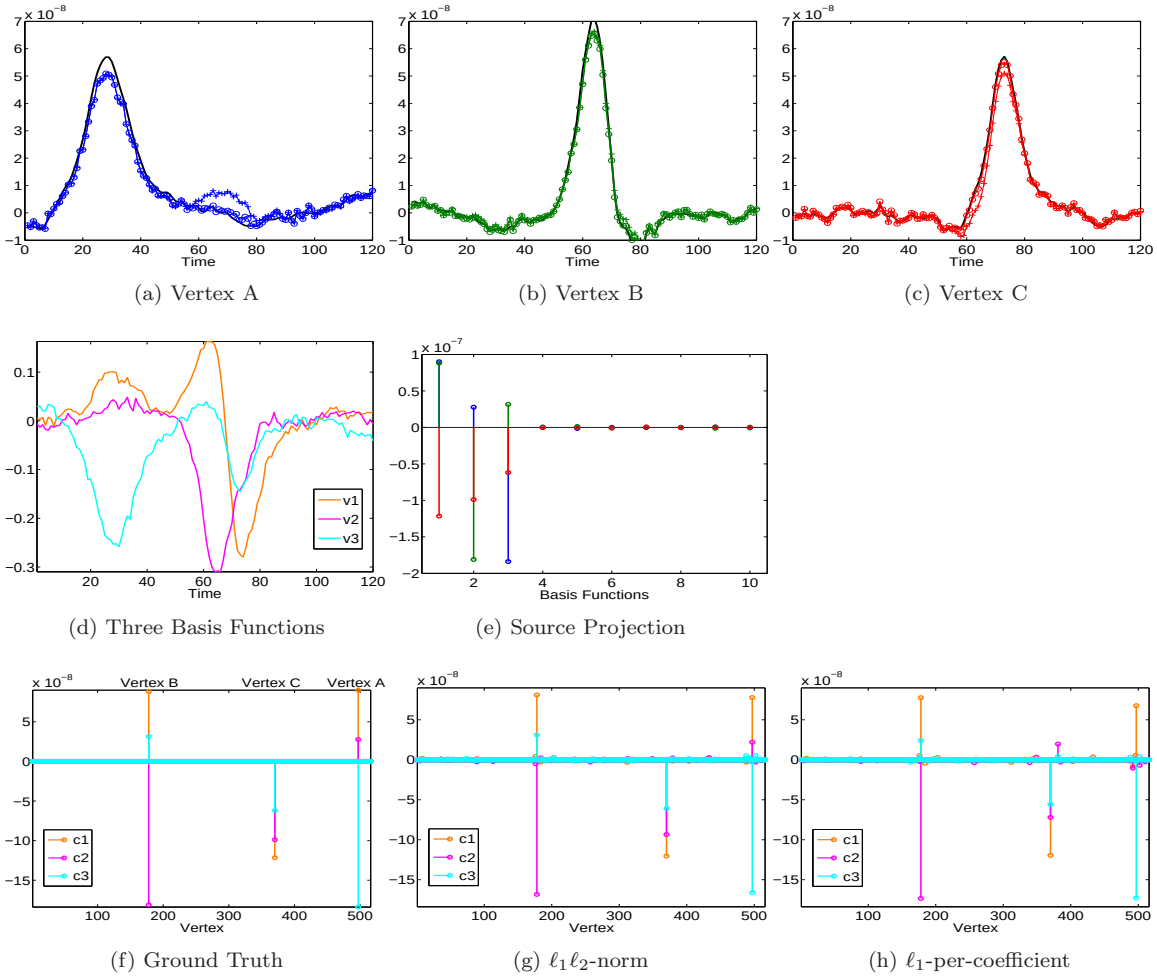


Figure 3.1. Activation maps at different time frames. (a) Ground truth activation maps at peak response time for three sources. (b) The spatial maps estimated using the $\ell_1\ell_2$-norm solver. (c) The spatial maps estimated using the ℓ_1-per-coefficient solver. The color codes in (a) do not indicate current directions. Hot/cold colors in (b,c) correspond to outward/inward current flow. The most active areas in the solutions are highlighted, and their estimated time courses are shown in Fig. 3.2a-c.

of other basis functions. It also achieves stable estimates 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. 3.1c depicts the estimation results, and the corresponding time courses of the three active vertices are presented in Fig. 3.2a-c (marked with '+'). Both $\ell_1\ell_2$-norm and ℓ_1-per-coefficient detect activations in the three vertices (Fig. 3.1), 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 estimated 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.



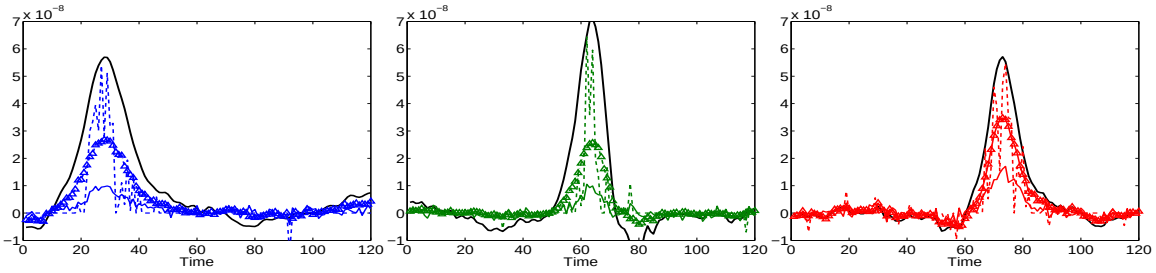
(a) Vertex A (b) Vertex B (c) Vertex C

(d) Three Basis Functions (e) Source Projection

(f) Ground Truth (g) $\ell_1\ell_2$-norm (h) ℓ_1-per-coefficient

Figure 3.2. Estimated source signals, the three basis functions, and the estimated coefficients $\widetilde{\mathbf{J}}$. The top row illustrates the simulated (black solid curves) and the estimated source time courses ($\ell_1\ell_2$-norm marked as 'o' and ℓ_1-per-coefficient marked as '+') for the three active vertices; the middle row presents the selected basis functions (d) and the projection coefficients of the simulated source time courses onto the top ten basis functions (e). (f) Projection coefficients of the simulated data onto the three basis functions, corresponding to the top three singular components. (g) Estimated coefficients from $\ell_1\ell_2$-norm. (h) Estimated coefficients from ℓ_1-per-coefficient.

The three selected basis functions are illustrated in Fig. 3.2d, and the estimated coefficients from these two algorithms are shown in Fig. 3.2(g,h). The projection coefficients of the simulated signals are also presented for comparison purpose (Fig. 3.2f). 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. 3.2g)



(a) Vertex A (b) Vertex B (c) Vertex C

Figure 3.3. Estimated source signals from MNE, MCE, and VESTAL. Each panel illustrates the simulated (black solid curves, same as Fig. 3.2a-c) and the estimated source time courses: MNE (solid), MCE (dashed), and VESTAL (marked as '△').

than in the ℓ_1-per-coefficient solution (Fig. 3.2h). 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 $\vec{\psi}_2$ and $\vec{\psi}_3$. That is illustrated by a more sparse distribution of the coefficients in Fig. 3.2g than that in Fig. 3.2h, a missing $\vec{\psi}_2$ component for Vertex A (Fig. 3.2h), and a false detection for a vertex close to Vertex B (Fig. 3.2h). The ℓ_2-norm regularizer essentially helps bundle basis functions $\vec{\psi}_2$ and $\vec{\psi}_3$ with those that are aligned well with the signal subspace, such as $\vec{\psi}_1$, 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. 3.2e presents the projection coefficients of the simulated source signals onto the temporal basis functions. The coefficients that correspond to basis functions $\{\vec{\psi}_k : k \geq 4\}$ are close to zero. We only displayed the coefficients corresponding to the first ten basis functions. For this simple example, Fig. 3.2e verifies that the selected basis well approximates the signal subspace of the simulated signals.

Comparison with MNE, MCE, and VESTAL

We also compared the proposed method with the standard MNE, MCE, and VESTAL (Fig. 3.6). 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

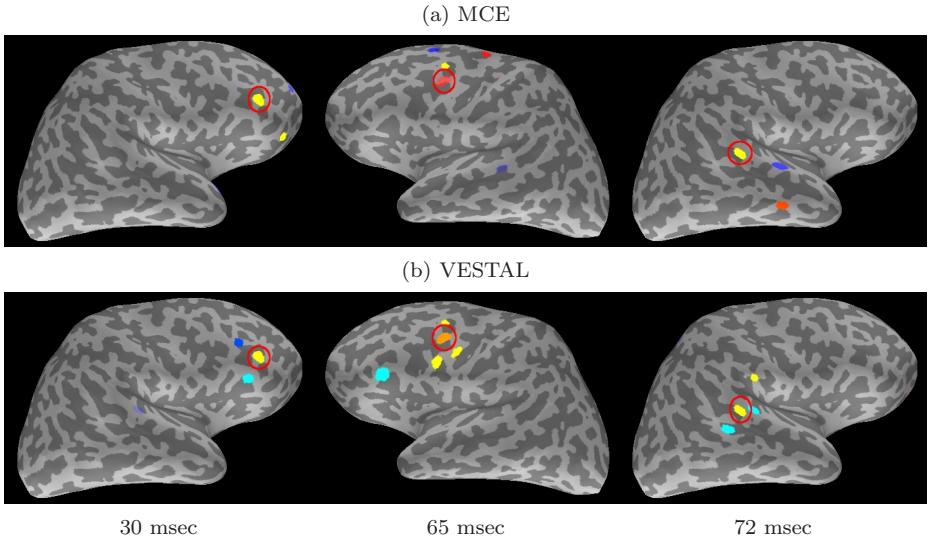


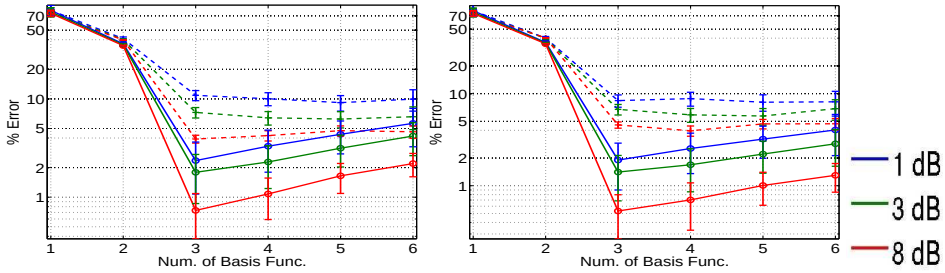
Figure 3.4. Activation maps for MCE and VESTAL at different time frames. (a) The spatial maps estimated using MCE. (b) The spatial maps estimated using VESTAL. Hot/cold colors correspond to outward/inward current flow. The ground truth activation maps are shown in Fig. 3.1(a).

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.

We also compare the estimated source maps from MCE and VESTAL (Fig. 3.4) with those from the $\ell_1\ell_2$-norm estimator (Fig. 3.4(b)). Due to the unstable estimates in MCE and the error propagation in the two-step procedure of VESTAL, MCE's and VESTAL's estimated activation maps have more false positves around the simulated sources compared to $\ell_1\ell_2$-norm results.

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 squares error (MSE) for the three active vertices and for all vertices of the ℓ_1-per-



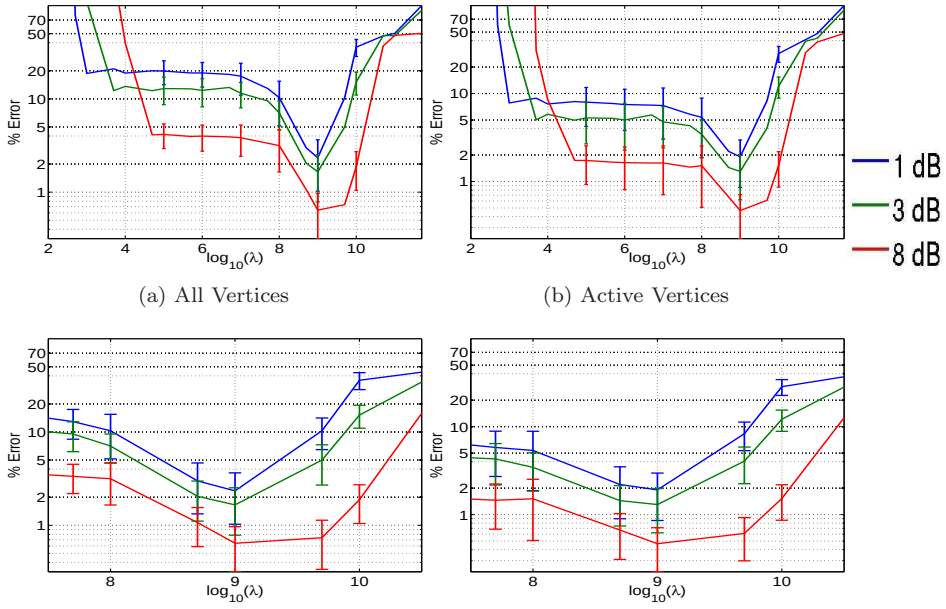
(a) All Vertices $\lambda = 10^9$ (b) Active Vertices $\lambda = 10^9$

Figure 3.5. Relative MSE for different number of selected basis functions. (a,b) present the percentage relative MSE of the estimation 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, respectively. Note that the error bars close to the bottom of the figures appear large due to the logarithmic scale.

coefficient and the $\ell_1\ell_2$-norm inverse solutions are shown in Fig. 3.5. We define the relative MSE as $\|\text{estimated signals} - \text{ground truth signals}\|_F^2 \,/\, \|\text{ground truth signals}\|_F^2$.

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 estimates 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 $\ell_1\ell_2$-norm 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.



(a) All Vertices (b) Active Vertices

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

Figure 3.6. Relative MSE for varying λ. (a,b) present the relative MSE of the estimation 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}. (c,d) are the corresponding zoomed-in versions. Note that the error bars close to the bottom of the figures appear large due to the logarithmic scale.

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. (3.5) is on the order of MK, where M is on the order of 10^2. For an activated vertex in our experiments, $\widetilde{j}_{nk}$ is on the order of 10^{-8}. Hence, $|\widetilde{\mathbf{J}}|_{\ell_1}^{\ell_2}$ is approximately $10^{-7}K$. Therefore, $\lambda = 10^9$ roughly balances between the data fidelity and the regularization terms in Eq. (3.5). In the experiments using human 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. 3.6). As we can see, λ around 10^9 provides accurate es-

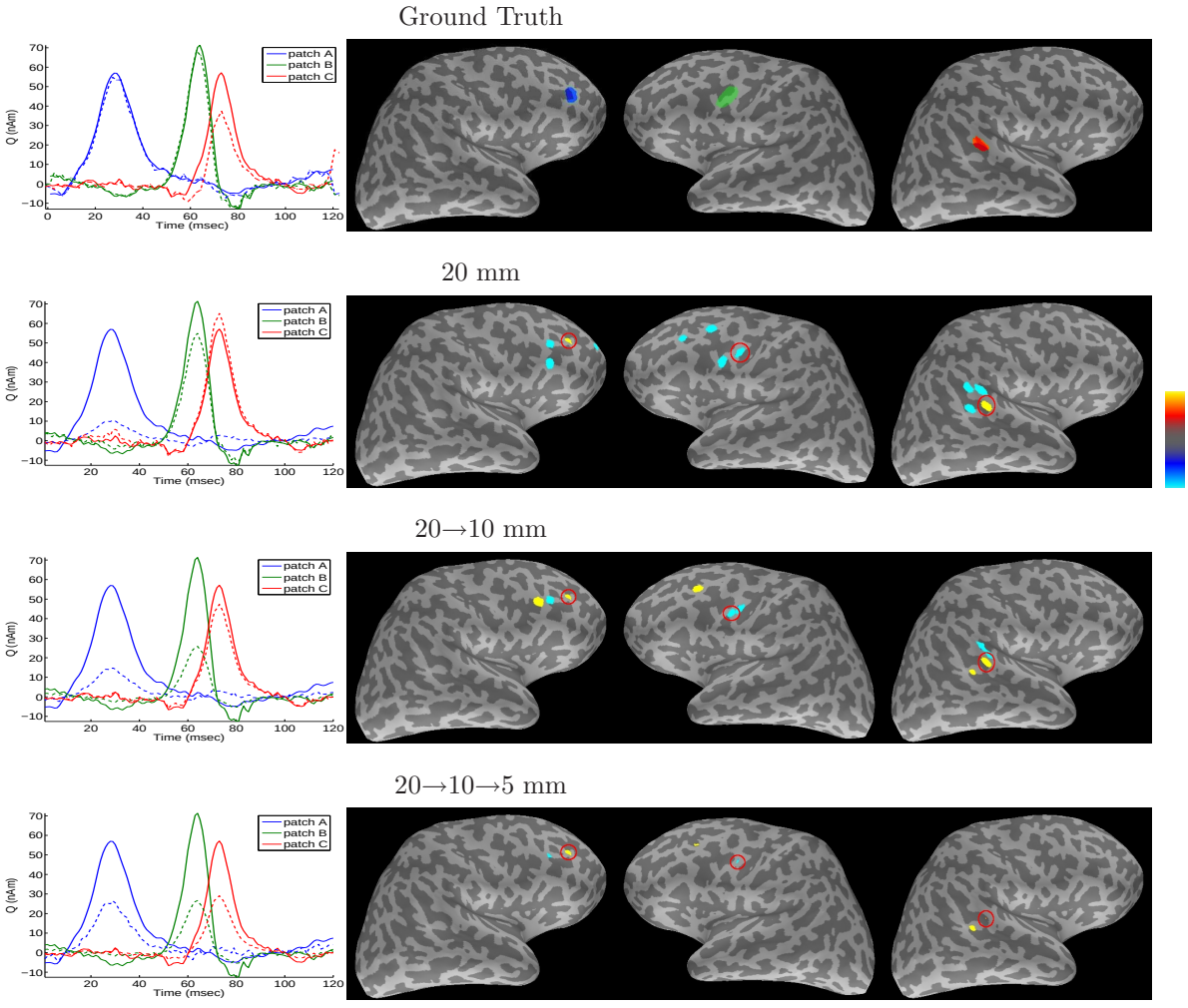
timation results. The regularization shows no effect for $\lambda < 10^3$; when $\lambda > 10^{10}$, the data fidelity term is effectively ignored in the optimization process. For $\lambda = 10^9$, the standard deviation of the MSE estimated from the 100 data sets is less than 1%.

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 3.3.1, 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. 3.7. The ground truth source signals are identical to those employed in Section 3.3.1, shown as solid curves in Fig. 3.7a. To generate the sensor signals, we added Gaussian noise with covariance matrix estimated from the pre-stimulus recordings of a human MEG data set, with a resulting SNR = 3 dB. 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. This experimental setting avoids the "inverse crime," which is defined as that the same forward model are employed to synthesize as well as to invert data in an inverse problem. Furthermore, we set $\lambda = 10^9$ and $K = 3$ in this experiment.

Fig. 3.7 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 blue-cyan or red-yellow corresponding to current flowing inward or outward with respect to the cortex. Each source estimate 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. 3.7a 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 estimated 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 estimated time courses by a factor of four for illustration purposes.

We observe that 20 mm resolution is too coarse for estimation, as reflected by more ambiguity in the source locations. The estimates results at the 5 mm resolution are too focal. At this resolution, the vertices in a simulated patch are close, and some



(a) (b) 30 msec (c) 65 msec (d) 72 msec

Figure 3.7. Source estimates 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 estimated time courses from $\ell_1\ell_2$-norm, which are scaled by a factor of four for illustration purposes.

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

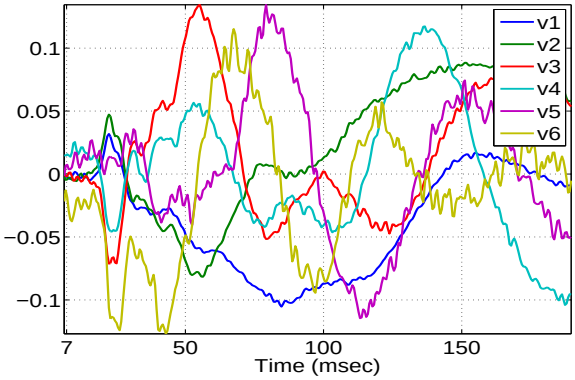


Figure 3.8. Six selected basis functions. The basis functions were obtained from SVD of the MEG measurements between 6 msec and 200 msec after stimulus onset.

allocates all source current to one of those vertices. Estimation 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 human MEG data.

■ 3.3.2 Human MEG Data

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 human 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.*, 1999b, 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.

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[1], 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. 3.8, and they explain 80% variance of the data. 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), medial and posterior to the SI cortex hand area, activates around 70-110 msec. This area, also 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 3.1.4. 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 comparison, we selected the threshold for dSPM so that all four regions of interest, cSI, cSII, iSII, and PPC, were included.

Fig. 3.9 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

[1]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.

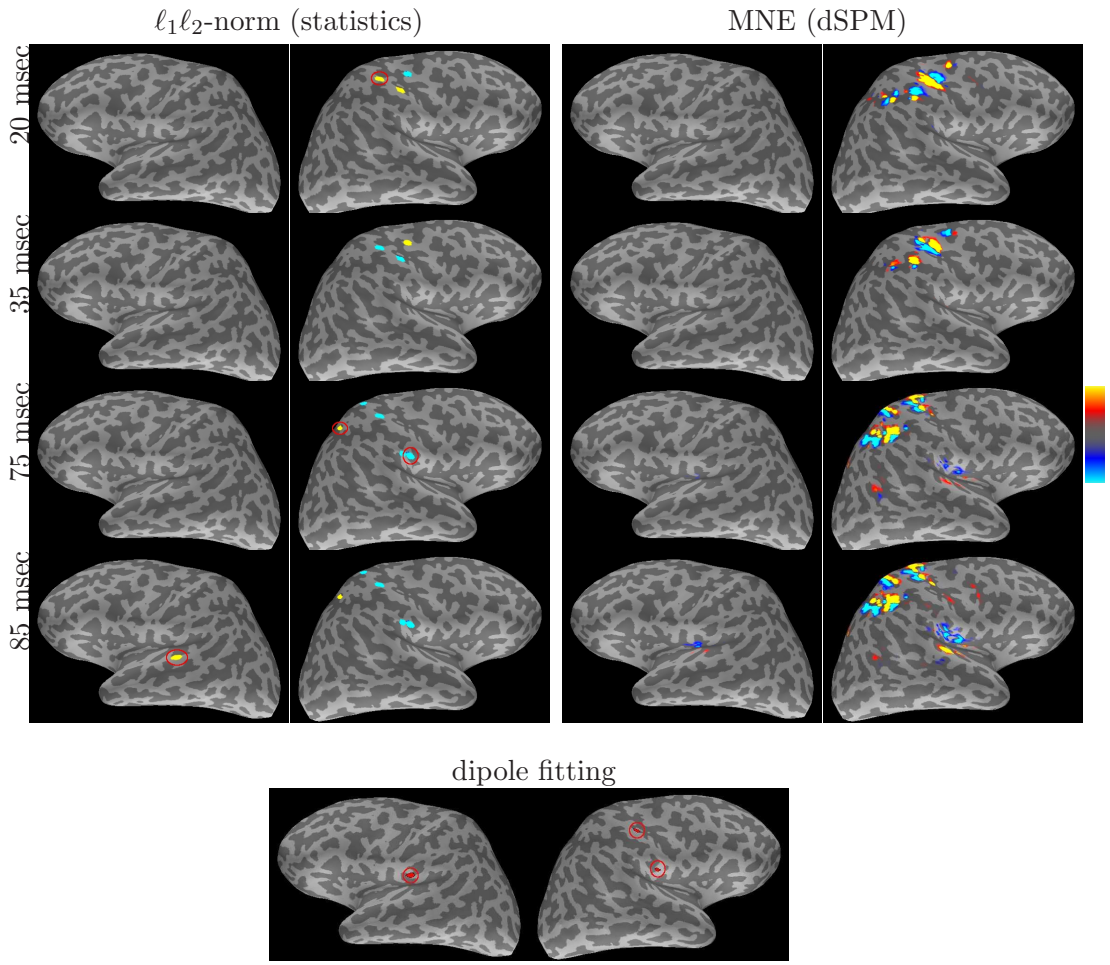
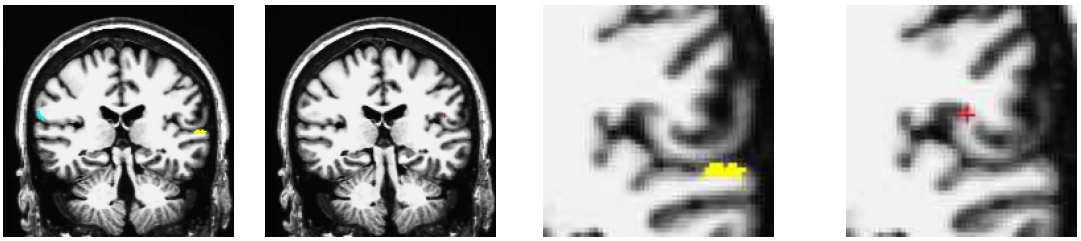


Figure 3.9. Source estimates. Top panel: significance statistics of the $\ell_1\ell_2$-norm solver and MNE for the median-nerve experiment. Hot/cold color corresponds to outward/inward current flow. The most active areas in the $\ell_1\ell_2$-norm solutions are highlighted, and their estimated time courses are shown in Fig. 3.11. Bottom panel: dipole fitting results with three sources. Compared to MNE, $\ell_1\ell_2$-norm pinponts the four regions of interest. Dipole fitting fails to correctly localize PPC since PPC is very close to cSI.

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,



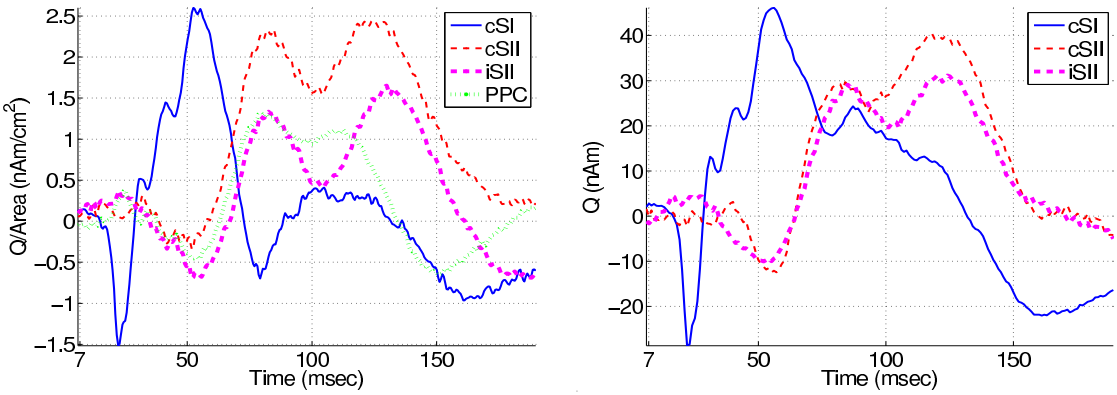
(a) $\ell_1\ell_2$-norm (b) dipole fitting (c) $\ell_1\ell_2$-norm (zoomed-in) (d) dipole fitting (zoomed-in)

Figure 3.10. 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.

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. 3.10), 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. 3.9(bottom), and the corresponding time courses are presented in Fig. 3.11(b). Dipole fitting does not correctly localize PPC from these data because PPC is very close to cSI. The locations for cSI and cSII identified by our solver 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 (Fig. 3.10). 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. 3.11 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



(a) $\ell_1\ell_2$-norm (b) dipole fitting

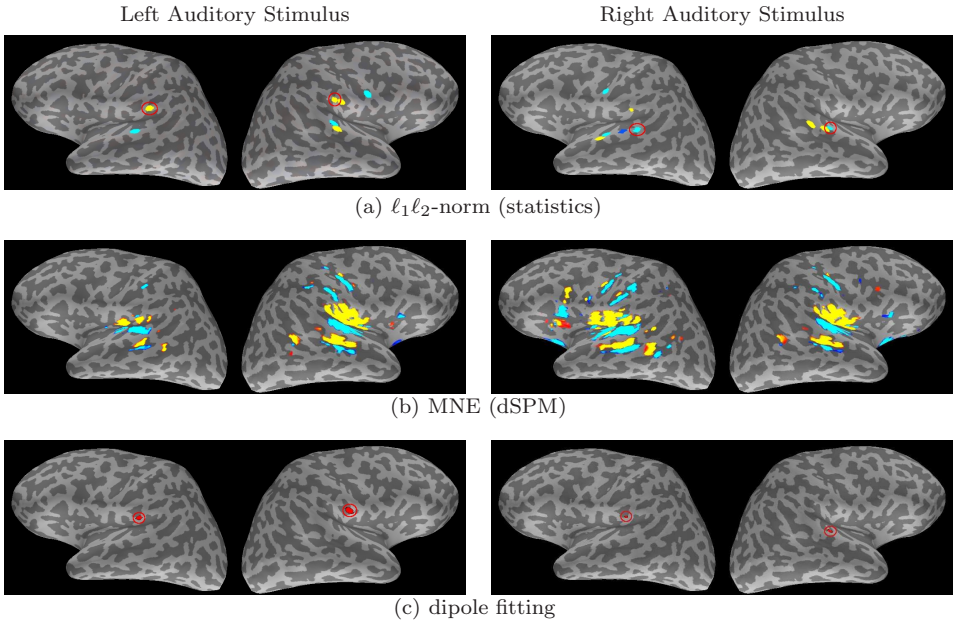
Figure 3.11. Estimated time courses obtained from the $\ell_1\ell_2$-norm solver and dipole fitting. The corresponding activation maps are reported in Fig. 3.9.

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. 3.11 (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 estimations of source signals. Furthermore, $\ell_1\ell_2$-norm does not restrict itself to a fixed number of dipole sources.

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. 3.12 illustrates one frame of the estimated 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



(a) $\ell_1\ell_2$-norm (statistics)

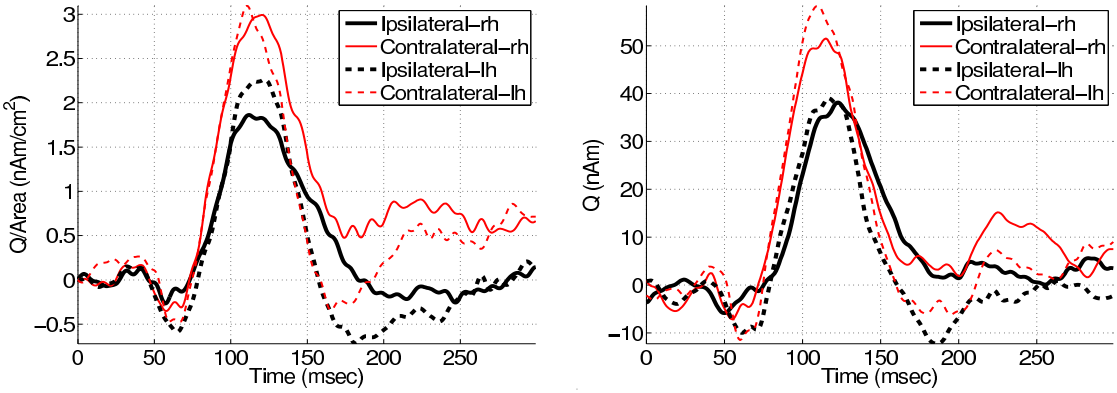
(b) MNE (dSPM)

(c) dipole fitting

Figure 3.12. Significance statistics of the $\ell_1\ell_2$-norm solver and MNE, and dipole fitting at 110 msec after left and right 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. 3.13.

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. 3.13. Both methods detected that the early auditory response occurs around 60 msec and that the contralateral auditory region activates slightly stronger than the ipsilateral 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 ipsilateral activations. Nevertheless, the $\ell_1\ell_2$-norm solver is more flexible than dipole fitting in capturing the spatial extent of the activation regions.



$\ell_1\ell_2$-Norm Dipole fitting

Figure 3.13. Estimated time courses from $\ell_1\ell_2$-norm and dipole fitting.

■ 3.4 Discussion

The proposed inverse solver utilizes an ℓ_1-norm regularizer to encourage spatial sparsity of the activations and an ℓ_2-norm regularizer 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 work 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 estimating the source signals.

Existing ℓ_1-norm related solvers also lack a consensus in handling the free-orientation source estimation. 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 Section 3.1.3.

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 3.3.2. 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 piecewise constant activation pattern. This idea combined with an MNE estimator is at the core of the LORETA estimation method [Pascual-Marqui *et al.*, 1994]. The work of [Auranen *et al.*, 2007, Nummenmaa *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].

■ 3.5 Summary

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 address the problem of 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 estimation in the signal subspace while jointly considering the coefficients for all selected basis functions leads to stable estimates and reduces the number of false positives as confirmed by our experiments using simulated and human MEG data. The $\ell_1\ell_2$-norm solver is formulated as an SOCP problem. The fast optimization method enables us to perform statistical significance testing 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.

Due to the limited spatial resolution of EEG/MEG, it is difficult to distinguish activations on two sides of a sulcal wall, as shown in Section 3.3.2. To improve the source estimation accuracy, in Chapter 5 we introduce a source estimator that takes into account fMRI information which has a better spatial resolution. Since EEG/MEG and fMRI have different physiological origins, we first study the neurovascular coupling relation in Chapter 4.

Chapter 4

Neurovascular Coupling

Since EEG/MEG and fMRI are induced by neural and vascular activities in the brain, we need a neurovascular coupling model to enable joint EEG/MEG-fMRI analysis. To study the neurovascular coupling, we employ recent advances in MEG and diffuse optical image (DOI) to simultaneously record neural and hemodynamic responses. With DOI we recover both oxy- and deoxy-hemoglobin (HbO and HbR) evoked responses in cSI for different stimulus train durations. From MEG data, we estimate the sites of neural activity and extract the cSI current dipole time course. By co-registering the two measurements with respect to each subject's structural MRI, we verified that there is spatial co-localization between neural and vascular activations. We then analyze the neurovascular coupling relationship using standard regression analysis. The work presented in this chapter was published in [Ou *et al.*, 2009b].

This chapter is organized as follows: we present the experimental details including stimulus selection in Section 4.1, MEG-DOI acquisition technique in Section 4.2, and data analysis in Section 4.3. Section 4.4 presents our experimental results in the spatial and the temporal domains, followed by a discussion and summary.

■ 4.1 Somatosensory Stimuli

Trains of median-nerve stimuli (slightly above motor threshold, current < 10 mA, 0.2-msec pulse duration, 4-Hz repetition rate) were presented at the right wrist according to an event-related protocol, with a random inter stimulus interval (ISI) varying between 2 and 18 s (mean 12 s). The train durations were 1, 2, 3, and 4 s, corresponding to 4, 8, 12, and 16 stimuli, respectively. These four conditions were presented in a pseudo-random order [Dale, 1999a] in each of the six runs (six minutes per run), with a minimum of 2-minute rest period between consecutive runs depending on the subject's

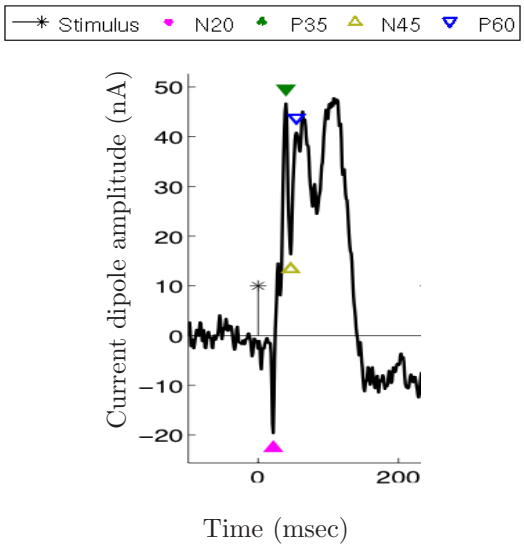


Figure 4.1. Subject $\mathbf{S}_1$'s estimated neural responses to a median-nerve stimulus. The stimulus onset and the detected peaks are marked.

condition. Ideally, we would like to explore a larger range of stimulus train durations; however, we have observed that the MEG response to stimulus trains beyond 4 s is often contaminated by eye-movements and blinks.

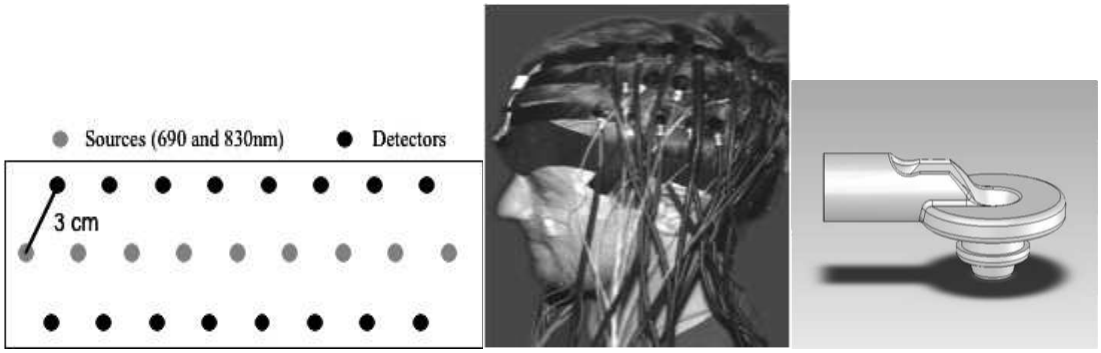
We chose median-nerve electrical stimulation because the corresponding somatosensory responses have been thoroughly studied in intra-cranial electrophysiological experiments in animals. Therefore, our results can be directly compared with, for example, the many forepaw stimulation studies in rats. In addition, the electrical response to median-nerve stimulation has been extensively studied in human with EEG and MEG [Brenner *et al.*, 1978, Goff *et al.*, 1978, Kaufman *et al.*, 1981, Hari *et al.*, 1984, Tiihonen *et al.*, 1989, Allison, 1992, Mauguiere *et al.*, 1997]. It has been shown that a median-nerve stimulus activates a complex somatosensory cortical network [Hari & Forss, 1999], including the contralateral primary sensory cortex (cSI), the bilateral secondary sensory cortex (SII), and other associated areas such as the posterior parietal cortex (PPC). Unlike SII, which lies in the inferior parietal region, deep in the highly folded cortex, the SI hand area, located on the posterior wall of the central sulcus, is relatively superficial and is easily studied with both MEG and DOI. PPC, located on the wall of the post-central sulcus, medial and posterior to the SI hand area, is also relatively superficial to the cortex compared with SII.

As shown in Fig. 4.1, the neural response to median nerve stimulation typically contains four main peaks: N20, P35, N45, and P60. The first peak N20, occurs at 20 msec after stimulus onset, corresponds to intra-cellular currents towards the surface of the cortex, induced by the early excitatory postsynaptic potentials (EPSPs) [Wikström *et al.*, 1996]. The cSI response then changes its polarity and reaches a maximum at 35 msec (P35), corresponding to intra-cellular currents flowing inward from superficial layers into the cortex. This peak has been suggested to be due to the early inhibitory postsynaptic potentials (IPSPs). P35 is followed by a narrow negative peak N45, corresponding to secondary EPSPs, and a broader positive peak P60. N20 and P60 are known to be robust with respect to stimulus repetition rate. Instead, P35 and N45 are highly sensitive to stimulus repetition rate. For instance, several studies have reported that N45 is enhanced with increase in the stimulus repetition rate [Wikström *et al.*, 1996]. In our case, using a stimulus repetition rate of 4 Hz, we observed a strong habituation of these two peaks during the stimulation trains.

■ 4.2 Acquisition

We studied 6 healthy, right-handed subjects (5 males, 1 female). The subjects were 24-61 years old (mean 37 ± 14). Informed consent was obtained from subjects in accordance with the Massachusetts General Hospital ethical committee prior to participation. Measurements were conducted in a magnetically shielded room, on subjects laying on a non-magnetic bed. During the measurements, subjects were alone in the room with the lights dimmed.

The DOI data were acquired with a multi-channel continuous-wave optical imager (CW4 system, TechEn Inc.) as described in [Franceschini *et al.*, 2003]. The CW4 instrument was kept in the control room and we used 10-m long fibers to deliver and collect light to and from the subjects in the magnetically shielded room. The source and detector fiber bundles were attached to three black rubber stripes and arranged in two rows of eight detectors separated by a row of nine sources, for a total of 32 source-detector combinations with 3-cm separation constituting the probe array (see Fig. 4.2a). In each of the nine sources, two laser diodes emitted light at 690 nm and 830 nm wavelength, respectively. The laser wavelengths were chosen to maximize the SNR of the measurements, while minimizing the cross-talk error between HbO and HbR [Yamashita *et al.*, 2001, Strangman *et al.*, 2003, Sato *et al.*, 2004a,



(a) (b) (c)

Figure 4.2. DOI probe. (a) Planar view of the DOI source/detector arrangement. (b) The DOI probe mounted on a subject's head. (c) Schematics of a custom made plug (1 cm thickness) that houses one fiber optic bundle.

Uludag *et al.*, 2004]. The sources were frequency-encoded by steps of approximately 200 Hz between 4.0 kHz and 7.4 kHz, and their signals were acquired simultaneously by the sixteen avalanche photo-diode detectors. The output from each detector was digitized at 40 kHz. The demodulation of the source signals was performed off-line by using an infinite-impulse-response filter, and finally the signals were converted to 10 Hz time series for analysis.

The optical probe covered an area of 6×16 cm^2 and was positioned to extend from left frontal to the left parietal areas with the center approximately above Brodmann area c3 (Fig. 4.2b). The fiber tips terminated with a 90 deg bend and were mounted to the probe by means of low-profile custom-made plugs (1 cm thickness) in order to fit the probe in the limited space between the head of the subject and the MEG dewar (Fig. 4.2c). All materials used in the optical probe were tested in the MEG environment, and special fiber jackets and epoxy, different from those typically used for optical alone or the optical-MRI measurements, were used.

The MEG data were acquired simultaneously with the optical data using a 306-channel Neuromag VectorView MEG device (Elekta-Neuromag Oy, Helsinki, Finland), in a magnetically shielded room (Imedco AG, Switzerland) equipped with adjustable lighting, an intercom, and a video camera. The VectorView sensor array consists of 204 planar gradiometers and 102 magnetometer detectors. Eye movements and blinks were monitored with electrooculogram (EOG). The MEG signals were recorded continuously

with a 1-kHz sampling rate and minimally filtered (0.01 to 250 Hz). There was no special setting adjustment needed for MEG recording in a simultaneous MEG-DOI acquisition, but the SNR of MEG signals suffers slightly from the additional distance between the sensors and the head introduced by the optical probe. The positions of the optical probes and the four head-position indicator coils for MEG were digitized before the measurement with a 3D magnetic digitizer (Polhemus).

In a separate session, MRI anatomical images were collected with a Siemens Avanto 1.5 T scanner with the T1-weighted sagittal MPRAGE protocol and the fast low angle shot (FLASH) protocol at 5 and 30 degrees flip angles. We segmented the head surface and the gray matter cortex from the T1-weighted images [Dale *et al.*, 1999b, Fischl *et al.*, 1999]. The head surface was then co-registered with the digitized points obtained in the joint MEG-DOI session. From the FLASH images, we extracted the inner skull surface to construct the MEG forward matrix using a single-compartment boundary-element model [Hämäläinen & Sarvas, 1989, Oostendorp & Van Oosterom, 1989].

■ 4.3 Analysis

In this section, we first identify the neural and vascular activation areas. We then extract the corresponding activation signals for neurovascular coupling analysis in time domain.

■ 4.3.1 Pre-processing

The optical data were temporally co-registered with the MEG signals by collecting the stimulation triggers with the two systems. The optical raw data were processed off-line using HomER (http://www.nmr.mgh.harvard.edu/DOT/resources/homer/home.htm), an in-house graphical user interface for visualization and analysis of NIRS data implemented in MATLAB (Mathworks Inc., Sherborn, MA). Data were band-pass filtered between 0.02 and 0.5 Hz to reduce the effect of slow drifts and heartbeat, and the temporal changes in the intensity were translated into temporal changes in absorption coefficients using the modified Beer-Lambert law. Concentration of HbO and HbR were derived from the absorption changes at the two wavelengths using hemoglobin extinction coefficients reported in [Wray *et al.*, 1988]. To reduce motion artifacts and systemic oscillation, principal component analysis (PCA) was applied by removing 1-3 components, as described in [Zhang *et al.*, 2005b]. Finally, the hemodynamic responses to each

stimulus train condition were calculated using an ordinary least squares linear deconvolution. The deconvolved hemoglobin responses were averaged over the measured runs to improve the SNR with respect to single-run responses.

Due to the sparse distribution of our source-detector pairs, we limited our analysis to a simple back-projection reconstruction method to identify the vascular activation area. From the 32 source-detector combinations, we identified the source-detector pair with the most statistically significant deoxy-hemoglobin activation (p-value for the selected source-detector pairs across six subjects < 0.05), and used the data of this combination to investigate the neurovascular coupling relation. For the chosen source-detector pair, we computed the hemoglobin time courses for each subject and each condition. Compared to HbO, HbR is known to be less sensitive to systemic oscillations, such as respiration and arterial pulsation [Obrig *et al.*, 2000, Franceschini *et al.*, 2003]. We performed identical analysis for HbR and HbO; however, due to the larger contamination to HbO, and hence a worse SNR, we only display the figures related to HbR in Section 4.4.

For MEG data, we first performed baseline correction and spatial whitening based on a noise covariance matrix computed from signals in a 200-msec window before the stimulus onset. With the MEG forward model and source space constructed as described in Section 3.2, we can estimate the neural response from the MEG measurements by solving the inverse problem [Baillet *et al.*, 2001]. In this work, we employed both dipole fitting [Mosher *et al.*, 1992] and the minimum-norm estimate (MNE) [Hämäläinen & Ilmoniemi,1984], constrained to the cortical gray matter [Dale & Sereno, 1993, Dale *et al.*, 2000]. Our study of the neurovascular coupling in the spatial domain is based on the MNE results. The analysis in the temporal domain is based on the estimated time courses from a standard dipole fitting procedure [Nelder & Mead, 1965, xfit software].

In dipole fitting, we separately fitted the evoked responses of each of the four duration conditions, using a selection of MEG channels (about 45 channels) covering an area above the left SI cortex, at 20-38 msec after the onset of the first stimulus in the train. Subsequently, the time course of this source over the whole epoch was determined with a linear least squares fit. In our single-dipole model, the estimated SI dipole time course usually includes contribution from nearby regions such as PPC. It may also include later activations which often have a more distributed spatial pattern. We believe the neurovascular coupling analysis is still valid despite the fact that the estimates may contain both SI and PPC activations. That is because the hemodynamic optical signals

often reflect a larger spatial pattern, which may very likely contain integrated effects including SI and PPC.

For each subject, after averaging runs for a particular duration condition, we identified N20, P35, N45, and P60 train response as indicated in Fig. 4.5 for condition one in Subject 1. For each subject and each condition, we determined the amplitudes of N20, P35, and P60 responses to each stimulus in the train (i.e., amplitude responses to 4 stimuli for condition one, 8 stimuli for condition two, and so on) by calculating the maximum (or minimum) in a narrow time window. Since the signals were baseline-corrected prior to dipole fitting, the amplitudes of the peaks were set relative to the zero baseline.While several studies have reported that N45 strengthens with reduced stimulus repetition rate, for the 250-msec setting in our experiments, N45 is relatively weak and its amplitude is further affected by P35 due to temporal proximity. Hence, we omitted this peak from the remaining analysis.

■ 4.3.2 Neurovascular Coupling Analysis

To confirm the non-linearity of the vascular responses when the stimulus duration was varied as shown in previous fMRI studies [Birn *et al.*, 2001, Bandettini *et al.*, 2002, Birn & Bandettini, 2005], we applied a linear model to the optical data to predict the responses to long stimulus train duration using the response to short stimulus train. For instance, to linearly predict the response to a 2-s stimulus train, we added the response to a 1-s train and its 1-s delayed version. To linearly predict the response to 3-s stimulus train, the 1-s response was replicated three times, shifted by 0, 1, and 2 s, and then summed. We performed similar linear predictions for the response to a 4-s train using the 1-s and the 2-s responses.

To evaluate the correlation between the hemoglobin and neural response, we followed two data analysis approaches. In the first approach, we evaluated the linearity of the responses by plotting the response amplitudes versus the number of stimuli and calculated the coefficient of determination (R^2) to compare results. In particular, for the hemodynamic response and for each subject, we calculated the area under the HbR and HbO responses for each condition and normalized them with respect to the largest response area. This normalization removes artifacts due to subject variability in head size and in probe location with respect to center of activation. For each condition, we then averaged the responses of the six subjects together and evaluated linearity with zero crossing imposed for HbO and HbR, as shown in Fig. 4.6.

To remove artifacts in MEG responses due to subject variability in source depth and head size, we normalized the peaks amplitudes with respect to the amplitude of the largest P35 in response to the first stimulus in the train among the four conditions. We then averaged the cumulative MEG responses to each stimulus across subjects, deriving the first 4 responses from condition one, the second 4 responses from condition two, responses to stimuli 9 to 12 from condition three, and finally responses to stimuli 13 to 16 from condition four, and we calculated R^2 and zero intersection of the three peaks, N20, P35, and P60.

In addition to the peak amplitudes, we also considered the sum of the peak amplitudes in absolute value, including P35-N20, P60-N20, and P35+P60-N20, and the root mean square (RMS) value of the dipole amplitude in the time window after N20 and before the response returning to baseline (see Fig. 4.5). For subject 2 and 5, the window is set as 30-160 msec after each stimulus onset in a train; for the rest of the subjects, the window is set as 30-140 msec. The chosen time window covers the early activations such as SI and PPC [Hari & Forss, 1999]. We also performed calculations including different time windows and did not see statistical differences on the results with respect to the chosen one.

In the second more rigorous approach, for each subject we estimated the linear convolution model between the measured hemoglobin time series to each condition and the measured neural response time series, including N20, P35, P60, sum of peak amplitudes in absolution values, and the RMS value, as described above.

For all experiments, the temporal length of the impulse response function (h) was set to 10 s, which corresponded to 100 time samples for the hemoglobin responses. We estimate a separate impulse response function for HbO ($h_{\mathrm{HbO}}(t)$) and HbR ($h_{\mathrm{HbR}}(t)$). This convolution model, also used in [Logothetis & Wandell, 2004, Martindale *et al.*, 2005], assumes that the neuronal activity and hemoglobin responses are related by

$$\begin{aligned} f_{\mathrm{HbR}}(t) &= h_{\mathrm{HbR}}(t) * j(t) + \epsilon_{\mathrm{HbR}}(t), \\ f_{\mathrm{HbO}}(t) &= h_{\mathrm{HbO}}(t) * j(t) + \epsilon_{\mathrm{HbO}}(t) \end{aligned} \tag{4.1}$$

where $j(t)$ is the time course of neuronal activity represented by a MEG component for each stimulus pulse in the pulse train. For example, when we compute the regression between peak N20 and the hemoglobin signals, $j(t)$ is a set of delta functions whose amplitudes equal to the corresponding N20 peak amplitudes to different stimuli in the stimulus train. Function $f_{\mathrm{HbO}}(t)$ and $f_{\mathrm{HbR}}(t)$ are the time courses of the hemoglobin

signals, and $h_{\mathrm{HbO}}(t)$ and $h_{\mathrm{HbR}}(t)$ are the hemoglobin response functions. The time windows of $f_{\mathrm{HbO}}(t)$ and $f_{\mathrm{HbR}}(t)$ for the linear convolution analysis are separately selected based on the zero crossing of the corresponding responses with respect to the baseline for each condition and each subject. In Eq. (4.1), $*$ indicates the convolution of two time series. $\epsilon_{\mathrm{HbO}}(t)$ and $\epsilon_{\mathrm{HbR}}(t)$ include background physiology and measurement noise, which we assume to be white noise. The impulse responses $h_{\mathrm{HbO}}(t)$ and $h_{\mathrm{HbR}}(t)$ are estimated for all stimulus conditions using least squares linear deconvolution for each subject separately, and we denote the corresponding estimates as $\widehat{h}_{\mathrm{HbO}}(t)$ and $\widehat{h}_{\mathrm{HbR}}(t)$. Predicting the variation in HbO and HbR responses to the different stimulus conditions is then possible using a linear transformation. Results for each subject are shown in Fig. 4.7.

A good linear fit is reflected by a good match between the predicted and measured hemoglobin response, which is mathematically evaluated using the coefficients of determination (R^2). For example, in the analysis of HbR, the corresponding R^2 is

$$R^2 = 1 - \frac{\mathrm{var}\left(f_{\mathrm{HbR}}(t) - \widehat{h}_{\mathrm{HbR}}(t) * j(t)\right)}{\mathrm{var}(f_{\mathrm{HbR}}(t))} \tag{4.2}$$

We then convert the R^2 value into the Fisher's z-score:

$$z = \frac{1}{2} \ln\left(\frac{1+R}{1-R}\right) \tag{4.3}$$

A two-sample t-test based on the Fisher's z-scores between the linear predictions from two types of neural components were calculated to assess the statistical significance (p-values ≤ 0.05 between different components).

4.4 Results

In this section, we first study the spatial relationship between neural and vascular activities. We then focus on the temporal relationship at the primary sensory area.

4.4.1 Co-localization of MEG Sources and DOI Signals

The median-nerve stimulus employed in this work is known to robustly activate cSI (Brodmann area 3) [Hari & Forss, 1999]. To verify the co-localization of the neural and the vascular activations, we present the estimated activation maps from MEG data (yellow area) using the MNE software [Hämäläinen, 2005] and the contour levels of HbR

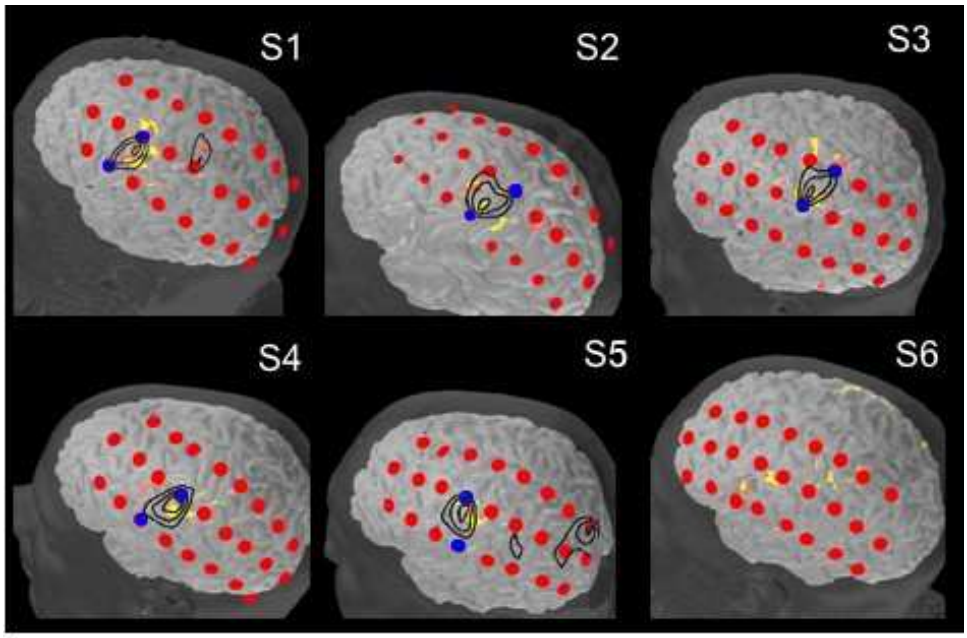


Figure 4.3. Activation maps for six subjects. Neural activation estimated from MEG measurements at 35 msec after the onset of the first stimulus in the stimulus train is depicted in yellow. The contour plot correspond to 90%, 75%, and 50% thresholds of maximum signal decrease in HbR estimated from the DOI measurements in the interval 3-5 s after stimulation onset. Three rows of DOI source-detector pairs are shown as red and blue circles. For each subject, the blue circles indicate the source-detector pair that detect the strongest vascular responses. For subject 6, the digitized points for the optical probes were lost. While the optical response concentrates in a small area similar to the one of the other subjects, localization of HbR with respect to the brain and MEG sources is not possible. Hence, we cannot show the overlaid HbR contour plot for this subject.

signal changes overlaid on the corresponding MRI cortices (Fig. 4.3). The reconstructed HbR map for each subject was obtained in HomER using a standard back-projection method. The contour levels represent the 90%, 75%, and 50% thresholds of maximum signal decrease between 3 – 5 s post-stimulation onset with respect to -2 – 0 s before stimulation onset for the 4-s stimulus train duration condition. The optode positions recovered with the 3D digitizer are depicted as red and blue circles in Fig. 4.3. For each subject, the source-detector pair with the most statistically significant HbR changes are shown in blue. In one subject ($\mathbf{S}_6$), the digitization of the optical probe was lost due to repositioning of the probe after digitization, and only the MEG source location is shown. In the other five subjects with digitized optode positions, the vascular responses,

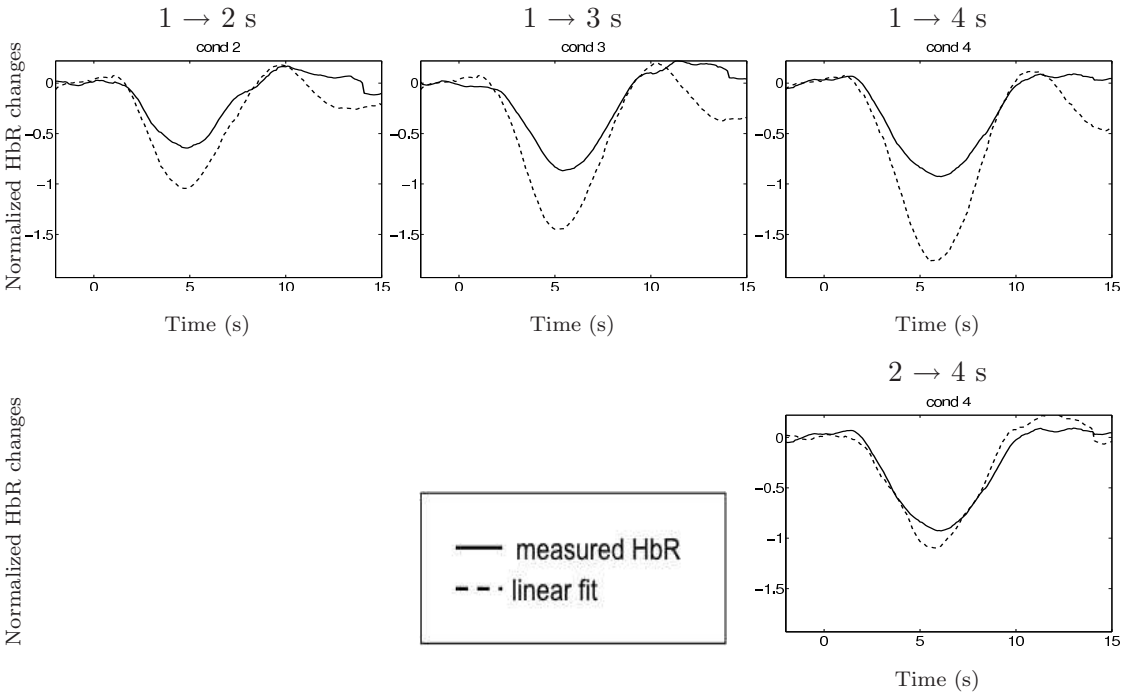


Figure 4.4. Linear predicted HbR from measured HbR to short stimulus duration. Solid curves show the grand-average HbR measured with DOI, and dashed curves are the corresponding predictions assuming a linear model. Top: linear prediction of responses to 2-, 3-, and 4-s stimulus train duration conditions based on the response to a 1-s train. Bottom: linear prediction of response to a 4-s stimulus train duration condition based on the response to a 2-s train.

reflected by the largest signal decreases in HbR, are spatially co-localized with the estimated neural activations from the MEG measurements. The results for HbO, not shown, are consistent with the ones for HbR.

■ 4.4.2 Non-Linear DOI Responses

The grand-averaged HbRs over all six subjects obtained from the selected source-detector pairs for 2-, 3-, and 4-s train durations are presented in Fig. 4.4 (solid curves). We used the HbR responses corresponding to 1- and 2-s trains separately to linearly predict the responses to longer stimulus train duration as described in Section 4.3.2. The linear predictions (dashed curves) are overlayed with the corresponding measured HbR.

We can see that the HbR response to a 1-s train overestimates the responses to longer train durations using a linear model, reflecting an over-shoot of vascular sup-

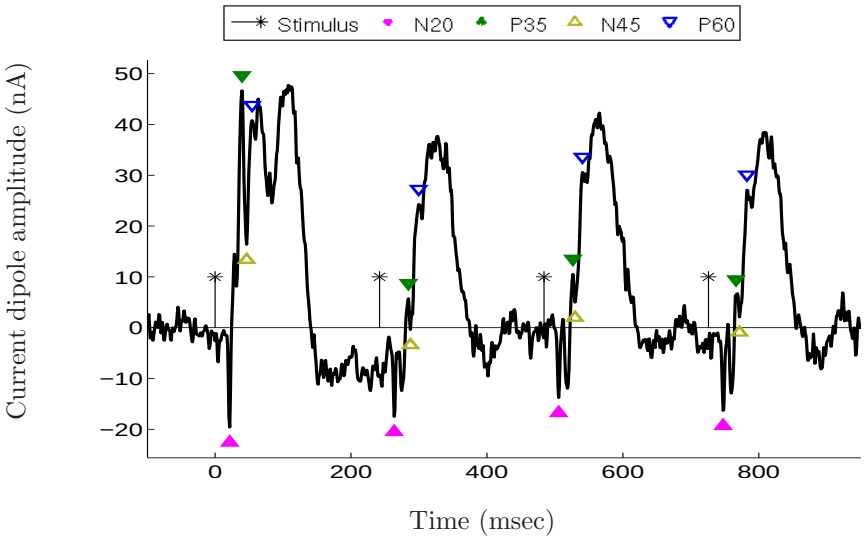


Figure 4.5. Subject $\mathbf{S}_1$'s estimated neural responses to the 1-s stimulus train duration condition (four stimuli). The stimulus onsets and the detected peaks are marked. For the neurovascular coupling analysis, we considered each individual peak amplitude, the difference of peak amplitudes, the sum of the peak amplitudes, and the RMS value of the current dipole amplitude in a time window between 30 and 140 msec after each stimulus onset for each of the stimulus train durations.

ply to short stimulus trains. In contrast, the HbR response to a 2-s train provides a good linear fit to the response to the 4-s train. Our DOI results confirm that the responses to short stimulus train duration are larger than those estimated using a linear model [Bandettini *et al.*, 2002, Birn *et al.*, 2001, Birn & Bandettini, 2005].

■ 4.4.3 Linear MEG Responses

Fig. 4.5 shows the evoked cSI response to the 1-s stimulus train in a subject ($\mathbf{S}_1$), with the major peaks highlighted. The evoked responses to stimulus 5 through stimulus 16 in the other duration conditions are similar to the response to the fourth stimulus in the 1-s condition, but slightly noisier.

For most of the subjects, the amplitudes of N20 and P60 are similar in all stimuli in a train. On the other hand, P35's amplitude drops significantly in response to the second to the last stimuli in a train. If the vascular response were linearly related to either N20 or P60, the vascular response would have been linearly related to the stimulus train duration for the four conditions that we examined. In contrast, if the vascular response were linearly related to P35, the vascular response would have a

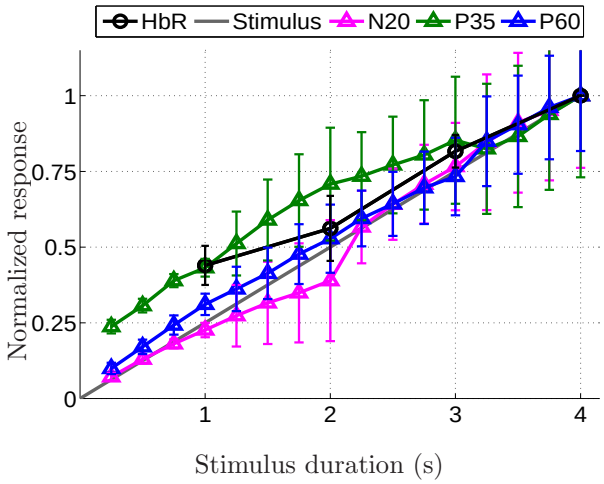


Figure 4.6. Average normalized responses of HbR (area) and average normalized amplitudes of the three peaks in the neural responses across subjects under the four stimulus duration conditions.

similar amplitude across the four conditions. Therefore, the linear relationship between HbR and any single peak is not adequate.

Fig. 4.6 shows the cumulated sum of the amplitudes of each of the three peaks averaged over the 6 subjects for each train duration. In contrast to the hemodynamic responses (also shown in Fig. 4.6 as the grand average of the areas under the HbR responses for each condition) which is not linear with respect to train duration ($R^2 = 0.96$), both N20 and P60 are linear with respect to stimulus train duration ($R^2 \geq 0.99$). On the other hand, P35 exhibits habituation effect in response to the second and later stimuli in a train, reflected by the smallest R^2 value ($R^2 = 0.93$) and the largest zero intersection (0.26).

■ 4.4.4 Hemodynamic Predictions

Fig. 4.7 and Fig. 4.8 illustrate the measured HbR, $f_{\text{HbR}}(t)$, and the linear prediction, $\widehat{h}_{\text{HbR}}(t) * j(t)$, using peak amplitudes, sum of the peak amplitudes in absolute value, and the RMS value of the neural response for the six subjects and the four stimulus train durations. It reveals that none of the peaks can individually explain the HbR response with a linear convolution model. It also reveals that the sum of three peaks (red dashed curves) provides relatively better linear fit to HbR consistently across all six subjects.

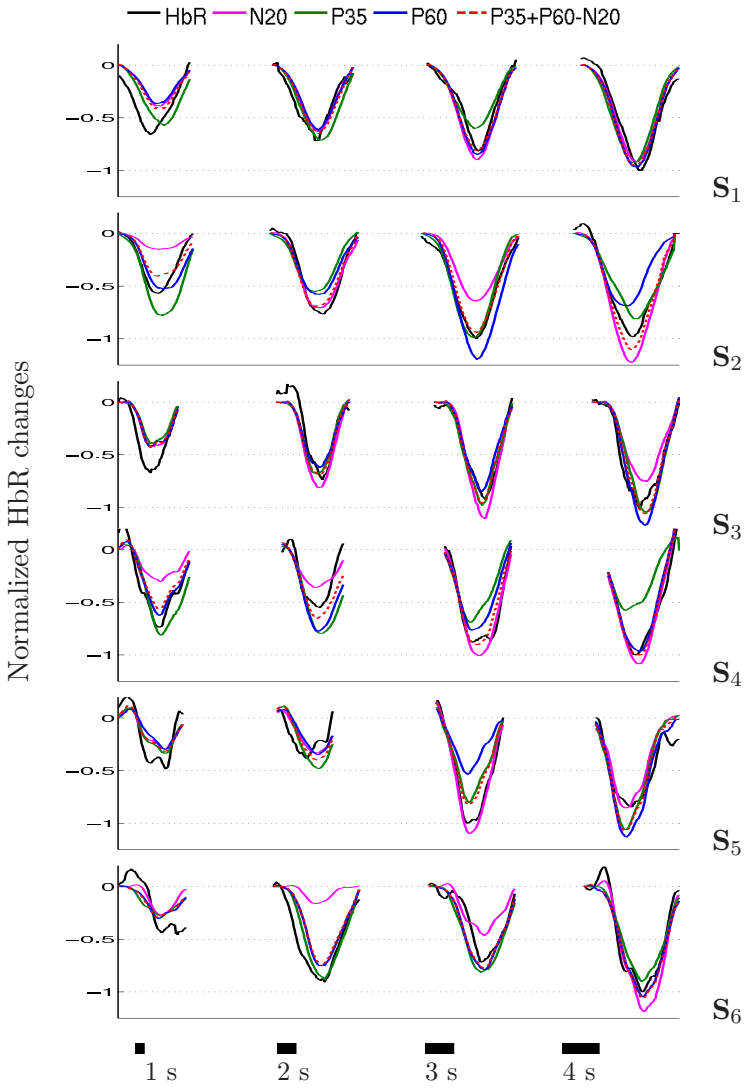


Figure 4.7. Linear convolution analysis. For the four stimulus train duration conditions and for each of the six subjects, the measured HbR using DOI (black curve) and the predicted HbR, $\widehat{h}_{\mathrm{HbR}}(t)$, based on the linear convolution model with different components in the neural responses are displayed above. The predictions are computed using N20 (pink curve), P35 (green curve), P60 (blue curve), and P35+P60−N20 (red dashed curve).

We evaluated the quality of the predictions based on differences in the coefficients of determination (R^2) between the predicted and the measured time courses of the hemodynamic responses, as shown in Fig. 4.9. Table 4.1 summarizes the results of a

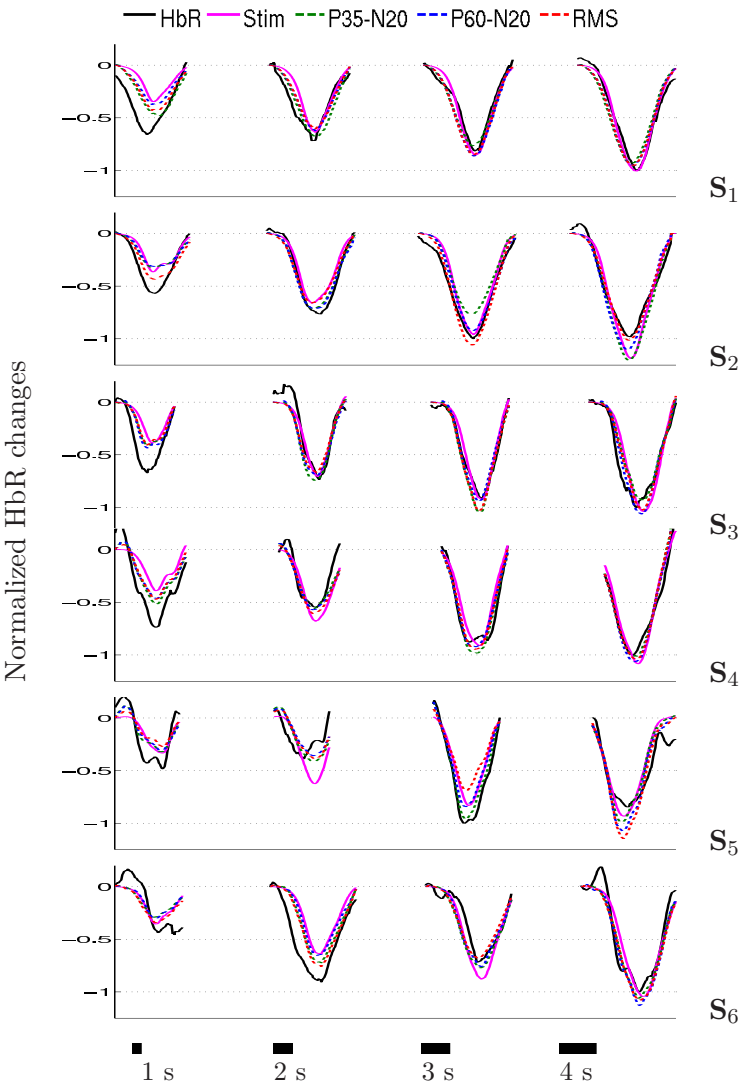


Figure 4.8. Linear convolution analysis. For the four stimulus train duration conditions and for each of the six subjects, the measured HbR using DOI (black curve) and the predicted HbR, $\widehat{h}_{\mathrm{HbR}}(t)$, based on the linear convolution model with different components in the neural responses are displayed above. The predictions are computed using the stimulus onset (pink curve), P35−N20 (green dashed curve), P60−N20 (blue dashed curve), and the RMS value (red dashed curve).

two-sample t-test. In particular, Table 4.1 (top) shows a cross-reference of p-values for statistically significant differences in R^2 values corresponding to different components in predicting HbR. Table 4.1 (bottom) shows the corresponding results for HbO. The

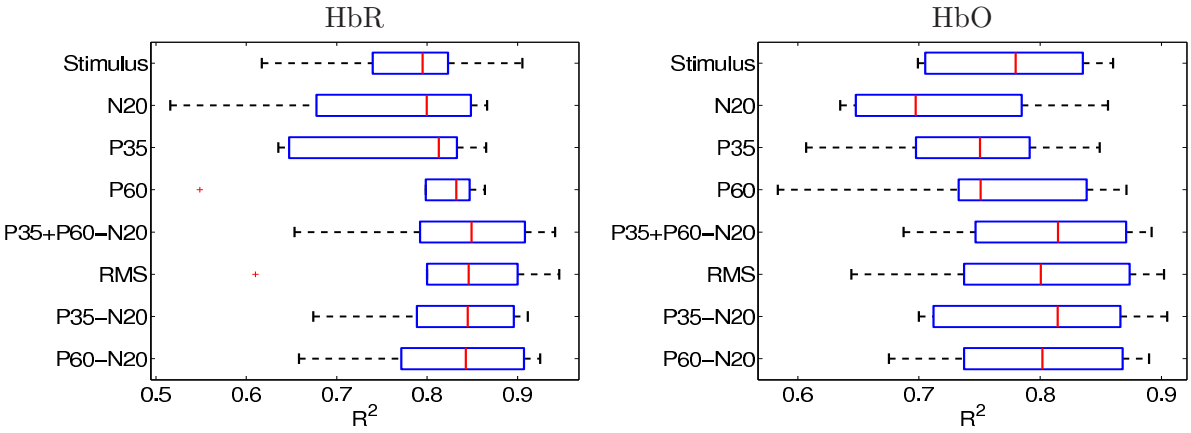


Figure 4.9. R^2 values of linear convolution analysis for different components in the neural responses. The corresponding results of two-sample t-test are reported in Table 4.1.

average R^2 across six subjects is presented in the right-most column. The sum of peak amplitudes and the RMS value of the neural response provide the best linear prediction compared to any single component, reflected by the largest R^2 values. All of them are statistically better than the input stimuli in predicting HbR, and the sum of the three peaks also achieves statistically better prediction than N20.

Since HbO is more sensitive to systemic physiological effects than HbR, the SNR is lower than that for HbR. Although the results for HbO are similar to those for HbR, the R^2 values of the HbO predictions are smaller than the corresponding ones for HbR prediction, and statistically significant better prediction is established for the sum of three peaks and the RMS value with respect to N20.

■ 4.5 Discussion

Previous studies using combination of MEG and DOI have employed direct current (DC) MEG measurement techniques to study DC or near-DC changes in the magnetic fields in a block-design finger-movement paradigm with 30-s blocks of finger movements alternating with rest periods [Mackert *et al.*, 2004, Sander *et al.*, 2007, Mackert *et al.*, 2008]. Limited to one type of stimulus (stimulus frequency and duration), these studies could not explore the linear/non-linear neurovascular coupling relationship. In contrast, here we investigated the relationship of the DOI vascular signals with the detailed time course of MEG current sources. The use of an event-related rather than a block design

Statistical Significance of better Linear Predictions in terms of R^2 between Measured and Predicted HbR

	Stimulus	N20	P35	P60	P35+P60-N20	RMS	P35-N20	P60-N20	Mean R^2
Stimulus	N/A								0.78
N20	0.42	N/A							0.75
P35	0.48	0.58	N/A						0.77
P60	0.52	0.33	0.42	N/A					0.79
P35+P60-N20	**0.01**	**0.05**	0.10	0.08	N/A				0.83
RMS	**0.01**	0.08	0.12	0.12	0.37	N/A			0.83
P35-N20	**0.05**	**0.04**	0.16	**0.05**	0.31	0.44	N/A		0.83
P60-N20	**0.02**	**0.05**	0.15	0.11	0.10	0.38	0.56	N/A	0.83

Statistical Significance of better Linear Predictions in terms of R^2 between Measured and Predicted HbO

	Stimulus	N20	P35	P60	P35+P60-N20	RMS	P35-N20	P60-N20	Mean R^2
Stimulus	N/A								0.78
N20	0.10	N/A							0.72
P35	0.36	0.53	N/A						0.74
P60	0.39	0.35	0.48	N/A					0.76
P35+P60-N20	**0.03**	**0.03**	0.17	0.08	N/A				0.80
RMS	0.13	**0.05**	0.24	0.12	0.27	N/A			0.79
P35-N20	0.08	**0.02**	0.19	0.09	0.56	0.47	N/A		0.80
P60-N20	0.07	**0.04**	0.22	0.14	0.02	0.56	0.44	N/A	0.80

Table 4.1. Cross-reference of p-values for significant differences in R^2 values for each subject and each component used to predict HbR (top) and HbO (bottom). p-values ≤ 0.05 (bold font) indicate statistically significantly better prediction of the component with a higher R^2 with respect to the corresponding component with a lower R^2. The average coefficients of determination, R^2, between the measured and the predicted HbR (or HbO) responses are reported in the last column.

allows us to employ shorter stimulation periods (1-4 s vs. 30 s) and to evaluate the contribution of single neural component to the hemodynamic responses. Our approach enables a direct comparison between our results and the vast number of neurovascular coupling findings in animals using invasive electric and vascular recordings.

The choice of limiting the study to four conditions is dictated by scanning time and subject's condition. In order to examine more duration conditions, it is necessary to extend the scanning time to ensure a sufficient SNR in the average signals for the neurovascular coupling analysis. The choice of 1 s as the shortest condition is due to the fact that hemoglobin response for shorter stimuli is too noisy; the choice of 4 s as the longest duration is due to the fact that MEG signal for longer stimuli become noisier because of eye blinks and muscle contractions. In any case, the 1 to 4 s duration chosen are within the range of non-linearity previously observed in the hemodynamic response

with BOLD fMRI [Bandettini *et al.*, 2002].

We observed a good spatial agreement between the activation areas determined by the two imaging modalities. The neural source location determined using MNE in five subjects is in good proximity with the maximum hemoglobin evoked changes measured with DOI. This result reveals the benefit of integrating other vascular measurements such as fMRI, which has a good spatial resolution, to help solve the ill-posed MEG inverse problem. Our result also indicates that a combined MEG-DOI inverse formulation may provide additional information for the two ill-posed inverse problems.

We verified that the hemodynamic response is non-linear with respect to stimulus train duration ($R^2 = 0.96$ without zero intersection constraint and $R^2 = 0.22$ with zero intersection constraint). Our result shows that the hemodynamic responses to short stimulus trains (< 2 s) measured with NIRS are larger than those predicted by a linear model, and it is in agreement with fMRI results in humans [Bandettini & Ungerleider, 2001, Birn *et al.*, 2001]. With the simultaneous MEG measurements, we tested whether the hemodynamic non-linearity can be justified by a habituation effect in the neural responses. For the tested stimulus train durations, we found that N20 and P60 show a linear relationship with train duration, achieving R^2 value above 0.99. On the other hand, due the strong habituation effect in P35, the R^2 value for P35 is 0.93; however, this habituation effect is too strong to explain the habituation effect in the hemodynamic response.

The initial hemoglobin overshoot observed in human is not supported by the findings in small animals, for which the habituation effect in hemoglobin response appears for longer stimulus trains ($8 - 20$ s) [Martindale *et al.*, 2005, Franceschini *et al.*, 2008]. This discrepancy may be due to differences between species (humans vs. rats) or anticipation effects, which are suppressed by anesthesia in rats.

In addition to linearity/non-linearity of the responses, we performed a convolution analysis to test whether individual components of the current amplitude can predict the hemodynamic response time courses better than input stimulus. To do so, we used either a current amplitude component or the input stimulus as an input in the linear convolution model to predict the hemodynamic responses and determined statistically which factor shows better predictive power. With the stimulus tested, we found that there is no significant difference between the predictions of HbR using N20, P35 or P60 versus using the input stimulus. Instead, the sum of the peak amplitudes and the RMS value can predict HbR response consistently better than the input stimulus. Further-

more, subtracting N20 helps remove possible hyperpolarising DC shift of the membrane potential [Hellweg *et al.*, 1977], both P35−N20 and P60−N20 achieve statistically better prediction of HbR than the stimulus.

The fact that the sum of three peaks and the RMS value of the current dipole amplitude are the best and N20 is the worst in predicting the hemodynamic responses is in agreement with recent findings in rat neurovascular coupling using EEG and DOI in rats [Franceschini *et al.*, 2008]. Similar to the human study, the animal study shows that the later peaks (N1 and P2) of the somatosensory evoked potentials are able to predict hemodynamic responses better than the first peak, P1. The EEG evoked potential peak P1 in rats is equivalent to MEG N20 deflection in humans. This initial peak constitutes the primary cortical response which is generated by synaptic excitation of middle layers by thalamocortical inputs [Li *et al.*, 1956, Mitzdorf, 1985, Di *et al.*, 1990]. In most invasive animal studies, this earliest peak is typically used to predict the hemodynamic response [Caesar *et al.*, 2003, Jones *et al.*, 2004, Sheth *et al.*, 2004, Iadecola, 2004]. Historically, the greater focus on P1 may be partially due to the fact that this peak is very stable and persists with deep anesthesia, while subsequent cortical activity is abolished with deep anesthesia. The later cortical activity is more spread both spatially and temporally, and it is less stable with changes in stimulus parameters [Cauller & Kulics, 1991]. Our findings that the hemodynamic responses correlate better with later peaks than with the primary cortical response suggest that the cortical hemodynamic response is largely controlled by synaptic activity related to intra-cortical processing rather than direct thalamic input. This is probably due to the majority of the synapses are triggered with column inputs from other neurons in the cortex, and the hemodynamic response is mostly sensitive to the level of synaptic activity [Mathiesen *et al.*, 1998]. Thus, our results indicate that later components need to be considered to correctly understand the neurovascular coupling.

Despite the better predictions of hemoglobin response by the sum of two or three peaks and the RMS value of the current dipole amplitude, the habituation effects exhibited in the hemodynamic response are not totally accounted for with these parameters. This suggests that the hemodynamic response cannot be described by a single component, but by a weighted combination of multiple components or some later components in the neural signals. A more complex experimental design which provides more variability in the neuronal responses will be necessary to obtain sufficient power in the statistical analysis of the different neurovascular coupling models.

4.6 Summary

By combining DOI and MEG, we investigated neurovascular coupling non-invasively in human subjects using median-nerve stimulation. Previous fMRI studies have shown a habituation effect in the hemodynamic BOLD response for stimulation periods longer than 2 s. With DOI and MEG we can test whether this effect in hemodynamic response can be accounted for by a habituation effect in the neural response. Our experimental results show that the habituation effect in the hemodynamic response is stronger than that in the earliest cortical neural response (N20). Using a linear convolution model to predict hemodynamic responses, we found that including late neural components ($\geq$ 30 ms) improves the prediction of the hemoglobin response. This finding suggests that in addition to the initial evoked-response deflections related to the thalamic afferent input, later cortical activity is needed to predict the hemodynamic response.

Our results indicate that neural activity and hemodynamic responses are approximately aligned in space, motivating the joint EEG/MEG-fMRI analysis method presented in the next chapter.

Chapter 5

fMRI-Informed Regional EEG/MEG Source Estimation (FIRE)

Due to the limited spatial resolution of EEG/MEG data, it is challenging to distinguish sources at close by locations in the brain. In this chapter, we focus on incorporating fMRI information to improve the EEG/MEG source estimation accuracy. Our model is motivated by findings of the neurovascular coupling discussed in Chapter 4. The resulting estimator, fMRI-informed regional EEG/MEG source estimation (FIRE) takes advantage of the spatial alignment between the neural and the vascular activities, while allowing for substantial differences in their dynamics. The work presented in this chapter was published in [Ou *et al.*, 2009c]. In addition, we have submitted an extended version of this work to NeuroImage journal.

In the following, we first discuss the model underlying FIRE, the inference procedure, and the implementation details in Section 5.1. We then present the experimental comparisons between FIRE and prior methods for joint E/MEG-fMRI analysis using both simulated and human data in Section 5.2, followed by a discussion and summary.

■ 5.1 Method

We first present the model assumptions in FIRE by dissecting its graphical model shown in Fig. 5.1. We then discuss the priors, the parameter setting, and the inference procedure to estimate the current source distribution.

■ 5.1.1 Neurovascular Coupling and Data Models

We assume that the source space comprises N discrete locations on the cortex parcelled into K brain regions. We denote the set indexing the discrete locations in region k

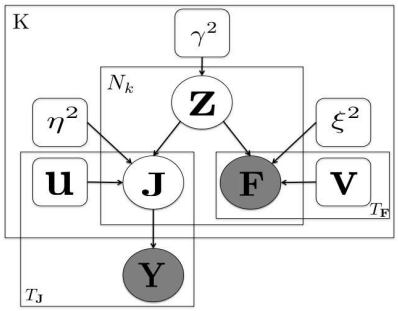


Figure 5.1. Graphical interpretation of FIRE. Circular nodes indicate random variables and square nodes indicate model parameters. The hidden activity z models the neurovascular coupling relationship. The hidden current source distribution $\mathbf{J}$ is measured by E/MEG, producing observation $\mathbf{Y}$. $\mathbf{F}$ denotes fMRI measurements. Vectors $\mathbf{u}$ and $\mathbf{v}$ are the unknown region-specific neural and vascular waveforms, respectively. The inner plate represents N_k vertices in region k; the outer plate represents K regions. The bottom left and right plates represent $T_\mathbf{J}$ and $T_\mathbf{F}$ time points in the neural and the vascular measurements, respectively.

by P_k and the cardinality of P_k by N_k. Hence, the outer plate in Fig. 5.1 describes K regions, and the inner plate captures N_k locations in region k.

The shapes of the source timecourses are identical within a region but may vary across regions. Specifically, we let $\mathbf{u}_k$ and $\mathbf{v}_k$ be the unknown waveforms in region k, associated with the neural and vascular activities, respectively. Examples $\mathbf{u}_k$ and $\mathbf{v}_k$ in two separate brain regions are illustrated as the black timecourses in Fig. 1.3. We model the strengths of neural and the vascular activities as being depended on a hidden vector variable $\mathbf{Z} = [z_1, z_2, \cdots, z_N]^\mathrm{T}$, where z_n indicates the activation strength at location n on the cortical surface. Thus, the probabilistic model for the neural activation timecourse $\mathbf{j}_n$ and the vascular activation timecourse $\mathbf{f}_n$ at location n in region k condition on the activation strength z_n can be expressed as

$$p\left(\mathbf{j}_n, \mathbf{f}_n | z_n; \mathbf{u}_k, \mathbf{v}_k, \eta_k^2, \xi_k^2\right) = p\left(\mathbf{j}_n | z_n; \mathbf{u}_k, \eta_k^2\right) p\left(\mathbf{f}_n | z_n; \mathbf{v}_k, \xi_k^2\right) \tag{5.1}$$

$$= \mathcal{N}\left(\mathbf{j}_n; z_n \mathbf{u}_k, \eta_k^2 \mathbf{I}\right) \mathcal{N}\left(\mathbf{f}_n; z_n \mathbf{v}_k, \xi_k^2 \mathbf{I}\right), \tag{5.2}$$

where η_k^2 and ξ_k^2 are noise variances. The blue and green timecourses in Fig. 1.3 represent examples of $\mathbf{j}_n$ and $\mathbf{f}_n$ for two locations within each region. Note that our neurovascular coupling model captures only the spatial alignment between the two types of activities; it does not impose temporal similarity of the signals. The neural timecourses $\mathbf{j}_n$ and the vascular signals $\mathbf{f}_n$ are conditionally independent given the hidden variable for brain

activity z_n. Although the parcellation is optimal when each parcel includes one type of neural and vascular waveform, our experimental results show that FIRE's performance is comparable to other methods when multiple waveforms are present within a parcel.

In the model description below, we construct all matrices such that rows represent locations or sensors and columns represent time points. Thus, we let $N \times T_{\mathbf{J}}$ matrix $\mathbf{J} = [\mathbf{j}_1, \mathbf{j}_2, \cdots, \mathbf{j}_N]^{\mathrm{T}}$ be the hidden neural current on the cortex for all $T_{\mathbf{J}}$ time points. We assume that the vascular signal $\mathbf{f}_n$ at location n is directly observable via fMRI. We let $N \times T_{\mathbf{F}}$ matrix $\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2, \cdots, \mathbf{f}_N]^{\mathrm{T}}$ be the fMRI measurements on the cortex at all $T_{\mathbf{F}}$ time points.

The neural currents $\mathbf{j}_n$ detected with E/MEG are characterized by the standard observation model. We let $M \times T_{\mathbf{J}}$ matrix $\mathbf{Y} = [\mathbf{y}(1), \mathbf{y}(2), \cdots, \mathbf{y}(T_{\mathbf{J}})]$ be the E/MEG measurements at all $T_{\mathbf{J}}$ time points. Column t of matrix $\mathbf{J}$, $\mathbf{j}(t)$, denotes the neural current distribution at time t. The quasi-static approximation of Maxwell's equations states that E/MEG signals at time t are instantaneous linear combinations of the currents at different locations in the source space:

$$\mathbf{y}(t) = \mathbf{A}\mathbf{j}(t) + \mathbf{e}(t) \qquad \forall\, t = 1, 2, \cdots, T_{\mathbf{J}}, \tag{5.3}$$

where $\mathbf{e}(t)$ is the measurement noise at time t. The $M \times N$ forward matrix $\mathbf{A}$ captures the electromagnetic properties of the head, the geometry of the sensors, and the locations of the sources. Similar to other source estimation methods, the forward matrix $\mathbf{A}$ is assumed to be known. We assume spatial whitening in the measurement (sensor) space so that $\mathbf{e}(t) \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. The number of sources N ($\sim 10^4$) is much larger than the number of measurements M ($\sim 10^2$), leading to an infinite number of solutions satisfying Eq. (5.3) even for $\mathbf{e}(t) = \mathbf{0}$. The plate at the bottom left corner of Fig. 5.1 corresponds to $T_{\mathbf{J}}$ temporal samples. In general, $\mathbf{j}_n$ should be modeled as three time-courses corresponding to the three Cartesian components of the current. However, due to the columnar organization of the cortex, we can further constrain the current orientation to be perpendicular to the cortical surface and consider a scalar current value at each location.

■ 5.1.2 Priors and Parameter Settings

To encourage the activation patterns to vary smoothly in space within a region, we impose a prior on the modulating variables $\mathbf{Z}$. Specifically, we define $\underline{\mathbf{z}}_k = \{z_n\}_{n \in P_k}$ and assume

$$p(\mathbf{Z}; \{\gamma_k\}_{k=1}^K) = \prod_{k=1}^K p(\underline{\mathbf{z}}_k; \gamma_k) = \prod_{k=1}^K \mathcal{N}\left(\mathbf{0}, \gamma_k^2 \mathbf{\Gamma}_k\right), \tag{5.4}$$

where the variance γ_k^2 indirectly models the strength of the activation magnitude z_n in region k, and $\mathbf{\Gamma}_k$ is a fixed matrix that acts as a regularizer by penalizing the sum of squared differences between neighboring locations. This spatial prior is particularly important for the brain regions where vascular activity is too weak to measure, but the neural activity can be detected by E/MEG. Without this prior, the estimated current source may have an unrealistic spatial distribution due to the ill-posed nature of the E/MEG inverse problem.

Our $\mathbf{\Gamma}_k$ is similar to the regularizer used in the Low Resolution Brain Electromagnetic Tomography (LORETA) [Pascual-Marqui *et al.*, 1994], except that we apply $\mathbf{\Gamma}_k$ to individual brain regions while LORETA's spatial regularizer is applied to the whole brain. We assume separate variance γ_k^2 for different brain regions since the strength of the currents is expected to vary significantly between regions with and without active sources. This choice is similar to the recent work in the application of ARD to E/MEG reconstruction [Sato *et al.*, 2004b, Wipf & Nagarajan, 2009], except that their work assumes independent variance γ^2 for each location in the brain.

Since the forward matrix $\mathbf{A}$ is underdetermined, the current distribution $\mathbf{J}$ produced by our neurovascular coupling model can fully explain the E/MEG data. In other words, without the noise term η_k^2 (i.e., $\mathbf{j}_n = z_n \mathbf{u}_k$), the fMRI data can exert too much influence on the reconstruction results. Although we can estimate the noise variance of the current source timecourses η_k^2 by extending the inference procedure, we find the corresponding estimate unstable without a prior. Based on the preliminary empirical testing, we fix $\eta_k^2 = 1$. With proper temporal whitening of the fMRI data, we can also assume that $\xi_k^2 = \eta_k^2$. Fixing $\eta_k^2 = \xi_k^2$ helps to significantly reduce the computational burden of the estimation.

To summarize, our model can be mathematically expressed as

$$p(\mathbf{F}, \mathbf{Y}, \mathbf{J}, \mathbf{Z}; \mathbf{\Theta}) = p(\mathbf{Y}|\mathbf{J}) p(\mathbf{J}, \mathbf{F}|\mathbf{Z}; \mathbf{\Theta}) p(\mathbf{Z}; \mathbf{\Theta}), \tag{5.5}$$

where $\mathbf{\Theta} = [\theta_1, \theta_2, \cdots, \theta_K]$ is the combined set of parameters, and $\theta_k = [\mathbf{u}_k, \mathbf{v}_k, \gamma_k^2]$ is the set of parameters for region k. $p(\mathbf{Y}|\mathbf{J})$ is the E/MEG data model in Eq. (5.3). As shown in Fig. 5.1, the E/MEG observation $\mathbf{Y}$ is conditionally independent of other variables

given the hidden source currents $\mathbf{J}$. $p(\mathbf{J},\mathbf{F}|\mathbf{Z};\boldsymbol{\Theta})$ is our neurovascular coupling model in Eq. (5.1), and $p(\mathbf{Z};\boldsymbol{\Theta})$ is the prior on $\mathbf{Z}$ in Eq. (5.4). Combining these elements of the joint likelihood model, we obtain

$$\begin{aligned}
\log p(\mathbf{Y},\mathbf{J},\mathbf{F},\mathbf{Z};\boldsymbol{\Theta}) &= \log p(\mathbf{Y}|\mathbf{J}) + \log p(\mathbf{J},\mathbf{F}|\mathbf{Z};\boldsymbol{\Theta}) + \log p(\mathbf{Z};\boldsymbol{\Theta}) \qquad (5.6)\\
&= \sum_{t=1}^{T_\mathbf{J}} \log \mathcal{N}\left(\mathbf{y}(t);\mathbf{A}\mathbf{j}(t),\mathbf{I}\right)\\
&\quad + \sum_{k=1}^{K}\sum_{n=1}^{N_k}\Big[\log \mathcal{N}\left(\mathbf{j}_n;z_n\mathbf{u}_k,\mathbf{I}\right) + \log \mathcal{N}\left(\mathbf{f}_n;z_n\mathbf{v}_k,\mathbf{I}\right)\Big]\\
&\quad + \sum_{k=1}^{K} \log \mathcal{N}\left(\underline{\mathbf{z}}_k;\mathbf{0},\gamma_k^2\boldsymbol{\Gamma}_k\right)\\
&= -\frac{1}{2}\left[MT_\mathbf{J}\log\left(2\pi\right) + \sum_{t=1}^{T_\mathbf{J}}\|\mathbf{y}(t)-\mathbf{A}\mathbf{j}(t)\|^2\right]\\
&\quad - \frac{1}{2}\sum_{k=1}^{K}\Big[N_kT_\mathbf{J}\log(2\pi) + \sum_{n\in P_k}\|\mathbf{f}_n - z_n\mathbf{v}_k\|^2\\
&\quad + N_kT_\mathbf{F}\log(2\pi) + \sum_{n\in P_k}\|\mathbf{j}_n - z_n\mathbf{u}_k\|^2\Big]\\
&\quad - \frac{1}{2}\sum_{k=1}^{K}\left[N_k\log(2\pi) + N_k\log(\gamma_k^2) + \log(\det(\boldsymbol{\Gamma}_k)) + \frac{\underline{\mathbf{z}}_k^\mathrm{T}\boldsymbol{\Gamma}_k^{-1}\underline{\mathbf{z}}_k}{\gamma_k^2}\right]
\end{aligned}$$

where $\det(\cdot)$ denotes matrix determinant. Note that although the latent activation strength $\mathbf{Z}$ is independent across regions *a priori*, the posterior estimates for $\mathbf{J}$, $\mathbf{Z}$, and Θ are spatially dependent due to the measurements $\mathbf{Y}$.

■ 5.1.3 Inference

Our goal is to estimate the current source $\mathbf{J}$ and the timecourses $\mathbf{u}$ and $\mathbf{v}$. A standard inference procedure is to compute the maximum likelihood (ML) estimate of $\boldsymbol{\Theta}$ while jointly considering the current source distribution $\mathbf{J}$ and the activation strength $\mathbf{Z}$ as hidden variables, followed by a maximum *a posteriori* (MAP) estimation of $\mathbf{J}$. However, this leads to a computationally intractable algorithm, as we discuss in Section 5.3. Instead, we alternate the optimization between estimating $\mathbf{J}$ and estimating $\boldsymbol{\Theta}$. While estimating $\boldsymbol{\Theta}$, we treat the activation strength $\mathbf{Z}$ as an auxiliary variable, and

marginalize it out in the analysis. Our inference procedure can be thus formulated as

$$\{\mathbf{J}^*,\ \boldsymbol{\Theta}^*\} = \arg\max_{\mathbf{J},\ \boldsymbol{\Theta}} \log p(\mathbf{F},\mathbf{Y},\mathbf{J};\boldsymbol{\Theta}) \tag{5.7}$$

$$= \arg\max_{\mathbf{J},\ \boldsymbol{\Theta}} \log \int_{\mathbf{Z}} p(\mathbf{Y}|\mathbf{J})p(\mathbf{F},\mathbf{J}|\mathbf{Z};\boldsymbol{\Theta})p(\mathbf{Z};\boldsymbol{\Theta})d\mathbf{Z} \tag{5.8}$$

$$= \arg\max_{\mathbf{J},\ \boldsymbol{\Theta}} \log p(\mathbf{Y}|\mathbf{J}) + \log p(\mathbf{F},\mathbf{J};\boldsymbol{\Theta}). \tag{5.9}$$

In Eq. (5.9), $p(\mathbf{F},\mathbf{J};\boldsymbol{\Theta})$ acts as the prior for $\mathbf{J}$. Since both $\mathbf{J}$ and $\mathbf{F}$ are linear functions of $\mathbf{Z}$, $p(\mathbf{F},\mathbf{J};\boldsymbol{\Theta})$ is a continuous Gaussian mixture model.

The difficulty in estimating the proposed model from the data is caused by the interactions between space and time variables, as reflected by the intersection of the temporal plates and the spatial plates in Fig. 5.1. It is easy to see from Eq. (5.3) that the output of a given E/MEG sensor is a mixture of signals from the entire source space. Moreover, $\mathbf{F}$, $\mathbf{J}$, and $\mathbf{Y}$ are jointly Gaussian. The correlation between different time points (i.e., between two E/MEG time points, between two fMRI time points, and between E/MEG and fMRI time points) is generally not zero. Hence, the inference must be performed for all time points and all locations simultaneously. FIRE is thus substantially more computationally demanding than the standard temporally independent E/MEG estimation or voxel-wise fMRI analysis that ignore these dependencies in modeling the observed signals. The benefit of this increased computational burden is more accurate inference across time points.

Due to the special structure of our model, we can derive an efficient gradient descent method with two alternating steps. In the first step, we fix $\boldsymbol{\Theta}$ and derive a closed-form solution for $\mathbf{J}$. In the second step, we fix $\mathbf{J}$ and show that $\boldsymbol{\Theta}$ can be efficiently estimated via the EM algorithm [Dempster *et al.*, 1977].

For a fixed $\boldsymbol{\Theta} = \widehat{\boldsymbol{\Theta}}$, $p(\mathbf{F},\mathbf{Y},\mathbf{J};\widehat{\boldsymbol{\Theta}})$ is a jointly-Gaussian distribution. As shown in Appendix A, the estimate of $\mathbf{J}$ is therefore equal to its conditional mean:

$$\mathrm{vec}\left(\widehat{\mathbf{J}}\right) = \arg\max_{\mathbf{J}} \log p(\mathbf{F},\mathbf{Y},\mathbf{J};\widehat{\boldsymbol{\Theta}}) = E\left[\mathbf{J}|\mathbf{F},\mathbf{Y};\widehat{\boldsymbol{\Theta}}\right] = \boldsymbol{\Gamma}_{\mathbf{W},\mathbf{J}}^{\mathrm{T}}\boldsymbol{\Gamma}_{\mathbf{W}}^{-1}\mathbf{W}, \tag{5.10}$$

where $\mathbf{W}^{\mathrm{T}} = \left[\left(\mathrm{vec}\left(\mathbf{F}\right)\right)^{\mathrm{T}}\left(\mathrm{vec}\left(\mathbf{Y}\right)\right)^{\mathrm{T}}\right]$ includes both E/MEG and fMRI measurements. Operator $\mathrm{vec}\left(\cdot\right)$ concatenates the columns of a matrix into a vector. $\boldsymbol{\Gamma}_{\mathbf{W}}$ is the covariance matrix of $\mathbf{W}$, and $\boldsymbol{\Gamma}_{\mathbf{W},\mathbf{J}}$ is the cross-covariance matrix between $\mathbf{W}$ and $\mathrm{vec}\left(\mathbf{J}\right)$. Appendix A presents the detailed derivations for $\boldsymbol{\Gamma}_{\mathbf{W}}$ and $\boldsymbol{\Gamma}_{\mathbf{W},\mathbf{J}}$. Eq. (5.10) implies that E/MEG and fMRI measurements jointly determine the estimate of the neural activity.

This equation is similar to the standard MNE solution [Hämäläinen *et al.*, 1993], but also includes the correlation between the observations $\mathbf{Y}$ and $\mathbf{F}$ and the correlation among different time points of the neural current $\mathbf{J}$.

For a fixed $\mathbf{J} = \widehat{\mathbf{J}}$, we estimate the parameters $\mathbf{\Theta}$:

$$\widehat{\mathbf{\Theta}} = \arg\max_{\mathbf{\Theta}} \log p(\mathbf{F}, \widehat{\mathbf{J}}; \mathbf{\Theta}) = \arg\max_{\mathbf{\Theta}} \sum_{k=1}^{K} \log p(\{\mathbf{f}_n, \widehat{\mathbf{j}}_n\}_{n\in P_k}; \theta_k). \tag{5.11}$$

It is easy to see that this optimization can be done for each region separately:

$$\widehat{\theta}_k = \arg\max_{\theta_k} \log p(\{\mathbf{f}_n, \widehat{\mathbf{j}}_n\}_{n\in P_k}; \theta_k) \quad \forall\ k = 1, \cdots, K. \tag{5.12}$$

As can be seen in Fig. 5.1, when the current distribution $\mathbf{J}$ is fixed, the E/MEG measurement $\mathbf{Y}$ does not provide additional information for the parameter estimation. Furthermore, each set of parameters θ_k can be efficiently estimated using the EM algorithm [Dempster *et al.*, 1977] by re-introducing the latent variable $\underline{\mathbf{z}}_k$ that describes activation strength of vertices within region k. This method can be thought of as an extension of the EM algorithm for probabilistic PCA [Tipping & Bishop, 1999] to two sets of data [Bach & Jordan, 2005].

Specifically, the parameter estimates $\widehat{\theta}_k$ for region k can be obtained by optimizing the lower bound of the log-probability:

$$\log p\Big(\{\mathbf{f}_n, \widehat{\mathbf{j}}_n\}_{n\in P_k}; \theta_k\Big) \geq \int_{\underline{\mathbf{z}}_k} q(\underline{\mathbf{z}}_k) \log p\Big(\{\mathbf{f}_n, \widehat{\mathbf{j}}_n\}_{n\in P_k}, \underline{\mathbf{z}}_k; \theta_k\Big) d\underline{\mathbf{z}}_k - \int_{\underline{\mathbf{z}}_k} q(\underline{\mathbf{z}}_k) \log p(q(\underline{\mathbf{z}}_k)) d\underline{\mathbf{z}}_k, \tag{5.13}$$

where $q(\underline{\mathbf{z}}_k) = p\left(\underline{\mathbf{z}}_k | \{\mathbf{f}_n, \widehat{\mathbf{j}}_n\}_{n\in P_k}; \widetilde{\theta}_k\right)$ is the posterior probability computed in the E-step, and $\widetilde{\theta}_k$ is the estimate from the last EM iteration. Since $\{\mathbf{f}_n, \widehat{\mathbf{j}}_n\}_{n\in P_k}$ and $\underline{\mathbf{z}}_k$ are jointly-Gaussian, $q(\underline{\mathbf{z}}_k)$ is also a Gaussian distribution, and the M-step update depends only on the first- and the second-order statistics of $\underline{\mathbf{z}}_k$. Due to this special structure, we first derive the M-step, followed by the E-step. To simplify notation, we use $\langle\cdot\rangle_q$ to denote the expectation with respect to the posterior distribution $q(\underline{\mathbf{z}}_k)$, i.e., $\langle\cdot\rangle_q \triangleq E\left[\cdot | \{\mathbf{f}_n, \widehat{\mathbf{j}}_n\}_{n\in P_k}; \widetilde{\theta}_k\right]$.

In the M-step, we fix $q(\underline{\mathbf{z}}_k)$ and optimize the right-hand side of Eq. (5.13), and get

$$\widehat{\theta}_k = \arg\max_{\theta_k} \int_{\underline{\mathbf{z}}_k} q(\underline{\mathbf{z}}_k) \log p(\{\mathbf{f}_n, \widehat{\mathbf{j}}_n\}_{n\in P_k}, z_n; \theta_k) d\underline{\mathbf{z}}_k \tag{5.14}$$

$$= \arg\max_{\theta_k} N_k \left(\mathbf{v}_k^{\mathrm{T}} \mathbf{f}_n + \mathbf{u}_k^{\mathrm{T}} \widehat{\mathbf{j}}_n\right) \langle z_n \rangle_q - \frac{N_k}{2} \left(\mathbf{v}_k^{\mathrm{T}} \mathbf{v}_k + \mathbf{u}_k^{\mathrm{T}} \mathbf{u}_k\right) \langle z_n^2 \rangle_q - \frac{1}{2} \frac{\langle \underline{\mathbf{z}}_k^{\mathrm{T}} \mathbf{\Gamma}_k^{-1} \underline{\mathbf{z}}_k \rangle_q}{\gamma_k^2} - N_k \log \gamma_k^2. \tag{5.15}$$

The detailed derivations of Eq. (5.15) are shown in Appendix B. Setting the derivatives of Eq. (5.15) with respect to the model parameter vector $\theta_k = [\mathbf{u}_k, \mathbf{v}_k, \gamma_k^2]$ to zero, we arrive at the update rules:

$$\widehat{\mathbf{u}}_k \leftarrow \frac{\sum_{n\in P_k}\langle z_n\rangle_q \widehat{\mathbf{j}}_n}{\mathrm{tr}(\langle \underline{\mathbf{z}}_k\underline{\mathbf{z}}_k^{\mathrm{T}}\rangle_q)}\ ,\ \widehat{\mathbf{v}}_k \leftarrow \frac{\sum_{n\in P_k}\langle z_n\rangle_q \mathbf{f}_n}{\mathrm{tr}(\langle \underline{\mathbf{z}}_k\underline{\mathbf{z}}_k^{\mathrm{T}}\rangle_q)}\ ,\ \widehat{\gamma_k^2} \leftarrow \frac{\langle \underline{\mathbf{z}}_k^{\mathrm{T}}\boldsymbol{\Gamma}_k^{-1}\underline{\mathbf{z}}_k\rangle_q}{N_k}. \tag{5.16}$$

Since the M-step depends only on quantities $\langle \underline{\mathbf{z}}_k\underline{\mathbf{z}}_k^{\mathrm{T}}\rangle_q$, $\langle \underline{\mathbf{z}}_k\rangle_q$, and $\langle \underline{\mathbf{z}}_k^{\mathrm{T}}\boldsymbol{\Gamma}_k^{-1}\underline{\mathbf{z}}_k\rangle_q$, we only need to evaluate those quantities in the E-step:

$$\langle \underline{\mathbf{z}}_k\underline{\mathbf{z}}_k^{\mathrm{T}}\rangle_q \leftarrow \left[\frac{1}{\gamma_k^2}\boldsymbol{\Gamma}_k^{-1} + \left(\mathbf{v}_k^{\mathrm{T}}\mathbf{v}_k + \mathbf{u}_k^{\mathrm{T}}\mathbf{u}_k\right)\mathbf{I}\right]^{-1} \tag{5.17}$$

$$\langle \underline{\mathbf{z}}_k\rangle_q \leftarrow \langle \underline{\mathbf{z}}_k\underline{\mathbf{z}}_k^{\mathrm{T}}\rangle_q \left[\left(\mathbf{v}_k^{\mathrm{T}}\mathbf{f}_1 + \mathbf{u}_k^{\mathrm{T}}\widehat{\mathbf{j}}_1\right), \cdots, \left(\mathbf{v}_k^{\mathrm{T}}\mathbf{f}_{N_k} + \mathbf{u}_k^{\mathrm{T}}\widehat{\mathbf{j}}_{N_k}\right)\right]^{\mathrm{T}} \tag{5.18}$$

$$\langle \underline{\mathbf{z}}_k^{\mathrm{T}}\boldsymbol{\Gamma}_k^{-1}\underline{\mathbf{z}}_k\rangle_q \leftarrow \langle \underline{\mathbf{z}}_k\rangle_q^{\mathrm{T}}\boldsymbol{\Gamma}_k^{-1}\langle \underline{\mathbf{z}}_k\rangle_q + \mathrm{tr}\left(\boldsymbol{\Gamma}_k^{-1}\langle \underline{\mathbf{z}}_k\underline{\mathbf{z}}_k^{\mathrm{T}}\rangle_q\right). \tag{5.19}$$

We iterate the EM algorithm until convergence which usually takes less than ten iterations. We then re-estimate $\mathbf{J}$ according to Eq. (5.10).

To summarize, the FIRE inference algorithm proceeds as follows:

(i) Initialize $\widehat{\mathbf{J}}$ as the MNE estimate: $\mathbf{J}^{(\mathrm{MNE})} = \mathbf{A}^{\mathrm{T}}\left(\mathbf{A}\mathbf{A}^{\mathrm{T}} + \lambda^2\mathbf{I}\right)^{-1}\mathbf{Y}$, where λ^2 is the regularization parameter related to the SNR of the data.

(ii) Repeat until convergence:

1. Compute $\widehat{\boldsymbol{\Theta}}$ using the EM algorithm:
   (a) E-step: construct $\{q(\underline{\mathbf{z}}_k)\}_{k=1}^K$ (Eq. (5.17-5.19))
   (b) M-step: estimate parameters $\boldsymbol{\Theta}$ (Eq. (5.16))
2. Update $\widehat{\mathbf{J}}$ for $\boldsymbol{\Theta} = \widehat{\boldsymbol{\Theta}}$ (Eq. (5.10)).

We also examine FIRE with different initializations. In particular, we use the fMNE estimate to initialize the algorithm and refer to this method as fFIRE. The fMNE estimate can be expressed as $\mathbf{J}^{(\mathrm{fMNE})} = \mathbf{R}\mathbf{A}^{\mathrm{T}}\left(\mathbf{A}\mathbf{R}\mathbf{A}^{\mathrm{T}} + \lambda^2\mathbf{I}\right)^{-1}\mathbf{Y}$, where $\mathbf{R}$ is a diagonal matrix of size N whose values depend on the thresholded fMRI-SPMs of the corresponding locations. A standard choice, as proposed in [Liu *et al.*, 1998], is 1 for locations with fMRI activation above a preselected threshold and 0.1 for those below the threshold.

5.2 Results

We first compare FIRE and fFIRE to MNE, fMNE, and fARD using simulated data, including three scenarios closely related to those typically observed in human experiments. We employ a Monte Carlo procedure to estimate performance statistics for each method. We then proceed to a comparison of the methods using human E/MEG-fMRI data from a somatosensory study and an attention-shift auditory study.

5.2.1 Simulation studies

To simulate MEG measurements, we created two patches on the cortical surface, with current source orientation along the outward normal to the cortical surface. Shown in the lateral-occipital view of the right hemisphere (Fig. 5.2), Patch A contains 20 vertices and is located in the inferior parietal region. Patch B contains 32 vertices and is located in the superior parietal region. The selection of the source patches is independent of the anatomical parcellation used in the source estimation. The anatomical parcels are used in the inference only. We simulated neural and vascular timecourses in these two patches for three different scenarios: no silent activity, silent vascular activity, and silent neural activity. In the two cases with silent activities, we kept the activity of patch B unchanged while silencing neural or vascular activity in patch A. The simulated neural signals are shown as solid black lines in the rightmost column of Fig. 5.2. The activation maps corresponding to the peaks of the two simulated neural signals are shown in the first column.

For the forward calculations, we employed the sensor configuration of the 306-channel Neuromag VectorView MEG system used in our human studies and added Gaussian noise to the signals. The resulting signals have a SNR of 3 dB, within the typical SNR range of real MEG data. Since the two patches are close to each other in the highly folded cortex and they exhibit neural activity during overlapping time intervals, it is particularly difficult to obtain accurate current source estimates for this configuration.

Columns two to five in Fig. 5.2 depict the current source estimates $\mathbf{J}^*$ obtained via different methods for the two time points corresponding to the peaks of activity. Following [Liu *et al.*, 1998], the fMNE weighting parameters are set to 1 and 0.1 for active and inactive fMRI locations, respectively. The hyper-parameters for fARD are selected according to [Sato *et al.*, 2004b]. The results from FIRE and fFIRE are quite similar in

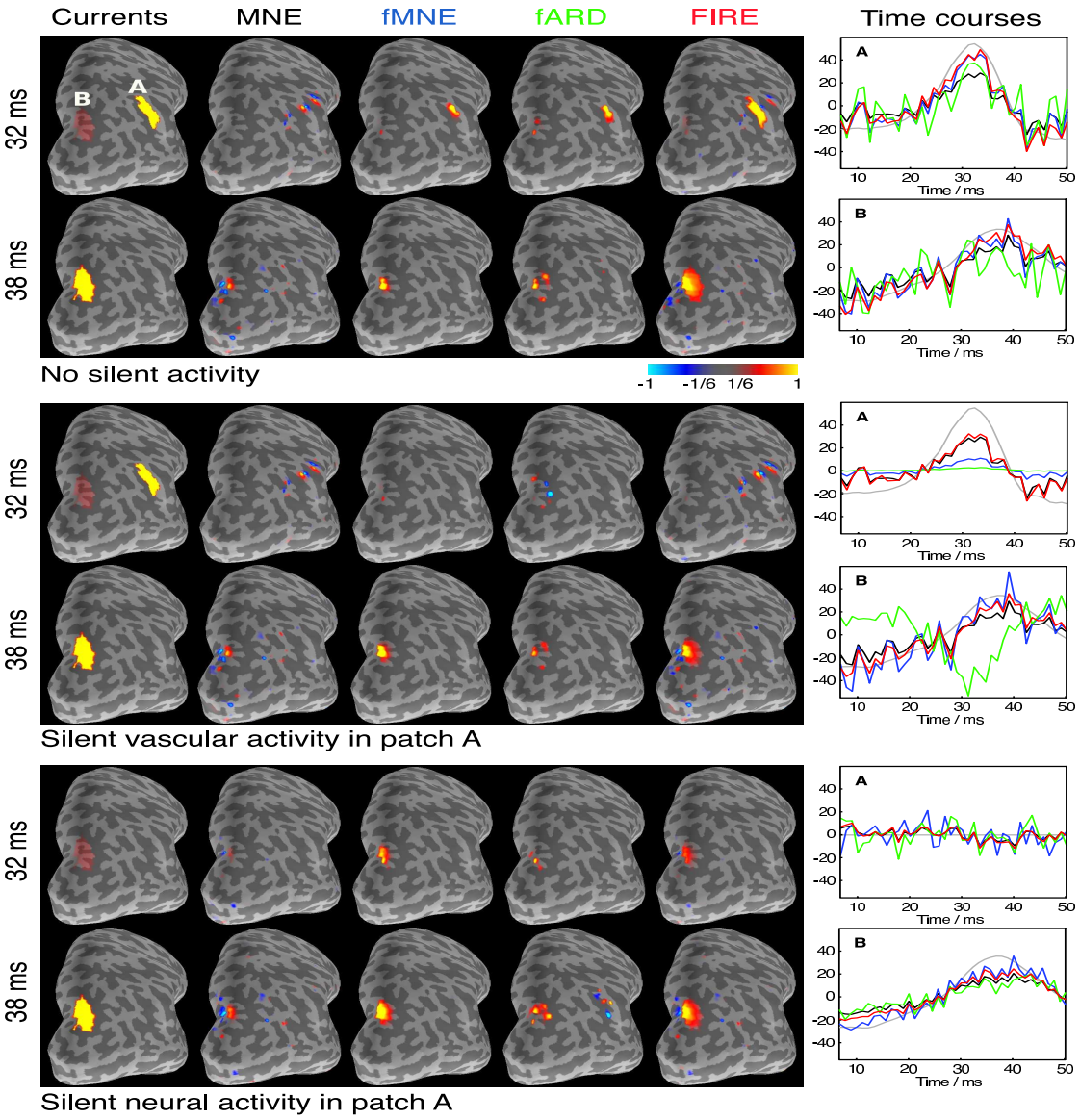


Figure 5.2. Current source estimates in three scenarios. Lateral-occipital view of the right hemisphere is shown. Patch A and patch B are highlighted in the top left panel; the rest of the figures follow the same convention. (Top) Neither neural nor vascular activity is silent. (Middle) Vascular activity in patch A is silent. (Bottom) Neural activity in patch A is silent. The first column illustrates the simulated current distributions with a selected threshold at the peak activations. The next four columns show the estimates from MNE, fMNE, fARD, and FIRE, respectively, for the time of peak activation for each patch. Hot/cold colors correspond to outward/inward current flow. The rightmost column shows the simulated (black solid) and the estimated (dashed) timecourses from the most active vertices in patches A and B. The color of the time course matches with those used for the name of the corresponding method.

this simulation setting. We therefore defer the evaluation of fFIRE to the Monte Carlo procedure presented later this session. Since the estimates from different methods are not directly comparable in amplitude, the threshold for each method is chosen to be 1/6 of the maximum absolute value of the corresponding current source estimates. The rightmost column in Fig. 5.2 presents the estimated timecourses (dashed) of the most active vertex, in terms of energy, in both patches.

No silent activity. As shown in Fig. 5.2(top), the MNE estimates extend across adjacent gyri. fMNE, fARD, and FIRE correctly localize the two patches at the peak activation, but FIRE provides a better estimate of the spatial extent of the activations. The fARD estimate is unstable, as reflected by the large fluctuations in the estimated timecourses, especially in patch B (Fig. 5.2(top), rightmost column, green).

Silent vascular activity. When the vascular activity in patch A is silent, fMNE shows excessive bias towards patch B. Without a large weight, the amplitude of the estimated timecourses (Fig. 5.2(middle), rightmost column, blue) in patch A is significantly lower than the corresponding estimates in patch B. It would be therefore easy to miss neural activation in patch A when interpreting the results (column three in Fig. 5.2(middle)). In contrast, by combining neural and vascular information in the re-weighted scheme, FIRE avoids such a bias. Its estimate in patch A (column five) is similar to that obtained via MNE (column one). Since the weight for patch B increases and the weight for patch A decreases during the fARD updates, the estimate in patch B explains the activation in patch A. As shown in the timecourse panel, the estimated timecourse in patch B (green) is similar in shape to the simulated timecourse in patch A (black solid). The change of sign is due to the fact that the outward normals for patch A and patch B are in approximately opposite directions.

Silent neural activity. As shown in Fig. 5.2(bottom), all methods correctly localize the neural activity in patch B, except for the small false positive in patch A for fARD. By assigning identical weights to patches A and B, fMNE estimates a timecourse for patch A (blue) that is noisier than the corresponding one produced by FIRE (red). FIRE suppresses the weights for patch A since the current estimates in that patch are close to zero; its results are closer to the simulations.

Monte Carlo simulation. We repeated the above experiments 100 times for each of the three scenarios. For each run, the locations of the simulated patches were randomly selected on the right hemisphere. Due to their spatial extent, the selected patches are likely to span portions of multiple anatomical parcels obtained from FreeSurfer: a

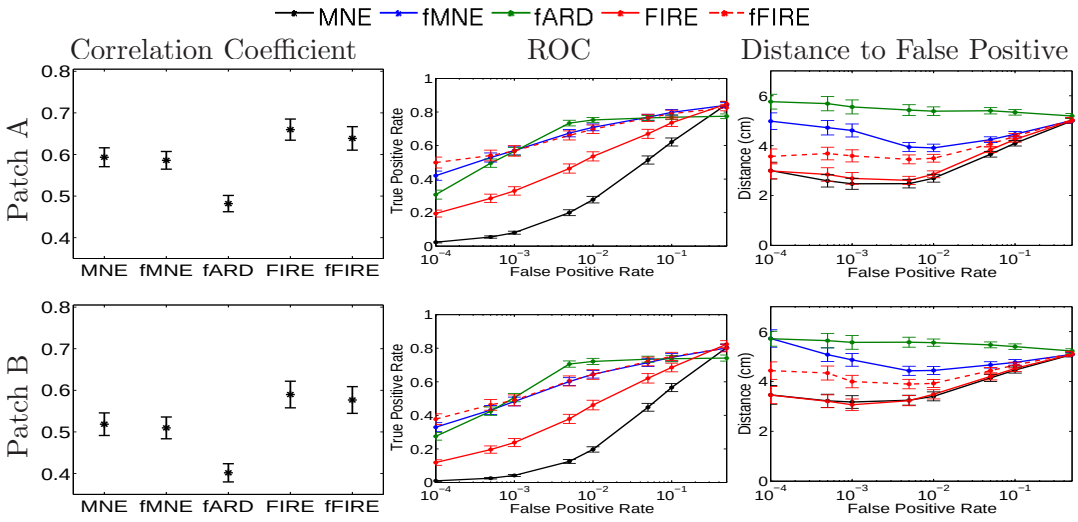
patch on average spans 3.5 anatomical parcels. Furthermore, in 30% of the trials in the simulation, the two selected patches cover the same anatomical parcel.

When comparing different estimation methods, we evaluated both the temporal and the spatial properties of the results. We used the correlation coefficient between the estimated timecourses and the ground truth ones, in the two patches separately, to evaluate the ability of the methods to reproduce the timecourses of the activity (Fig. 5.3, left column). To compare spatial accuracy of the methods, we computed the receiver operating characteristic (ROC) curve (Fig. 5.3, middle column) and the average distance between the simulated patches and the falsely detected locations (Fig. 5.3, right column). To compute the ROC curve, we selected the current estimates $\mathbf{J}^*$ at two time points corresponding to the peak activation. For each time point, we then separately varied the threshold and compared it with the ground truth to compute the true positive and false positive rates. To compute the distance to false positives, we varied the false positive rate, and computed the average distance between the falsely detected vertices and the ground truth activation patches.

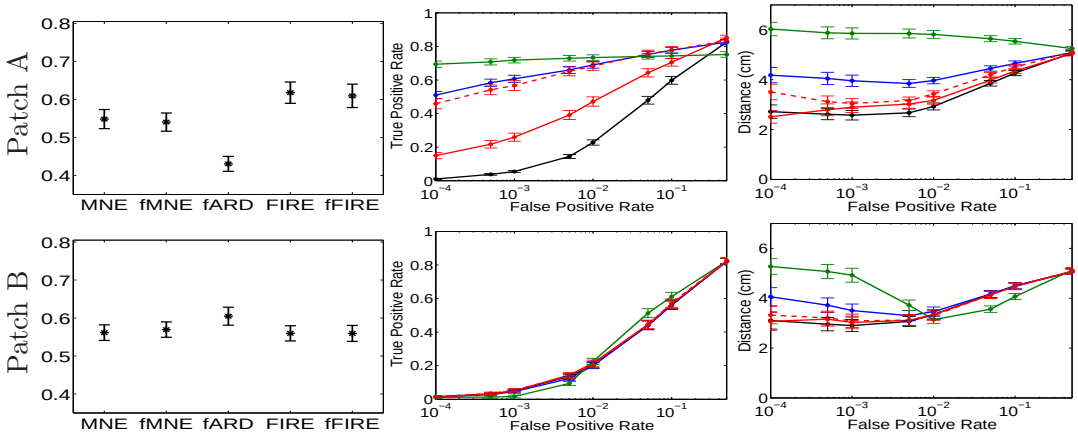
We first focus on the temporal correlation (Fig. 5.3, left column). For the three scenarios, FIRE and fFIRE achieve the highest temporal correlation (approximately 0.65), for patches exhibiting both neural and vascular activities (rows 1–3 and 5). The combination of static fMRI-SPM and the shrinkage prior in the ARD framework causes unstable timecourse estimates, reflected in low temporal correlation for fARD. When the patch exhibits neural activity, but no vascular activity, the temporal correlations are similar across all source estimation algorithms we examined (approximately 0.55).

The ROC curves (Fig. 5.3, middle column) demonstrate that fARD, fMNE, and fFIRE achieve similar detection accuracy for patches exhibiting both neural and vascular activities. When a patch shows neural activity, but no vascular activity, all algorithms have similar detection accuracy. As shown in the right column of Fig. 5.3, the falsely detected locations obtained from MNE, FIRE, and fFIRE tend to be close to the ground truth patches. In contrast, the falsely detected locations for fARD are relatively far away from the simulated patches, 5 to 6 cm on average. The standard error decreases as the false positive rate increases as there are more false positives involved in the computation of the average distance. Among all algorithms that we examined, fFIRE provides the best trade-off between the spatial and the temporal accuracy.

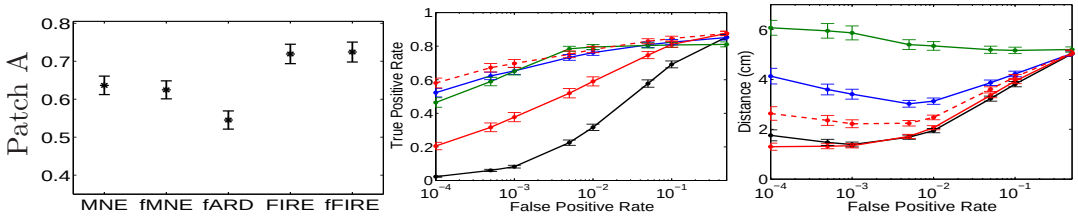
We analyzed separately the performance of the 30% of the trials where the two source patches are located within the same anatomical parcel. Since the two activations



(a) No silent activity



(b) Silent vascular activity in patch A



(c) Silent neural activity in patch A

Figure 5.3. Performance statistics in three scenarios. Left: the correlation coefficient between the estimated timecourses and the ground truth ones in patch A (top) and patch B (bottom). Middle: the ROC curves evaluated at the peak activation of the two patches. Right: the average distance from the simulated patches to the falsely detected locations.

are close to each other in space, the current estimation is particularly challenging in this case. We see significant reduction in performance across all methods, and their performance becomes more similar: the temporal correlation coefficients are 0.16 for MNE and fMNE, and 0.18 for FIRE and fFIRE; at false positive rate 0.005, the true positive rates are 0.20 for fMNE and fFIRE, 0.13 for FIRE, and 0.06 for MNE. We observe that the results for FIRE and fFIRE are quite robust with respect to the choice of anatomical parcellations. Although FIRE and fFIRE use a less-than-optimal parcellation in these 30% trials, the performance is equivalent to that of MNE and fMNE.

■ 5.2.2 Human Experiments

We also tested the method using human experimental data. E/MEG and fMRI measurements were obtained in separate sessions. The MEG data were acquired simultaneously using a 306-channel Neuromag VectorView MEG system; the EEG data were acquired with a 70-channel MEG-compatible EEG system. A 200 ms baseline before the stimulus was used to estimate the noise covariance matrix of the MEG sensors and EEG electrodes. fMRI data were obtained with a 3T Siemens TimTrio scanner. Anatomical images, from a 3T scanner, were used to construct the source space and the forward model. Informed consent in accordance with the Massachusetts General Hospital ethical committee was obtained from subjects prior to participation.

Median-nerve experiments

The median nerve at the right wrist was stimulated according to an event-related protocol, with a random inter-stimulus-interval ranging from 3 to 14 seconds. This stimulus activates a complex cortical network [Hari & Forss, 1999], including the contralateral primary somatosensory cortex (cSI) and bilateral secondary somatosensory cortices (cSII and iSII).

An average MEG signal, computed from approximately 100 trials, was used as the input to each method. In this experiment, EEG data were not acquired. The fMRI images were acquired using a Siemens 3T scanner (TR=1.5 s, 64×64×24, 3×3×6 mm^3, single channel head coil).

In the leftmost column in Fig. 5.4, the approximate locations for cSI (solid), cSII (dashed), and iSII (dashed) are highlighted on the fMRI activation maps ($p \leq 0.005$ uncorrected). Given the expected activations, we partitioned the contralateral activa-

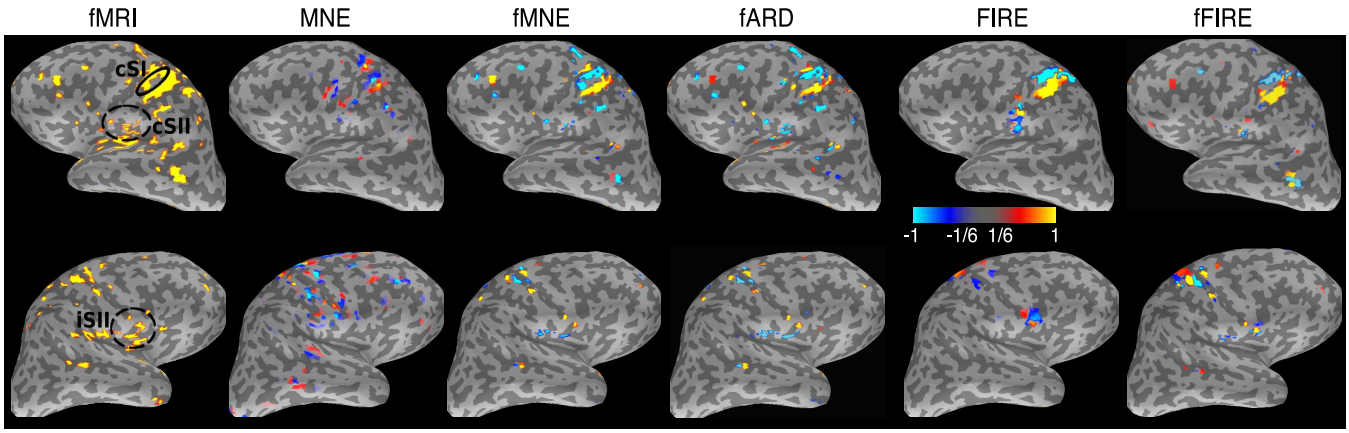


Figure 5.4. Human median-nerve experiments. In the first column, approximate locations for cSI (solid), cSII (dashed), and iSII (dashed) are highlighted on the fMRI activation maps ($p \leq 0.005$). Columns two to five show the current estimates obtained via MNE, fMNE, fARD, FIRE, and fFIRE respectively, at 75 ms after the stimulus onset. Hot/cold colors indicate outward/inward current flow.

tion into two regions, separately covering cSI and cSII. Note that in the noisy SPM, the sites of fMRI activations do not exactly agree with the locations of the expected current sources.

Columns two to five in Fig. 5.4 present the estimates at 75 ms after stimulus onset. At this time, cSI, cSII, and iSII should be activated. The threshold was set separately for each hemisphere since the activation in iSII is much weaker than that in cSI and cSII. For each method, the threshold is set to be 1/6 of the maximum absolute value of the corresponding current estimates. MNE produces a diffuse estimate, including physiologically unlikely activations at the gyrus anterior to the cSI area. In contrast, FIRE and fFIRE pinpoint cSI to the post-central gyrus. With the prior knowledge from fMRI, the detected cSII and iSII activations using fMNE, fARD, FIRE, and fFIRE are within the expected areas. The fMNE and fARD show stronger weighting toward the fMRI, reflected by the activations in the temporal lobes. fFIRE further detects activation in the visual area, the middle temporal area (MT) of the left hemisphere. This false detection is primarily caused by the strong activation in this area present in the initialization for fFIRE. Due to the highly folded nature of the cortex and uncertainties in MRI-fMRI registration, fMRI cannot distinguish between the walls of the central sulcus and the post-central sulcus, causing both walls to show strong vascular activity after mapping of the fMRI volume onto the cortex. Hence, fMNE, fARD, FIRE, and

fFIRE estimates extend to both sulcal walls.

Attention-Shift Auditory Experiments

An auditory attention task was utilized to investigate activations elicited by occasional attention shifting cues during dichotic stimulation. These activations were presumed to spread from the primary auditory cortex (Heschl's gyrus, HG) to surrounding association areas within the superior temporal plane (superior temporal gyrus, STG; planum temporale, PT) [Ahveninen *et al.*, 2006, Hart *et al.*, 2002, Rauschecker, 1998] and the superior temporal sulcus (STS) [Altmann *et al.*, 2008, Lu *et al.*, 1992] [Williamson *et al.*, 1991] before extending to higher-order parieto-frontal areas associated with attention shifting. Functional characterization of different subregions of the auditory cortex has been difficult in humans. In this experiment, we focus on the performance of each source estimation method in characterizing different activation patterns in the auditory cortex.

Three subjects were recruited for this study, and their task was to press a button upon hearing a target stimulus (quarter-tone or semitone deviants among standard tones) in the designated ear and to ignore sounds in the opposite ear. The stimulus of interest was an occasional "novel" buzzer sound that instructed the subject to shift attention to the cued ear. During E/MEG acquisition, the cue was presented after every 30 seconds. During fMRI acquisition, attention shifting cues were presented between clustered EPI acquisitions, after every other TR.

An average MEG signal, computed from approximately 40 trials, was used as the input to each method. In a separate session, sparse-sampling [Hall *et al.*, 1999] auditory fMRI data was acquired with a block design (3T Siemens TimTrio, TR=11.7 s, TE=30 ms; 48 axial slices 2.25 mm thick, 0.75 mm gap, $3\times3\,\mathrm{mm}^2$ in-plane). Each run was composed of three blocks, and each block consisted of two active stimulation periods (11.7 s each) interleaved with one silent baseline period (11.7 s).

The top panel in Fig. 5.5 shows the fMRI activation maps ($p \leq 0.0005$ uncorrected), with approximate locations for HG and STS areas highlighted. In this noisy SPM, some strong vascular activity appears in unexpected locations.

The bottom panel of Fig. 5.5 presents the current source estimates at 92, 125, and 225 ms after stimulus onset for different methods. The threshold is set to be 1/6 of the maximum absolute value of the corresponding current estimates, similar to other experiments presented in this section. Similar to the previous experiments, the results

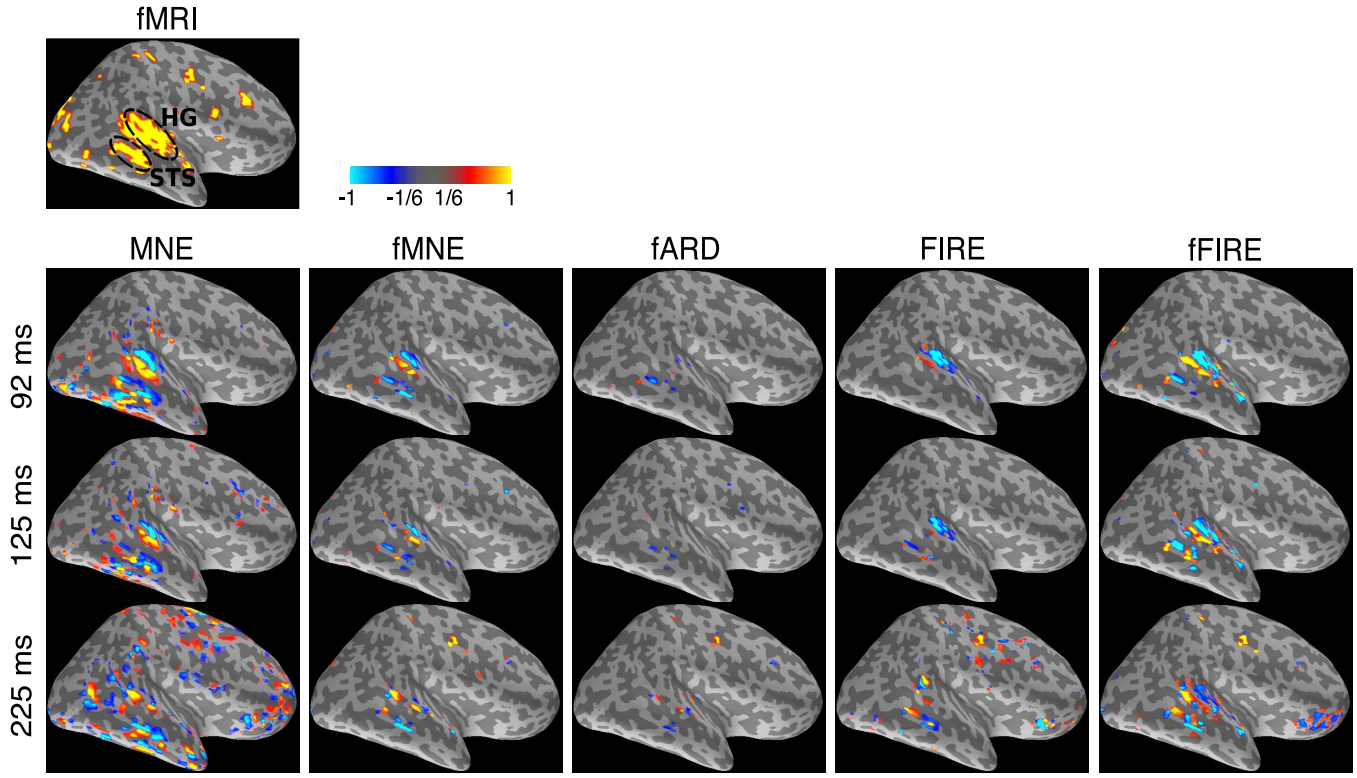


Figure 5.5. Human attention-shift auditory experiments. The top panel shows the fMRI activation maps associated with the attention left-shifting task ($p \leq 0.0005$). The bottom panel shows the current estimates obtained via MNE, fMNE, fARD, FIRE, and fFIRE at 92, 125, and 225 ms after stimulus onset, respectively. Hot/cold colors indicate outward/inward current flow.

obtained from MNE are too diffuse, especially for the late time frames. The detected areas using fARD are spatially sparse. fMNE, FIRE, and fFIRE produce physiologically sound estimates. Compared to fMRI, FIRE and fFIRE remove several diffuse activation areas in MNE results; the resulting estimates are more similar to fMRI-SPM. Both FIRE and fFIRE consistently retain the anterior-frontal area, which is present in MNE but not in fMRI-SPM, indicating that the vascular activation in this area is too weak for the fMRI measurements.

We further analyzed the estimated timecourses in Heschl's gyrus (HG) and the STS (Fig. 5.6). The neural timecourses in the auditory cortex estimated using different methods are similar, with the peak time at 92 ms (Fig. 5.6(a)). However, MNE, fMNE, and fARD cannot distinguish the later activity in STS from the early activity in HG, reflected in a strong estimated activation before 100 ms in the STS timecourses. In

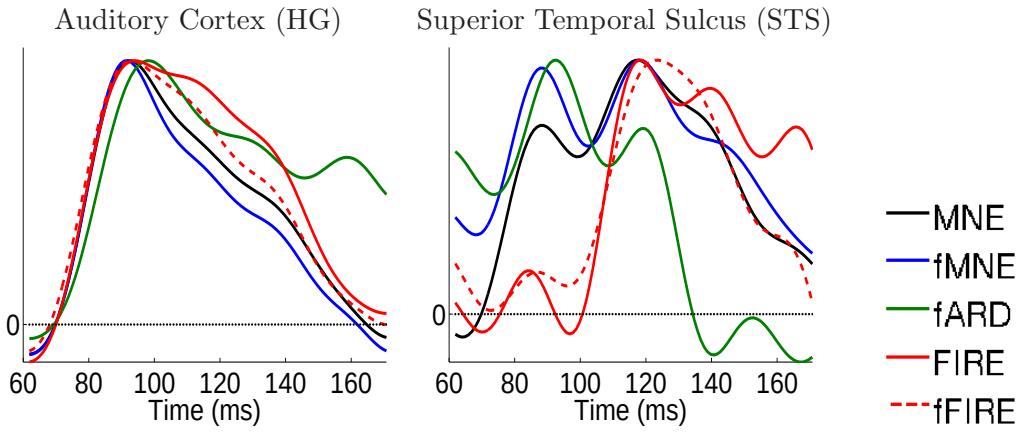


Figure 5.6. Estimated neural activity timecourses at the auditory cortex and STS using different methods.

contrast, FIRE and fFIRE differentiate activations in these two areas, detecting the activation in STS that peaks at approximately 120 ms. The results for the other two subjects we analyzed are similar to those presented here. Their early auditory activation in HG peaks at 93 and 97 ms, and STS peaks at 120 and 111 ms, respectively.

■ 5.3 Discussion

The computation for source space and lead-field matrix $\mathbf{A}$ is discussed in Section 3.2.1. The cortical regions for modeling purposes are defined by parceling the cortical surface using the FreeSurfer software, resulting in 35 parcels per hemisphere [Fischl *et al.*, 2002]. The boundaries of adjacent parcels are defined along sulci. We merge adjacent parcels that contain fewer than 30 vertices. We also tried another parcellation provided by FreeSurfer with 85 parcels per hemisphere. The results for this parcellation tend to be less stable, especially for parcels with few vertices. FIRE is also compatible with a data-driven parcellation. However, this approach may create a undesirable bias due to the use of the data in both parcel generation and current estimation. This bias can be avoided if the data-driven parcellation is obtained using a separate independent functional data set.

Under the orientation constrains, most forward models $\mathbf{A}$ follow a local orientation convention: the currents flowing outward of the cortex are considered positive, the currents flowing inward are viewed as negative. That means if a region includes two

sides of a gyrus, the positive on the two sulcal walls corresponds to currents flowing in opposite directions. Hence, the local time courses will have opposite signals violating the assumption of a single time course. In this work, we set the regional orientation reference to be the largest left singular vector of the matrix formed by the outward cortical normals within a region. We then modify the sign of the columns in the forward matrix $\mathbf{A}$ corresponding to vertices based on the angle between their normal vectors and the reference vector. We reverse the sign of a column if the angle between the normal and the reference vector is greater than 90 deg. To display the estimated current $\mathbf{J}^*$, we reverse the sign alternation and display results using the local orientation convention mentioned above.

We apply the standard preprocessing to fMRI data, then estimate the hemodynamic response function (HRF) at each voxel with a finite impulse response regressor covering a 20-s time window using the FS-FAST software (MGH, Boston, MA). The estimated HRF is used as the hemodynamic data $\mathbf{f}_n$ in our model.

For a source space of $N \sim 10,000$ vertices and timecourses of $T_{\mathbf{J}} \sim 100$ and $T_{\mathbf{F}} \sim 10$ samples, FIRE takes less than 20 iterations until the energy function is reduced by less than 0.1% in the next iteration. In each iteration of the coordinate descent algorithm, the estimation of $\boldsymbol{\Theta}$ takes 30 seconds; the estimation of $\mathbf{J}$ takes 4 minutes on a standard PC (2.8 GHz CPU and 8 GB RAM), leading to the total run time of approximately 1.5 hours. Estimating $\mathbf{J}$ involves an inversion of an $(MT_{\mathbf{J}} + NT_{\mathbf{F}}) \times (MT_{\mathbf{J}} + NT_{\mathbf{F}})$ dense symmetric matrix $\boldsymbol{\Gamma}_{\mathbf{W}}$, which is too large to store in memory. Instead, we employ the conjugate gradient descent method to solve the corresponding system of linear equations. It usually takes 100 iterations until convergence. Using the same computational resources, the run time for fARD is shorter since it estimates the current source distribution at each time point separately, ignoring the dependencies in the signal across time. The total run time for fARD with 10 iterations is about 5 minutes.

The coupling between the spatial and temporal domains in the joint E/MEG-fMRI analysis has restricted many previous models to operate on a coarse source space. The use of a regional neurovascular coupling model proposed in this work reduces the computational burden, leading to a tractable reconstruction on a densely sampled source space, similar to that typically used in MNE.

In practice, it is often necessary to use different fMRI vs. MEG stimulation schedules. In this work, we used two types of experiments to test FIRE: (a) a somatosen-

sory experiment with identical MEG and fMRI paradigms and (b) an auditory experiment, which provides an example of a situation where exact matching fMRI vs. MEG paradigms may be suboptimal. Specifically, in the auditory experiment we used a blocked sparse-sampling fMRI design to mitigate acoustical scanner noise; adding similar interleaved EPI noise/silent baseline periods would have made the corresponding MEG measurement simply too long for the subjects. These experimental designs were based on previous studies [Ahveninen *et al.*, 2006, Jaaskelainen *et al.*, 2004]. Although the general activation patterns are the same across different modalities, we expect that minor discrepancies remain. Thus, the question on how well we can deal with such differences using FIRE depends on the underlying theoretical neurophysiological assumptions. One of the major advantages of FIRE is that its underlying generative model does not force perfect match in activations (Eq. (5.1)), avoiding excessive bias toward fMRI information.

As mentioned earlier, a more standard inference procedure for our graphical model would jointly consider $\mathbf{J}$ and $\mathbf{Z}$ as latent variables in the EM framework, while maximizing the log likelihood with respect to the model parameters $\mathbf{\Theta}$. Since $[\mathbf{J}, \mathbf{Z}]$ and the measurement are jointly Gaussian given the model parameters $\mathbf{\Theta}$, the posterior probability distribution of the latent variables $[\mathbf{J}, \mathbf{Z}]$ is also Gaussian, leading to a closed-form update. Similar to the derivations in Eq. (5.16), the M-step updates depend on the first- and second-order statistics of the latent variables, computed in the E-step. Since $\mathbf{J}$ is not fixed in this EM procedure, the estimate at each location depends on the estimate at all other locations in the source space, as opposed to the region-based estimation in Eq. (5.12) when $\mathbf{J}$ is fixed. Computing the second-order statistics involves solving $NT_{\mathbf{J}} + N$ systems of linear equations, each of which is of size $NT_{\mathbf{F}} + MT_{\mathbf{J}}$. In other words, we need to apply the conjugate gradient solver $NT_{\mathbf{J}} + N$ times, exacerbating the memory and runtime requirements for the procedure similar to the bottleneck step in our coordinate descent approach (Step ii-2 in the algorithm summary). Therefore, treating both $\mathbf{J}$ and $\mathbf{Z}$ as hidden variables is infeasible except for an extremely coarse discretization of the source space. Similarly, it is currently computationally infeasible to compute the variance of vec $\left(\widehat{\mathbf{J}}\right)$, since it requires applying the conjugate gradient solver $NT_{\mathbf{J}}$ times. Instead of the variance of the estimate, we provide an alternative way to study the sensitivity of the solver through Monte-Carlo simulations.

The estimation of $\mathbf{u}_k$ and $\mathbf{v}_k$ is closely related to the Canonical Correlation Analysis (CCA), which seeks vectors for projecting two high dimensional data sets $(\{\widehat{\mathbf{j}}_n\}_{n\in P_k}$

and $\{\mathbf{f}_n\}_{n\in P_k}$ in our case) to a low dimensional space so as to maximize the correlation coefficient between the resulting projection coordinates. The probabilistic interpretation of CCA, established in [Bach & Jordan, 2005], offers a generative perspective on the method. Moreover, the probabilistic interpretation also helps to naturally extend the CCA model by incorporating prior information such as prior distributions on the waveforms $\mathbf{U}$ and $\mathbf{V}$.

Since the cost function is not convex, our method depends on the initialization. MNE estimate is a reasonable choice for initialization since it is unbiased while fMNE is a good alternative as MNE estimates may be too diffuse in certain brain regions. Moreover, maximizing the cost function does not necessarily correspond to the best ROC performance. For the Monte Carlo simulation trials where value of the likelihood achieved by fFIRE is greater than that of FIRE, the ROC performance of fFIRE is often better than that of FIRE. A good ROC performance indicates the results are close to the ground truth, but it is not perfectly correlated with the likelihood values, which are based on an approximate model inference.

Our neurovascular coupling model is designed for fixed-orientation current estimates, since the latent-variable model assumes that the spatial concordance of neural and vascular activities is characterized by a scalar. For free-orientation current estimates, the neurovascular coupling model would have to be adjusted to handle the correspondence between the current flow in three directions and a single vascular activation timecourse at a certain location. Moreover, FIRE assumes a single activation waveform pair, $\mathbf{u}$ and $\mathbf{v}$, in a region. The validity of this assumption depends on the size of the region and the distance between two activation sources. We cannot easily extend FIRE to multiple activation waveform pairs per region, since such an extension does not capture the fact that the shape of the vascular activation timecourses from two distinct sources is often highly similar but the neural processes are different. In the situation where there are two distinct current sources in one region, our preliminary results demonstrate that FIRE can localize the two current sources, but the estimated timecourses are combinations of the true timecourses. We defer the extension for free-orientation estimate and the extension for multiple activation sources per region to future work.

■ 5.4 Summary

In contrast to most joint E/MEG-fMRI models, we explicitly take into account the inherent differences in the data measured by E/MEG and fMRI, allowing for common situations in real experiments where either neural or vascular activity is silent. The current source estimates can be computed efficiently with an iterative procedure which bears similarity with re-weighted MNE methods, except that the weights are based on both the current estimates in the last iteration and the fMRI data via the proposed spatial neurovascular coupling model. This construction of the weights reduces the excessive sensitivity to fMRI present in many joint E/MEG-fMRI analysis methods and leads to more accurate current estimates as demonstrated by our experimental results of both simulated and human data.

Chapter 6

Conclusions and Future Work

In this thesis, we proposed two novel EEG/MEG inverse solvers: the $\ell_1\ell_2$-norm estimator and FIRE.

To overcome the noise sensitivity of most focal source estimators, such as MCE, we took advantage of the relatively smooth nature of the underlying EEG/MEG source signals and computed the current sources for all temporal samples simultaneously. The resulting $\ell_1\ell_2$-norm estimator maintains the focal spatial pattern via ℓ_1-norm regularization in space and encourages smooth temporal dynamics through the use of temporal basis functions and ℓ_2-norm regularization. The $\ell_1\ell_2$-norm estimator is efficient as it can be formulated as an SOCP problem and solved using the interior-point method. Our experiments using both simulated and real EEG/MEG data demonstrate that the method achieves more accurate source estimation than the traditional approaches.

Despite improvements in estimation accuracy, the $\ell_1\ell_2$-norm estimator still experiences challenge in differentiating activations on two sides of a sulcal wall due to limited spatial resolution of EEG/MEG measurements. We incorporated fMRI information, which has better spatial resolution, into the EEG/MEG source estimation to further improve the accuracy. Our method FIRE, in general, encourages spatial similarity between neural and hemodynamic activities. Moreover, it explicitly takes into account the inherent differences in the data measured by EEG/MEG and fMRI, allowing for either the neural or the vascular activity to be silent in the EEG/MEG and fMRI measurements, respectively.

The modeling assumptions underlying FIRE are based on the findings from our neurovascular coupling experiments using median-nerve stimulation in human subjects using simultaneous MEG and DOI recordings. Our results show that the neural and hemodynamic activities generally match in space. Furthermore, temporal analysis of these two types of activities reveals that the habituation effect in the hemodynamic

response is stronger than that in the earliest neural response.

Allowing for the different physiological origins of EEG/MEG and fMRI signals, FIRE avoids excessive bias toward fMRI information exhibited by many EEG/MEG-fMRI analysis methods. It can further distinguish sources at close spatial locations as shown in both simulated and real data.

This thesis focuses on subject-specific EEG/MEG source estimation. Looking forward, there are two main avenues for future exploration: group analysis of EEG/MEG data and source estimation for other image modalities such as DOI.

There is still a wealth of untapped information that could be used for group analysis in spatial domain, temporal domain, frequency domain, or all. We categorize group analysis into two main approaches: statistical modeling of the estimated sources across subjects and a joint source estimation by combining data from multiple subjects.

For the former approach, Oken *et al.* studied the relationship between inter-subject EEG spectrum variability and disease therapy [Oken & Chiappa, 1988]. Work by Maestú *et al.* compared spatio-temporal differences in current sources between a group of patients with the Alzheimer's disease and a group of control subjects using MEG recordings [Maestú *et al.*, 2001]. We studied the temporal variation across subjects and across MEG systems using variance component analysis [Ou *et al.*, 2007]. Our analysis shows that inter-subject differences are the biggest factor in the signal variability. We demonstrated that the timing of the peaks is consistent in the early somatosensory response, which justifies a direct comparison of peak times acquired from different visits, subjects, and systems. Compared with peak times, peak magnitudes vary substantially across sites; modeling of this variability is necessary for data pooling.

To the best of our knowledge, no joint EEG/MEG source estimation method has been demonstrated so far that combines data from multiple subjects. This type of analysis is commonly seen in fMRI studies, for example, via random effect analysis [Worsley *et al.*, 2002]. Adopting random effect analysis in EEG/MEG source localization, one can estimate the common spatio-temporal source distribution across subjects while allowing for inter-subject variability.

Although the $\ell_1\ell_2$-norm estimator is designed for the EEG/MEG source estimation, we believe that its general concept is applicable to DOI source estimation [Graber *et al.*, 1995, O'Leary, 1996, Arridge *et al.*, 1999]. As mentioned in Section 2.4.2, DOI records near-infrared light reaching each detector from a light source. One can use the data to infer scattering and absorption as a function of position in the tissue. Furthermore, the DOI

scattering/absorption inference share the ill-posed nature similar to the EEG/MEG source estimation problem.

We expect that the spatial extent of the hemodynamic activity underlying DOI signals is larger than the neural activity measured by EEG/MEG, so the ℓ_1-norm regularization in space may not be optimal. Instead, one can encourage spatial continuity by applying the ℓ_1-norm regularization to the spatial derivatives in order to make the method more suitable to the DOI source estimation. The underlying hemodynamic source signal of DOI measurements is expected to be smoother than its neural counterpart. Hence, the use of temporal basis functions and ℓ_2-norm regularization are directly applicable.

EEG/MEG and DOI measure neural and hemodynamic activities, respectively. Moreover, the spatial resolution of EEG/MEG is generally higher than that of DOI data. By combining EEG/MEG and DOI, we expect to produce more consistent source estimates, especially for DOI source estimates. To achieve joint EEG/MEG-DOI analysis, we can extend the neurovascular coupling model in FIRE to help bridge the gap between these two types of measurements.

To summarize, the methods proposed and demonstrated in this thesis improve the accuracy of source estimation from EEG/MEG data and promise benefits for reconstruction methods in other fields of imaging.

Appendix A

Second-Order Cone Programming

■ A.1 From Quadratic Constraint to Second-Order Cone Constraint

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:

$$\|\mathbf{y} - \mathbf{A}\mathbf{j}\|_{\ell_2}^2 \leq w \tag{A.1}$$

With straightforward derivations including expanding the ℓ_2-norm and completing squares, we can show that Eq. (A.1) is equivalent to

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

Eq. (A.1) also implies $1 + 2\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{j} - \mathbf{y}^{\mathrm{T}}\mathbf{y} + w \geq 0$. Therefore, setting $x_0 = \frac{1+2\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{j}-\mathbf{y}^{\mathrm{T}}\mathbf{y}+w}{2}$ and $\bar{\mathbf{x}} = \begin{bmatrix} \mathbf{A}\mathbf{j} \\ \frac{1-2\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{j}+\mathbf{y}^{\mathrm{T}}\mathbf{y}-w}{2} \end{bmatrix}$, 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 $M + 2$.

■ A.2 Second-Order Cone Programming

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:

$$\min \mathbf{c}^{\mathrm{T}}\mathbf{x} \tag{A.3}$$

$$\text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b} \tag{A.4}$$

$$\mathbf{x} \in \mathcal{Q} \tag{A.5}$$

where $\mathcal{Q} = \{\mathbf{x} : \mathbf{x}_0 \geq \|\bar{\mathbf{x}}\|_{\ell_2}\}$. $\mathcal{Q}$ is also referred to as the Lorentz cone; it is self-dual. The dual cone $\mathcal{Q}^C$ is defined as

$$\mathcal{Q}^C := \left\{\mathbf{y} : \forall\, \mathbf{x} \in \mathcal{Q}, \mathbf{y}^{\mathrm{T}}\mathbf{x} \geq 0\right\} \tag{A.6}$$

It is straightforward to prove that the self-dual property, $\mathcal{Q} = \mathcal{Q}^C$, 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. (A.3)-(A.5) is defined as follows:

$$\max \mathbf{b}^{\mathrm{T}}\mathbf{y} \tag{A.7}$$

$$\text{s.t. } \mathbf{A}^{\mathrm{T}}\mathbf{y} + \mathbf{z} = \mathbf{c} \tag{A.8}$$

$$\mathbf{z} \in \mathcal{Q}^{\mathcal{C}} = \mathcal{Q} \tag{A.9}$$

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

$$\mathbf{A}\mathbf{x} = \mathbf{b} \quad \mathbf{x} \in \mathcal{Q} \tag{A.10}$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{y} + \mathbf{z} = \mathbf{c} \quad \mathbf{z} \in \mathcal{Q} \tag{A.11}$$

$$\mathbf{x}^{\mathrm{T}}\mathbf{z} = 0 \tag{A.12}$$

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. (A.10) and Eq. (A.11). In fact, one can combine the two conic constraints and the complementary slackness condition in Eq. (A.12), and reduce them

to a more suitable form (Eq. (A.15)) for numerical solvers. The revised system of linear equations becomes

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{A.13}$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{y} + \mathbf{z} = \mathbf{c} \tag{A.14}$$

$$\mathbf{x} \circ \mathbf{z} = \mathbf{0} \tag{A.15}$$

$$\text{where} \quad \mathbf{x} \circ \mathbf{z} := \begin{bmatrix} \mathbf{x}^{\mathrm{T}}\mathbf{z} \\ x_0\bar{\mathbf{z}} + z_0\bar{\mathbf{x}} \end{bmatrix} = \begin{bmatrix} \mathbf{x}^{\mathrm{T}}\mathbf{z} \\ x_0 z_1 + z_0 x_1 \\ \vdots \\ x_0 z_n + z_0 x_n \end{bmatrix} \tag{A.16}$$

We refer readers to [Alizadeh & Goldfarb, 2001] for detailed derivations.

The primal-dual interior-point method solves this system of linear equations (Eq. (A.13)-(A.15)) using Newton's method. The optimization begins with a relaxed version of the complementary condition (Eq. (A.15)), and slowly strengthens this condition as iterations proceed. Iterations stop once the residual is less than a pre-selected threshold.

Appendix B

FIRE Inference Details

■ B.1 Estimating J

In this Appendix, we describe the estimation procedure for $\mathbf{J}$ based on the standard jointly Gaussian distribution. As mentioned before, $\mathbf{W}$ and $\mathbf{J}$ are jointly Gaussian. For fixed $\widehat{\mathbf{\Theta}}$, if we define

$$\mathbf{W} = \left[\ \mathbf{f}^{\mathrm{T}}(1), \cdots, \mathbf{f}^{\mathrm{T}}(T_{\mathbf{F}}), \mathbf{y}^{\mathrm{T}}(1), \cdots, \mathbf{y}^{\mathrm{T}}(T_{\mathbf{J}}) \ \right]^{\mathrm{T}}, \tag{B.1}$$

then

$$\mathbf{W} \sim \mathcal{N}\left(\mathbf{0}, \mathbf{\Gamma}_{\mathbf{W}}\right) = \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \mathbf{\Gamma}_{\mathbf{F}} & \mathbf{\Gamma}_{\mathbf{F},\mathbf{Y}} \\ \mathbf{\Gamma}_{\mathbf{F},\mathbf{Y}}^{\mathrm{T}} & \mathbf{\Gamma}_{\mathbf{Y}} \end{bmatrix}\right) \tag{B.2}$$

where $\mathbf{\Gamma}_{\mathbf{X},\mathbf{Y}}$ are the covariance matrix of the corresponding random variables $\mathbf{X}$ and $\mathbf{Y}$. Furthermore,

$$\left[\ \mathbf{W}^{\mathrm{T}}, \mathbf{j}^{\mathrm{T}}(1), \cdots, \mathbf{j}^{\mathrm{T}}(T_{\mathbf{J}}) \ \right]^{\mathrm{T}} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \mathbf{\Gamma}_{\mathbf{W}} & \mathbf{\Gamma}_{\mathbf{W},\mathbf{J}} \\ \mathbf{\Gamma}_{\mathbf{W},\mathbf{J}}^{\mathrm{T}} & \mathbf{\Gamma}_{\mathbf{J}} \end{bmatrix}\right). \tag{B.3}$$

Here, we only show the derivations of the covariance matrices for a single region ($K = 1$). The extension to multiple regions is straightforward. Based on the definition of covariance, we obtain

$$\begin{aligned} \mathbf{\Gamma}_{\mathbf{F}} &= \left(\mathbf{v}\mathbf{v}^{\mathrm{T}} + \xi^2 \mathbf{I}_{T_{\mathbf{F}}}\right) \otimes \mathbf{I}_N \\ \mathbf{\Gamma}_{\mathbf{Y}} &= \left(\mathbf{u}\mathbf{u}^{\mathrm{T}} + \eta^2 \mathbf{I}_{T_{\mathbf{J}}}\right) \otimes \mathbf{A}\mathbf{A}^{\mathrm{T}} + \mathbf{I}_{MT_{\mathbf{J}}} \\ \mathbf{\Gamma}_{\mathbf{F},\mathbf{Y}} &= \left(\mathbf{v}\mathbf{u}^{\mathrm{T}}\right) \otimes \mathbf{A}^{\mathrm{T}} \\ \mathbf{\Gamma}_{\mathbf{W},\mathbf{J}} &= \left[\left(\mathbf{v}\mathbf{u}^{\mathrm{T}}\right) \otimes \mathbf{I}_N, \quad \left(\mathbf{u}\mathbf{u}^{\mathrm{T}} + \eta^2 \mathbf{I}_{T_{\mathbf{J}}}\right) \otimes \mathbf{A}\right] \end{aligned} \tag{B.4}$$

where $\mathbf{I}_N$ indicates identify matrix of size N. Eq. (B.4) assumes a normalized E/MEG sensory noise with unit variance. Matrix Kronecker product $\otimes$ stems from the interactions between space and time in the model.

We can then express the conditional distribution $p(\mathbf{J}|\mathbf{W})$ using the Bayes' rule:

$$p(\mathrm{vec}\,(\mathbf{J})\,|\mathbf{W}) = \frac{p(\mathbf{W}, \mathrm{vec}\,(\mathbf{J}))}{p(\mathbf{W})} = \mathcal{N}\left(\mathbf{\Gamma}_{\mathbf{W},\mathbf{J}}^{\mathrm{T}}\mathbf{\Gamma}_{\mathbf{W}}^{-1}\mathbf{W}, \mathbf{\Gamma}_{\mathbf{J}} - \mathbf{\Gamma}_{\mathbf{W},\mathbf{J}}^{\mathrm{T}}\mathbf{\Gamma}_{\mathbf{W}}^{-1}\mathbf{\Gamma}_{\mathbf{W},\mathbf{J}}\right).$$

■ B.2 The M-step

In this Appendix, we derive the M-step for estimating Θ. When $\widehat{\mathbf{J}}$ is fixed, we employ the EM algorithm to optimize the model parameters $\theta_k = [\mathbf{u}_k, \mathbf{v}_k, \gamma_k]$ for each region separately:

$$\widehat{\theta}_k = \arg\max_{\theta_k} \int_{\underline{\mathbf{z}}_k} q(\underline{\mathbf{z}}_k) \log p(\{\mathbf{f}_n, \widehat{\mathbf{j}}_n\}_{n\in P_k}, z_n; \theta_k) d\underline{\mathbf{z}}_k \tag{B.5}$$

$$= \arg\max_{\theta_k} \int_{\underline{\mathbf{z}}_k} q(\underline{\mathbf{z}}_k) \log p(\{\mathbf{f}_n, \widehat{\mathbf{j}}_n\}_{n\in P_k}|z_n; \mathbf{u}_k, \mathbf{v}_k) p(\underline{\mathbf{z}}_k; \gamma_k) d\underline{\mathbf{z}}_k$$

$$= \arg\max_{\theta_k} \int_{\underline{\mathbf{z}}_k} q(\underline{\mathbf{z}}_k) \sum_{n\in P_k} \left(\log p(\mathbf{f}_n|z_n; \mathbf{v}_k) + \log p(\widehat{\mathbf{j}}_n|z_n; \mathbf{u}_k)\right) d\underline{\mathbf{z}}_k + \int_{\underline{\mathbf{z}}_k} q(\underline{\mathbf{z}}_k) \log p(\underline{\mathbf{z}}_k; \gamma_k) d\underline{\mathbf{z}}_k$$

$$= \arg\max_{\theta_k} \left[-\frac{1}{2} \sum_{n\in P_k} \langle \|\mathbf{f}_n - z_n\mathbf{v}_k\|^2 \rangle_q - \frac{1}{2} \sum_{n\in P_k} \langle \|\widehat{\mathbf{j}}_n - z_n\mathbf{u}_k\|^2 \rangle_q - \frac{1}{2} \frac{\langle \underline{\mathbf{z}}_k^{\mathrm{T}} \mathbf{\Gamma}_k^{-1} \underline{\mathbf{z}}_k \rangle_q}{\gamma_k^2} - N_k \log \gamma_k^2 \right] \tag{B.6}$$

$$= \arg\max_{\theta_k} \Big[N_k \mathbf{v}_k^{\mathrm{T}} \mathbf{f}_n \langle z_n \rangle_q - \frac{N_k}{2} \mathbf{v}_k^{\mathrm{T}} \mathbf{v}_k \langle z_n^2 \rangle_q + N_k \mathbf{u}_k^{\mathrm{T}} \widehat{\mathbf{j}}_n \langle z_n \rangle_q - \frac{N_k}{2} \mathbf{u}_k^{\mathrm{T}} \mathbf{u}_k \langle z_n^2 \rangle_q$$

$$- \frac{1}{2} \frac{\langle \underline{\mathbf{z}}_k^{\mathrm{T}} \mathbf{\Gamma}_k^{-1} \underline{\mathbf{z}}_k \rangle_q}{\gamma_k^2} - N_k \log \gamma_k^2 \Big]$$

$$= \arg\max_{\theta_k} \left[N_k \left(\mathbf{v}_k^{\mathrm{T}} \mathbf{f}_n + \mathbf{u}_k^{\mathrm{T}} \widehat{\mathbf{j}}_n\right) \langle z_n \rangle_q - \frac{N_k}{2} \left(\mathbf{v}_k^{\mathrm{T}} \mathbf{v}_k + \mathbf{u}_k^{\mathrm{T}} \mathbf{u}_k\right) \langle z_n^2 \rangle_q - \frac{1}{2} \frac{\langle \underline{\mathbf{z}}_k^{\mathrm{T}} \mathbf{\Gamma}_k^{-1} \underline{\mathbf{z}}_k \rangle_q}{\gamma_k^2} - N_k \log \gamma_k^2 \right]. \tag{B.7}$$

Eq. (B.6) is obtained with parameter setting $\eta_k^2 = \xi_k^2 = 1$ as discussed in Section 5.1.2. Therefore, we only need to update $\langle z_n \rangle_q$, $\langle z_n^2 \rangle_q$, and $\langle \underline{\mathbf{z}}_k^{\mathrm{T}} \mathbf{\Gamma}_k^{-1} \underline{\mathbf{z}}_k \rangle_q$ in the E-step. By equating the derivatives of Eq. (B.7) with respect to $\mathbf{u}_k$, $\mathbf{v}_k$, and γ_k^2 to zero, we obtain the M-step update in Eq. (5.16).

Bibliography

[Ahlfors & Simpson, 2004] Ahlfors, S. and Simpson, G. Geometrical interpretation of fMRI-guided MEG/EEG inverse estimates. *NeuroImage*, 22:323-32, 2004. 17, 37

[Ahveninen *et al.*, 2006] Ahveninen, J., Jaaskelainen, I.P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., Levanen, S., Lin, F.H., Sams, M., Shinn-Cunningham, B.G., Witzel, T., Belliveau, J.W. Task-modulated "what" and "where" pathways in human auditory cortex. *Proc Natl Acad Sci U S A*, 103:14608-13, 2006. 17, 102, 106

[Alizadeh & Goldfarb, 2001] Alizadeh, F. and Goldfarb, D. Second-order cone programming. Technical Report. RRR Report number 51-2001, RUTCOR, Rutgers University, 2001. 19, 43, 115

[Allison *et al.*, 1989] Allison, T., McCarthy, G., Wood, C.C., Darcey, T.M., Spencer, D.D., and Williamson, P.D. Human cortical potentials evoked by stimulation of the median nerve. I. cytoarchitectonic areas generating short-latency activity. *J. Neurophysiol.*, 62:694-710, 1989. 18, 31

[Allison, 1992] Allison, T., Scalp and cortical recordings of initial somatosensory cortex activity to median nerve stimulation in man. *Ann NY Acad Sci.*, 388:671-678, 1992. 20, 68

[Altmann *et al.*, 2008] Altmann, C.F., Henning, M., Doring, M.K., Kaiser, J. Effects of feature-selective attention on auditory pattern and location processing. *NeuroImage*, 41:69-79, 2008. 102

[Arridge *et al.*, 1999] Arridge, S. Optical tomography in medical imaging. *NeuroImage*, 15:41-93, 1999. 33, 35, 110

[Auranen *et al.*, 2007] Auranen, T., Nummenmaa, A., Hämäläinen, M.S., Jääskeläinen I.P., Lampinen, J., Vehtari, A., and Sams, M. Bayesian analysis of the neuromagnetic inverse problem with l^p-norm priors. *NeuroImage*, 26:870-884, 2007. 65

[Bach & Jordan, 2005] Bach, F. and Jordan M. A probabilistic interpretation of canonical correlation analysis. Technical Report 688, UC Berkeley, 2005. 93, 107

[Baillet & Sereno, 1997] Baillet, S. and Garnero, L. A Bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem. *IEEE Trans. Biomed. Eng.*, 44:374-385, 1997. 32

[Baillet *et al.*, 2001] Baillet, S., Mosher, J.C., and Leahy, R.M., Electromagnetic brain mapping *IEEE Signal Processing Magazine*, 18:14-30, 2001. 17, 72

[Bandettini & Ungerleider, 2001] Bandettini, P.A. and Ungerleider, L.G., From neuron to BOLD: new connections. *Nat Neurosci* 4:864-866, 2001. 84

[Bandettini *et al.*, 2002] Bandettini, P.A., Birn, R.M., Saad, Z.S., and Kelley, D., Dynamic nonlinearities in BOLD contrast: neuronal or hemodynamic? *International Congress Series*, 1235:73-85, 2002. 73, 78, 84

[Barth *et al.*, 1982] Barth, D.S., Sutherling, W., Engel, J., and Beatty, J. Neuromagnetic localization of epileptiform spike activity in the human brain. *Science*, 218:891–894, 1982. 18, 31

[Bertsimas & Tsitsiklis, 2008] Bertsimas, D. and Tsitsiklis, J. Introduction to Linear Optimization. Athena Scientisc, Belmont, MA, 2008. 31

[Bijma *et al.*, 2005] Bijma, F., de Munck, J.C., and Heethaar, R.M. The spatiotemporal MEG covariance matrix modeled as a sum of Kronecker products. *NeuroImage*, 27:402-415, 2005. 40

[Birn *et al.*, 2001] Birn, R.M., Saad, Z.S., and Bandettini, P.A., Spatial heterogeneity of the nonlinear dynamics in the fMRI BOLD response. *NeuroImage*, 14:817–826, 2001. 36, 73, 78, 84

[Birn & Bandettini, 2005] Birn, R.M. and Bandettini, P.A., The effect of stimulus duty cycle and "off" duration on BOLD response linearity. *NeuroImage*, 27:70–82, 2005. 36, 73, 78

[Boas *et al.*, 2001] Boas, D., Brooks, D., Miller, E., DiMarzio, C., Kilmer, M. Gaudette, R. and Zhang, Q. Imaging the body with diffuse optical tomography. *IEEE Signal Processing Magazine*, 18:57-75, 2001. 33, 35

[Boynton *et al.*, 1996] Boynton, G.M., Engel, S.A., Glover, G.H., and Heeger, D.J., Linear systems analysis of functional magnetic resonance imaging in human V1. *J. NeuroSci*, 16:4207–4221, 1996.

[Brenner *et al.*, 1978] Brenner, D., Lipton, J., Kaufman, L., and Williamson, S.J., Somatically evoked magnetic fields of the human brain. *Science*, 199:81-83, 1978. 20, 68

[Brooks *et al.*, 1999] Brooks, D.H., Ahmad, G.F., MacLeod, R.S., and Maratos, G.M. Inverse electrocardiography by simultaneous imposition of multiple constraints. *IEEE Trans. Biomed. Eng.*, 44:374-385, 1997. 32

[Butti *et al.*, 2006] Butti M., Pastori, A., Merzagora, A., Zucca, C., Bianchi, A., Reni, G., and Cerutti, S., Multimodal analysis of a sustained attention protocol: Continuous Performance Test assessed with Near Infrared Spectroscopy and EEG. *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, 1:1040-1043, 2006. 36

[Buxton *et al.*, 1998] Buxton, R., Wong, E., and Frank, L., Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. *Magn. Reson. Med.*, 39:855-64, 1998. 38

[Cauller & Kulics, 1991] Cauller, L.J. and Kulics, A.T., The neural basis of the behaviorally relevant N1 component of the somatosensory-evoked potential in SI cortex of awake monkeys: evidence that backward cortical projections signal conscious touch sensation. *Exp Brain Res*, 84:607-619, 1991. 85

[Caesar *et al.*, 2003] Caesar, K., Thomsen, K., and Lauritzen, M., Dissociation of spikes, synaptic activity, and activity-dependent increments in rat cerebellar blood flow by tonic synaptic inhibition. *Proc Natl Acad Sci U. S. A.*, 100:16000-16005, 2003. 85

[Chafetz & Cadilhac, 1954] Chafetz, M. and Cadilhac, J., A new procedure for a study of barbiturate effect and evoked potentials in the EEG. *Electroencephalogr. Clin. Neurophysiol.*, 6:465-72, 1954. 17

[Cottereau *et al.*, 2007] Bottereau, B., Jerbi, K., and Baillet, S. Multiresolution imaging of MEG cortical sources using an explicit piecewise model. *NeuroImage* 38:439–451, 2007. 66

[Cuffin & Cohen, 1977] Cuffin, B.N. and Cohen, D. Magnetic fields of a dipole in special volume conductor shapes. *IEEE Trans. Biomed. Engr.* 24:372–381, 1977. 28

[Custo & Wells, 2006] Custo, A. and Wells, W. Combined optical and magnetic functional brain imaging. *In Proc. HBM: Human Brain Mapping*, 2006. 35

[Dale & Sereno, 1993] Dale A.M. and Sereno M. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. *J. Cogn. Neurosci.* 5:162–176, 1993. 18, 28, 29, 30, 44, 46, 72

[Dale, 1999a] Dale, A.M., Optimal experimental design for event-related fMRI. *Hum. Brain Mapp.*, 8:109-114, 1999. 67

[Dale *et al.*, 1999b] Dale, A.M., Fischl, B., and Sereno, M. Cortical surface-based analysis: I. segmentation and surface reconstruction. *NeuroImage* 9:179-194, 1999. 46, 58, 71

[Dale *et al.*, 2000] Dale, A.M., Liu, A.K., Fischl, B., Buckner, R.L., Belliveau, J.W., Lewine, J.D., and Halgren, E., Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron*, 26:55-67, 2000. 72

[Daunizeau *et al.*, 2007] Daunizeau, J., Grova, C., Marrelec, G., Mattout, J., Jbabdi, S., Pelegrini-Issac, M., Lina, J., and Benali, H. Symmetrical event-related EEG/fMRI information fusion in a variational Bayesian framework. *NeuroImage*, 36:69-87, 2007. 37, 38

[Dempster *et al.*, 1977] Dempster, A., Laird, N., and Rubin, D. Maximum likelihood from incomplete data via the EM algorithm. *J of Roy. Stat. Soc. B*, 39:1-38, 1977. 22, 92, 93

[Deneux & Faugeras, 2006] Deneux, T. and Faugeras, O. EEG-fMRI fusion of non-triggered data using Kalman filtering. *ISBI*, 1068-71, 2006. 37, 38

[Devor *et al.*, 2003] Devor, A., Dunn, A.K, Andermann, M.L., Ulbert, I., Boas D.A., and Dale, A.M., Coupling of total hemoglobin concentration, oxygenation, and neural activity in rat somatosensory cortex. *Neuron*, 39:353–359, 2003. 36

[Di *et al.*, 1990] Di, S., Baumgartner, C., and Barth, D.S., Laminar analysis of extracellular field potentials in rat vibrissa/barrel cortex. *J Neuroscience*, 63:832-840, 1999. 85

[Ding & He, 2007] Ding, L. and He, B. Sparse source imaging in electroencephalography with accurate field modeling. *Hum. Brain Mapp.*, in press. 45

[Donoho & Elad, 2003] Donoho, D.L. and Elad, M. Maximal sparsity representation via L_1 minimization. *Proc. Natl. Acad. Sci. U.S.A.*, 100:2197-2202, 2003. 41

[Dunsby & French, 2003] Dunsby, C. and French, P. Techniques for depth-resolved imaging through turbid media including coherence gated imaging. J. Phys. D: Appl. Phys., 36:207-27, 2003. 35

[Efron & Tibshirani, 2002] Efron, B. and Tibshirani, R. Empirical bayes methods and false discovery rates for microarrays. Genetic Epidemiology, 23:70-86, 2002. 46

[Engel *et al.*, 2001] Engel, A., Fries, P., and Singer W. Dynamic predictions: oscillations and synchrony in top-down processing. Nat. Rev. Neurosci., 2:704-16, 2001. 17

[Fischl *et al.*, 1999] Fischl, B., Sereno, M., and Dale, A.M. Cortical surface-based analysis: II. inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9:195-207, 1999. 46, 58, 71

[Fischl *et al.*, 2002] Fischl, B., Salat, D., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., and Dale, A. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33:341-55, 2002. 21, 104

[Forss *et al.*, 1994] Forss, N., Hari, R., Salmelin, R., Ahonen, A., Hämäläinen, M., Kajola, M., Knuutila, J., and Simola, J. Activation of the human posterior parietal cortex by median nerve stimulation. *Exp. Brain Res.*, 99:309–15, 1994. 61

[Franceschini *et al.*, 2003] Franceschini, M.A., Fantini, S. Thompson, J.H. Culver, J.P., and Boas, D.A., Hemodynamic evoked response of the sensorimotor cortex measured

noninvasively with near-infrared optical imaging. *Psychophysiology*, 40:548-560, 2003. 69, 72

[Franceschini *et al.*, 2008] Franceschini, M.A., Nissilä, I., Wu, W., Diamond, S.G., Bonmassar, G., and Boas, D.A., Coupling between somatosensory evoked potentials and hemodynamic response in the rat. *NeuroImage*, 41:189-203, 2008. 84, 85

[Friston *et al.*, 1995] Friston, K., Holmes, A., Worsley, K., Poline, J-B., Frith, C., and Frackowiak, R., Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain Mapp.*, 2:189-210, 1995. 35

[Friston *et al.*, 2000] Friston, K., Mechelli, A., Turner, R., and Price, C., Nonlinear responses in fMRI: the Balloon model, Volterra kernels, and other hemodynamics. *NeuroImage*, 12:466-77, 2000. 38

[Fujimaki *et al.*, 2002] Fujimaki, N., Hayakawa, T., Nielsen, M., Knsche, T.R., Miyauchi, S. An fMRI-constrained MEG source analysis with procedures for dividing and grouping activation. *NeuroImage*, 17:324-43, 2002.

[Galka *et al.*, 2004] Galka, A., Yamashita, O., Ozaki, T., Biscay, R., and Valdes-Sosa P. A solution to the dynamical inverse problem of EEG generation using spatiotemporal Kalman filtering. *NeuroImage*, 23:435-453, 2004. 32

[Genovese *et al.*, 2002] Genovese, C.R., Lazar, N.A., and Nichols, T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, 15:870–878, 2002. 46

[Geva, 1998] Geva, A.B., Bioelectric sources estimation using spatio-temporal matching pursuit. *J. Appl. Sigal Processing*, 5:195–208, 1998. 32

[George *et al.*, 1995] George, J.S., Aine, C.J., Mosher, J.C., Schmidt, D.M., Ranken, D.M., Schlitt, H.A., Wood, C.C., Lewine, J.D., Sanders, J.A., Belliveau, J.W. Mapping function in the human brain with magnetoencephalography, anatomical magnetic resonance imaging, and functional magnetic resonance imaging. *J. Clin. Neurophysiol.*, 12:406-31, 1995.

[Goff *et al.*, 1978] Goff, W.R., Allison, T., and Vaughan, H.G., The functional neuroanatomy of event related potentials. *In* Event-Related Brain Potentials in Man.

Callaway, E., Tueting P., and Koslow, S.H. Eds. Academic Press, New York (1978), 1-79. 20, 68

[Gorodnitsky & Rao, 1997] Gorodnitsky, I.F. and Rao, B.D. Sparse signal reconstruction from limited data using FOCUSS: a re-weighted minimum norm algorithm. *IEEE Trans. Signal Processing*, 45:600–616, 1997. 22, 31

[Graber *et al.*, 1995] Graber, H., Barbour, R., and Chang, J., Algebraic reconstruction of images of a diffusive medium containing strong absorbers: comparative study of different illumination schemes and the effect of restricted view angle. *In Proc. SPIE, Optical Tomography: Photon Migration and Spectroscopy of Tissue and Model Media: Theory, Human Studies, and Instrumentation*, 2389:448-64, 1995. 110

[Gramfort, 2009] Gramfort, A., Mapping, timing and tracking cortical activations with MEG and EEG: Methods and application to human vision. *Ph.D. Thesis*, INRIA Sophia-Antipolis, 2009. 43

[Grill-Spector *et al.*, 2006] Grill-Spector, K., Henson, R., and Martin, A., Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci.*, 10:14-23, 2006. 17, 36

[Gruber *et al.*, 2002] Gruber, T., and Müller, M., Effects of picture repetition on induced gamma band responses, evoked potentials, and phase synchrony in the human EEG. *Cogn. Brain Res.*, 13:377-92, 2002. 36

[Hadamard, 1902] Hadamard J. Sur les problems aux derivees partielles et leur signification physique. *Princeton University Bulletin*, 49–52, 1902. 17, 29

[Hall *et al.*, 1999] Hall, D.A., Haggard, M.P., Akeroyd, M.A., Palmer, A.R., Summerfield, A.Q., Elliott, M.R., Gurney, E.M., Bowtell, R.W. "Sparse" temporal sampling in auditory fMRI. *Hum. Brain Mapp.*, 7:213-23, 1999. 102

[Hämäläinen & Ilmoniemi,1984] Hämäläinen, M.S. and Ilmoniemi R. Interpreting measured magnetic fields of the brain: Estimates of current distributions. Techical Report TKK-F-A559, 1984. 17, 18, 29, 30, 72

[Hämäläinen & Sarvas, 1989] Hämäläinen, M.S. and Sarvas, J. Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. *IEEE Trans. Biomed. Eng.*, 36:165–171, 1989. 28, 46, 58, 71

[Hämäläinen *et al.*, 1993] Hämäläinen, M.S., Hari, R., Ilmoniemi, R., Knuutila, J., and Lounasmaa, O.V. Magnetoencephalography - theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, 65:413–497, 1993. 17, 27, 40, 93

[Hämäläinen, 2005] Hämäläinen, M.S. MNE software user's guide. NMR Center, Mass General Hospital, Harvard University. 58, 59, 75

[Hari *et al.*, 1984] Hari, R., Reinikainen, K., Kaukoranta, E., Hämäläinen, M., Ilmoniemi, R., Penttinen, A., Salminen, J., and Teszner, D., Somatosensory evoked cerebral magnetic fields from SI and SII in man. *Electroencephalogr. Clin. Neurophysiol.*, 57:254-263, 1984. 20, 68

[Hari & Forss, 1999] Hari, R. and Forss N. Magnetoencephalography in the study of human somatosensory cortical processing. *Philos. Trans. R. Soc. Lond. B, Biol. Sci.*, 354:1145–1154, 1999. 17, 20, 29, 59, 61, 68, 74, 75, 100

[Hart *et al.*, 2002] Hart, H.C., Palmer, A.R., Hall, D.A. Heschl's gyrus is more sensitive to tone level than non-primary auditory cortex. *Hear. Res.* 171:177-90, 2002. 102

[Hellweg *et al.*, 1977] Hellweg, F., Schultz, W., and Creutzfeldt, O. Extracellular and intracellular recordings from cat's cortical whisker projection area: thalamocortical response transformation. *J. Neurophysiol.*, 40:463–479, 1977. 85

[Herrmann *et al.*, 2008] Herrmann, M.J., Huter, T., Plichta, M.M., Ehlis, A.C., Alpers, G.W., Mühlberger, A., Fallgatter, A.J., Enhancement of activity of the primary visual cortex during processing of emotional stimuli as measured with event-related functional near-infrared spectroscopy and event-related potentials *Hum. Brain Mapp.*, 29:28-35, 2008. 36

[Huang *et al.*, 1998] Huang, M., Aine, C.J., Supek, S., Best, E., Ranken, D., and Flynn, E.R. Multi-start downhill simplex method for spatio-temporal source localization in magnetoencephalography. *Electroencephalogr. Clin. Neurophysiol.* 108:32–44, 1998. 29

[Huang *et al.*, 2006] Huang, M.X., Dale, A.M., Song, T., Halgren, E., Harrington, D.L., Podgomy, I., Canive, J.M., Lewis, S., and Lee, R.R. Vector-based spatial-temporal minimum L1-norm solution for MEG. *NeuroImage* 31:1025–1037, 2006. 32, 39, 65

[Huizenga *et al.*, 2002] Huizenga, H.M., de Munck, J.C., Waldorp, L.J., and Grasman, R.P.P.P. Spatiotemporal EEG/MEG source analysis based on a parametric noise covariance model. *IEEE Trans. Biomed. Eng.*, 49:533-539, 2002. 40

[Iadecola, 2004] Iadecola, C., Neurovascular regulation in the normal brain and in Alzheimer's disease. *Nat. Rev. Neurosci.*, 5:347-360, 2004. 85

[Jaaskelainen *et al.*, 2004] Jääskeläinen I.P., Ahveninen, J., Bonmassar, G., Dale, A.M., Ilmoniemi, R.J., Levänen, S., Lin, F.H., May, P., Melcher, J., Stufflebeam, S., Tiitinen, H., Belliveau, J.W. Human posterior auditory cortex gates novel sounds to consciousness. *PNAS* 101:6809-14, 2004. 17, 106

[Jerbi *et al.*, 2004] Jerbi, K., Baillet, S., Mosher, J.C., Nolte, G., Gamero, L., and Leahy, R.M. Localization of realistic cortical activity in MEG using current multipoles. *NeuroImage*, 22:779–793, 2004. 29

[Jezzard *et al.*, 2002] Jezzard, F., Matthews, P.M. and Smith, S.M. Functional MIR – An Introduction to Methods. *OXFORD*, 2002. 32, 33

[Jones *et al.*, 2001] Jones, M., Berwick, J., Johnston, D., and Mayhew, J., Concurrent optical imaging spectroscopy and laser-Doppler flowmetry: the relationship between blood flow, oxygenation, and volume in rodent barrel cortex. *Neuroimage*, 13:1002-1015, 2001. 36

[Jones *et al.*, 2004] Jones, M., Hewson-Stoate, N., Martindale, J., Redgrave, P., and Mayhew, J., Nonlinear coupling of neural activity and CBF in rodent cortex. *Neuroimage*, 22:956-965, 2004. 85

[Jun *et al.*, 2008] Jun, S.C., George, J.S., Kim, W., Paré-Blagoev, J., Plis, S., Ranken, D.M., Schmidt, D.M. Bayesian brain source imaging based on combined MEG/EEG and fMRI using MCMC. *NeuroImage* 40:1581-94, 2008.

[Kaipio & Somersalo, 2004] Kaipio, J. and Somersalo, E. Statistical and Computational Inverse Problems. Springer, 2004. 47

[Kaufman *et al.*, 1981] Kaufman, L., Okada, Y., Brenner, D., Williamson, S.J., On the relation between somatic evoked potentials and fields. *Int J Neurosci*, 15:223-239, 1981. 20, 68

[Kershman *et al.*, 1951] Kershman J., Vasquez, J., and Golstein, S., The incidence of focal and non-focal EEG abnormalities in clinical epilepsy. *Electroencephalogr. Clin. Neurophysiol.*, 3:15-24, 1951. 17

[Koch *et al.*, 2008] Koch, S.P., Koendgen, S., Bourayou, R., Steinbrink, J., and Obrig, H., Individual alpha-frequency correlates with amplitude of visual evoked potential and hemodynamic response. *NeuroImage*, 41:233-242, 2008. 36

[Kass *et al.*, 1979] Kass, J., Nelson, R., Sur, M., Lin, C-S, and Merzenich, M. Multiple representations of the body within the primary somatosensory cortex of primates. *Science*, 204:521–523, 1979. 59

[Knake *et al.*, 2006] Knake, S. *et al.* The value of multichannel MEG and EEG in the presurgical evaluation of 70 epilepsy patients. *Epilepsy Res.*, 69:80–86, 2006. 17, 25

[Kwong *et al.*, 1992] Kwong, K., Belliveau, J., Chesler, D., Goldberg, I., Weisskoff, R., Poncelet, B., Kennedy, D., Hoppel, B., Cohen, M., Turner, R., Cheng, H., Brady, T., and Rosen, B., Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *PNAS*, 89:5675-9, 1992. 33

[Lachaux *et al.*, 1999] Lachaux, J., Rodriguez, E., Martinerie, J., Varela, F. Measuring phase synchrony in brain signals. *Hum. Brain Mapp.* 8:194-208, 1999. 17

[Lamus *et al.*, 2007] Lamus, C., Long, C.J., Hämäläinen, M.S., Brown, E.N., and Purdon, P.L. Parameter estimation and dynamic source localization for the magnetoencephalography (MEG) inverse problem. *IEEE International Symposium on Biomedical Imaging*, 2007. 32

[Li *et al.*, 1956] Li, C.L., Cullen, C., and Jasper, H.H. Laminar microelectrode studies of specific somatosensory cortical potentials. *J Neurophysiol* 19:111-130, 1956. 85

[Limpiti *et al.*, 2006] Limpiti, T., Van Veen, B.D., and Wakai, R.T. Cortical patch basis model for spatially extended neural activity. *IEEE Trans. Biomed. Eng.*, 53:1740-1754, 2006. 66

[Lin *et al.*, 2004] Lin. F-H., Witzel, T., Hämäläinen, M.S., Belliveau, and Stufflebeam, S.M. Spectral spatiotemporal imaging of cortical oscillations and interactions in the human brain. *NeuroImage*, 23:582-95, 2004. 17

[Lin *et al.*, 2006] Lin. F-H., Belliveau, J.W., Dale, A.M., and Hämäläinen, M.S. Distributed current estimates using cortical orientation constraints. *Hum. Brain Mapp.*, 27:1–13, 2006. 47, 65

[Liu *et al.*, 1998] Liu, A.K., Belliveau, J.W., and Dale, A.M.. Spatiotemporal imaging of human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations. *Proc. Natl. Acad. Sci. U.S.A.* 95:8945-50, 1998. 37, 61, 94, 95

[Logothetis & Wandell, 2004] Logothetis, N.K. and Wandell, B.A. Interpreting the BOLD signal. *Annu. Rev. Physiol.* 66:735-769, 2004. 19, 36, 74

[Lu *et al.*, 1992] Lu, Z.L., Williamson, S.J., Kaufman, L. Human auditory primary and association cortex have differing lifetimes for activation traces. *Brain Res.* 572:236-41, 1992. 102

[Mackert *et al.*, 2004] Mackert, B.M., Wubbeler, G., Leistner, S., Uludag, K., Obrig, H., Villringer, A., Trahms, L., and Curio, G., Neurovascular coupling analyzed noninvasively in the human brain. *Neuroreport* 15:63-66, 2004. 82

[Mackert *et al.*, 2008] Mackert, B.M., Leistner, S., Sander, T., Liebert, A., Wabnitz, H., Burghoff, M., Trahms, L., Macdonald, R., and Curio, G. Dynamics of cortical neurovascular coupling analyzed by simultaneous DC-magnetoencephalography and time-resolved near-infrared spectroscopy. *Neuroimage* 39:979-986, 2008. 36, 82

[Maestú *et al.*, 2001] Maestú F., Fernández, A., Simos, P., Gil-Gregorio, P., Amo, C., Rodriguez, R., Arrazola, J., and Ortiz, T. Spatio-temporal patterns of brain magnetic activity during a memory task in Alzheimer's disease. *Neuroreport*, 12:3917-22, 2001. 25, 110

[Malioutov *et al.*, 2005] Malioutov, M., Çetin, M., and Willsky, A.S. A sparse signal reconstruction perspective for source localization with sensor arrays. *IEEE Trans. Signal Processing*, 53:3010–3022, 2005. 39, 45

[Mantini *et al.*, 2007] Mantini, D., Perrucci, M.G., Cugini, S., Ferretti, A., Romani, G.L., Del Gratta, C., Complete artifact removal for EEG recorded during continuous fMRI using independent component analysis. *NeuroImage*, 34:598-607, 2007. 36

[Martindale *et al.*, 2005] Martindale, J., Berwick, J., Martin, C., Kong, Y., Zheng, Y., and Mayhew, J., Long duration stimuli and nonlinearities in the neural-haemodynamic coupling. *J of Cereb Blood Flow Metab*, 25:651-661, 2005. 36, 74, 84

[Mathiesen *et al.*, 1998] Mathiesen, C., Caesar, K., Akgören, N., and Lauritzen, M., Modification of activity-dependent increases of cerebral blood flow by excitatory synaptic activity and spikes in rat cerebellar cortex. *J Physiol*, 512:555-566, 1998. 85

[Matsuura & Okabe, 1999] Matsuura, K. and Okabe, Y. Multiple current-dipole distribution reconstruction reconstructed by modified selective minimum-norm method. Advances in Biomagnetism Res.: Biomag96, Springer, New York, 290–293. 65

[Mayberg *et al.*, 2000] Mayberg, H., Brannanab, S., Tekellb, J., Silvab, J., Mahurinab, R., McGinnisa, S., Jerabekac, P., Regional metabolic effects of fluoxetine in major depression: serial changes and relationship to clinical response. *Biological Psychiatry*, 48:830-43, 2000. 25

[Mauguiere *et al.*, 1997] Mauguiere, F., Merlet, I., Forss, N., Vanni, S., Jousmäki, V., Adeleine, P., and Hari, R., Activation of a distributed somatosensory cortical network in the human brain. A dipole modelling study of magnetic fields evoked by median nerve stimulation. Part I: Location and activation timing of SEF sources. *Electroencephalogr Clin Neurophysiol.* 104:281-289, 1997. 20, 68

[Mitzdorf, 1985] Mitzdorf, U., Current source-density method and application in cat cerebral cortex: Investigation of evoked potentials and EEG phenomena. *Physiol. Rev.*, 65:37-100, 1985. 85

[Mosher *et al.*, 1992] Mosher, J.C., Lewis, P.S., and Leahy, R.M. Multiple dipole modeling and localization from spatio-temporal MEG data. *IEEE Trans. Biomed. Eng.*, 39:541–557, 1992. 18, 29, 44, 72

[Negishi *et al.*, 2004] Negishi, M., Abildgaard, M., Nixon, T., and Constable, R.T., Removal of time-varying gradient artifacts from EEG data acquired during continuous fMRI. *Clin Neurophysiol*, 115:2181-2192, 2004. 36

[Nelder & Mead, 1965] Nelder, J.A. and Mead, R. A Simplex Method for Function Minimization. *Computer Journal* 7:308–313, 1965. 61, 72

[Nesterov & Nemirovski, 1994] Nesterov, Y., and Nemirovski, A. Interior-point polynomial methods in convex programming. *Studies in Applied Mathematics*, SIAM, Philadelphia, PA, 1994. 114

[Ngai *et al.*, 1999] Ngai, A.C., Jolley, M.A., D'Ambrosio, R., Meno, J.R., and Winn, H.R., Frequency-dependent changes in cerebral blood flow and evoked potentials during somatosensory stimulation in the rat. *Brain Res.* 837:221-228, 1999. 36

[Norup Nielsen & Lauritzen, 2001] Norup Nielsen, A., and Lauritzen, M., Coupling and uncoupling of activity-dependent increases of neuronal activity and blood flow in rat somatosensory cortex. *J of Physiol*, 533:773-785, 2001. 36

[Nunez, 1981] Nunez, P.L., Electric fields fo the brain. New York: Oxford University Press, 1981. 26

[Nummenmaa *et al.*, 2007] Nummenmaa, A., Auranen, T., Hämäläinen, M., Jääskeläinen, I., Lampinen, J., Sams, M., and Vehtari, A., Hierarchical Bayesian estimates of distributed MEG sources: theoretical aspects and comparison of variational and MCMC methods. *Neuroimage* 35:669-85, 2007. 65

[Obrig *et al.*, 2000] Obrig, H., Neufang, M., Wenzel, R., Kohl, M., Steinbrink, J., Einhäupl, K., and Villringer, A., Spontaneous low frequency oscillations of cerebral hemodynamics and metabolism in human adults. *Neuroimage* 12:623-639, 2000. 72

[Obrig *et al.*, 2002] Obrig, H., Israel, H., Kohl-Bareis, M., Uludag, K., Wenzel, R., Müller, B., Arnold, G., and Villringer, A., Habituation of the visually evoked potential and its vascular response: implications for neurovascular coupling in the healthy adult. *NeuroImage*, 17:1-18, 2002. 36

[Ogawa *et al.*, 1990] Ogawa, S., Lee, T., Nayak, A., and Glynn, P., Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magn. Reson. Med.*, 14:68-78, 1990. 33

[O'Leary, 1996] O'Leary, M., Imaging with diffuse photon density waves. PhD Thesis, Dept. Physics, Univ. Pennsylvania, Philadelphia, PA, 1995. 110

[Oken & Chiappa, 1988] Oken, B. and Chiappa, K., Short-term variability in EEG frequency analysis. *Electroencephalogr. Clin. Neurophysiol.*, 69:191-8, 1988. 110

[Oostendorp & Van Oosterom, 1989] Oostendorp, T.F. and Van Oosterom, A. Source parameter estimation in inhomogeneous volume conductors of arbitrary shape. *IEEE Trans. Biomed. Eng.*, 36:382–391, 1989. 28, 46, 58, 71

[Ou *et al.*, 2007] Ou, W., Golland, P., and Hämäläinen, M. Sources of Variability in MEG. In Proc. MICCAI: International Conference on Medical Image Computing and Computer Assisted Intervention, LNCS 4792:751-9, 2007. 110

[Ou *et al.*, 2009a] Ou, W., Hämäläinen, M., and Golland, P. A distributed spatio-temporal EEG/MEG inverse solver. *NeuroImage*, 44:932-46, 2009. 39

[Ou *et al.*, 2009b] Ou, W., Nissilä I., Radhakrishnan, H., Boas, D., Hämäläinen, M., and Franceschini, M. Study of neurovascular coupling in humans via simultaneous Magnetoencephalography and diffuse optical imaging acquisition. *NeuroImage*, 46:624-32, 2009. 67

[Ou *et al.*, 2009c] Ou, W., Nummenmaa, A., Hämäläinen, M., and Golland, P. Multimodal functional imaging using fMRI-informed regional EEG/MEG source estimation. *In Proc. IPMI: International Conference on Information Processing and Medical Imaging*, LNCS 5636:88-100, 2009. 87

[Pantazis *et al.*, 2005] Pantazis, D., Nichols, T.E., Baillet, S., and Leahy, R.M. A comparison of random field theory and permutation methods for the statistical analysis of MEG data. *NeuroImage* 25:383–394, 2005. 45

[Pascual-Marqui *et al.*, 1994] Pascual-Marqui R.D., Michel, C.M., and Lehmann, D. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *Int. J. of Psychophysiol.*, 18:49–65, 1994. 44, 65, 90

[Plonsey 1969] Plonsey, R. Bioelectric Phenomena. New York, McGraw-Hill, 1969. 26

[Rao & Kreutz-Delgado, 1999] Rao, B.D. and Kreutz-Delgado, K. An affine scaling methodology for best basis selection. *IEEE Trans. Signal Processing* 47: 187-200, 1999. 31

[Rauschecker, 1998] Rauschecker, J.P. Cortical processing of complex sounds. *Curr. Opin. Neurobiol.* 8:516-21, 1998. 102

[Riera *et al.*, 2007] Riera, J.J., Jimenez, J.C., Wan, X., Kawashima, R., and Ozaki, T., Nonlinear local electrovascular coupling II: From data to neuronal masses. *Hum. Brain Mapp.*, 28:225-254, 2007. 36

[Ritter & Villringer, 2006] Ritter, P. and Villringer, A., Simultaneous EEG-fMRI. *Neurosci Biobehav Rev*, 30:823-838, 2006. 36

[Rovati *et al.*, 2007] Rovati, L., Salvatori, G., Bulf, L., and Fonda, S., Optical and electrical recording of neural activity evoked by graded contrast visual stimulus. *Biomed Eng Online*, 6:28, 2007. 36

[Rowe *et al.*, 1996] Rowe, M., Turman, A., Murray, G., and Zhang, H. Parallel organization of somatosensory cortical areas I and II for tactile processing. *Clin. Exp. Pharmacol. Physiol.*, 23:931–938, 1996. 59

[Royl *et al.*, 2006] Royl, G., Leithner, C., Sellien, H., Müller, J.P., Megow, D., Offenhauser, N., Steinbrink, J., Kohl-Bareis, M., Dirnagl, U., Lindauer, U., Functional imaging with Laser Speckle Contrast Analysis : Vascular compartment analysis and correlation with Laser Doppler Flowmetry and somatosensory evoked potentials *Brain Res.*, 1121:95-103, 2006. 36

[Sander *et al.*, 2007] Sander, T.H., Liebert, A., Mackert, B.M., Wabnitz, H., Leistner, S., Curio, G., Burghoff, M., Macdonald, R., and Trahms, L. DC-magnetoencephalography and time-resolved near-infrared spectroscopy combined to study neuronal and vascular brain responses. *Physiol Meas*, 28:651-664, 2007. 82

[Sato *et al.*, 2004a] Sato, H., Kiguchi, M., Kawaguchi, F., and Maki, A. Practicality of wavelength selection to improve signal-to-noise ratio in near-infrared spectroscopy. *NeuroImage*, 21:1554-1562, 2004. 70

[Sato *et al.*, 2004b] Sato, M., Yoshioka. T., Kajihara, S., Toyama, K., Goda, N., Doya, K., and Kawato, M. Hierarchical Bayesian estimation for MEG inverse problem. *NeuroImage*, 23:806-26, 2004. 37, 38, 90, 95

[Scherg & Von Cramon, 1985] Scherg M. and Von Cramon, D. Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model. *Electroencephalogr. Clin. Neurophysiol.*, 62:32–44, 1985. 18, 29

[Schmid *et al.*, 2006] Schmid, M.C., Oeltermann, A., Juchem, C., Logothetis, N.K., Smirnakis, S.M., Simultaneous EEG and fMRI in the macaque monkey at 4.7 Tesla. *Magn Reson Imaging*, 24:335-342, 2006. 36

[Schmitt *et al.*, 2001] Schmitt, U., Louis, A.K., Darvas, F., Buchner, H., and Fuchs, M. Numerical aspects of spatio-temporal current density reconstruction from EEG-/MEG-data. *IEEE Trans. Med. Imaging*, 20:314–324, 2001. 32

[Sharon *et al.*, 2007] Sharon, D., Hämäläinen, M., Tootell, R., Halgren, E., Belliveau, J., The advantage of combining MEG and EEG: comparison to fMRI in focally stimulated visual cortex. *NeuroImage*, 36:1225-35, 2007. 19

[Sheth *et al.*, 2003] Sheth S.A., Nemoto, M., Guiou, M., Walker, M., Pouratian, N., and Toga, A.W., Evaluation of coupling between optical intrinsic signals and neuronal activity in rat somatosensory cortex. *NeuroImage*, 19:884-894, 2003. 36

[Sheth *et al.*, 2004] Sheth S.A., Nemoto, M., Guiou, M., Walker, M., Pouratian, N., and Toga, A.W., Linear and nonlinear relationships between neuronal activity, oxygen metabolism, and hemodynamic responses. *Neuron*, 42:347-355, 2004. 85

[Strangman *et al.*, 2003] Strangman, G., Franceschini, M.A., and Boas, D.A., Factors affecting the accuracy of near-infrared spectroscopy concentration calculations for focal changes in oxygenation parameters. *Neuroimage*, 18:865-879, 2003. 70

[Sturm *et al.*, 2001] Sturm, J.F., Romanko, O., Polik, I., and Terlaky, T. Self-Dual-Minimization (SeDuMi): solver for optimization problems over symmetric cones. http://sedumi.mcmaster.ca/. 48, 50

[Tiihonen *et al.*, 1989] Tiihonen, J., Hari, R., and Hämäläinen, M., Early deflections of cerebral magnetic responses to median nerve stimulation. *Electroencephalogr Clin Neurophysiol.*, 74:290-296, 1989. 20, 68

[Tipping & Bishop, 1999] Tipping, M. and Bishop, C. Probabilistic principal component analysis. *J Royl. Stat. Soc. B*, 61:611-22, 1999. 93

[Tipping, 2001] Tipping, M. Sparse Bayesian learning and the relevance vector machine. *J of Machine Learning Research*, 1:211-44, 2001. 38

[Trujillo-Barreto *et al.*, 2007] Trujillo-Barreto, N., Aubert-Vazquez, E., and Penny, W. Bayesian M/EEG source reconstruction with spatio-temporal priors. *NeuroImage*, 39:318–335, 2007. 32

[Tuomisto *et al.*, 1983] Tuomisto, T., Hari, R., Katila, T., Poutanen, T., and Varpula, T. Studies of auditory evoked magnetic and electric responses: Modality specificity and modelling. *Nuovo Cimento*, D 2:471–483, 1983. 63

[Uludag *et al.*, 2004] Uludag, K., Steinbrink, J., Villringer, A., and Obrig, H., Separability and cross talk: optimizing dual wavelength combinations for near-infrared spectroscopy of the adult head. *NeuroImage*, 22:583-589, 2004. 70

[Ureshi *et al.*, 2004] Ureshi, M., Matsuura, T., and Kanno, I., Stimulus frequency dependence of the linear relationship between local cerebral blood flow and field potential evoked by activation of rat somatosensory cortex. *Neurosci Res*, 48:147-153, 2004. 36

[Uutela *et al.*, 1998] Uutela, K., Hämäläinen, M.S., and Salmelin, R.. Global optimization in the localization of neuromagnetic sources. *IEEE Trans. Biomed. Eng.* 45:716–723, 1998. 29

[Uutela *et al.*, 1999] Uutela, K., Hämäläinen, M.S., and Somersalo, E.. Visualization of magnetoencephalographic data using minimum current estimates. *NeuroImage* 10:173–180, 1999. 17, 18, 29, 31, 48, 65

[Vanni *et al.*, 2004] Vanni, S., Warnking, J., Dojat, M., Delon-Martin, C., Bullier, J., Segebarth, C. Sequence of pattern onset responses in the human visual areas: an fMRI constrained VEP source analysis. *NeuroImage* 21:801-17, 2004.

[Vasios *et al.*, 2006] Vasios, C.E., Angelone, L.M., Purdon, P.L., Ahveninen, J., Belliveau, J.W., Bonmassar, G., EEG/(f)MRI measurements at 7 Tesla using a new EEG cap ("InkCap"). *NeuroImage*, 33:1082-1092, 2006. 36

[Vazquez & Noll, 1998] Vazquez, A.L. and Noll, D.C., Nonlinear aspects of the BOLD response in functional MRI. *Neuroimage*, 7:108-118, 1998. 36

[Wang *et al.*, 1993] Wang, J.Z., Williamson, S.J., and Kaufman, L. Magnetic source imaging based on the minimum-norm least-squares inverse. *Brain Topography*, 5:365–371, 1993. 17, 18, 29, 30

[Weerd & Kap, 1981] Weerd, J.P.C. and Kap, J.I. Spectro-temporal representations and time-varying spectra of evoked potentials. *Biol. Cybern.*, 41:101–107, 1981. 40

[Williamson *et al.*, 1991] Williamson, S.J., Lu, Z.L., Karron, D., Kaufman, L.. Advantages and limitations of magnetic source imaging. *Brain Topography* 4:169-80, 1991. 102

[Wikström *et al.*, 1996] Wikström, H., Huttunen, J., Korvenoja, A., Virtanen, J., Salonen, O., Aronen, H., and Ilmoniemi, R.J. Effects of interstimulus interval on somatosensory evoked magnetic fields (SEFs): a hypothesis concerning SEF generation at the primary sensorimotor cortex. *Electroencephalogr. Clin. Neurophysiol.*, 100:479–487, 1996. 21, 60, 69

[Wipf & Rao, 2004] Wipf, D. and Rao, B. Sparse Bayesian learning for basis selection. *IEEE Trans. Signal Processing*, 52:2153-64, 2004. 38

[Wipf & Nagarajan, 2009] Wipf, D. and Nagarajan, S. A unified Bayesian framework for MEG/EEG source imaging. *NeuroImage*, 44:947-66, 2009. 90

[Wood, 1982] Wood C.C. Application of dipole localization methods to source identification of human evoked potentials. *Ann. N.Y. Acad. Sci.*, 388:139–155, 1982. 18, 29

[Worsley *et al.*, 2002] Worsley, K., Liao, C., Aston, J., Petre, V., Duncan, G., Morales, F., and Evans, A. A general statistical analysis for fMRI data. *NeuroImage*, 15:1–15, 2002. 35, 110

[Wray *et al.*, 1988] Wray, S., Cope, M., and Delpy, D.T. Characteristics of the near infrared absorption spectra of cytochrome aa3 and hemoglobin for the noninvasive monitoring of cerebral oxygenation. *Biochim Biophys Acta*, 933:184-192, 1988. 71

[Yamashita *et al.*, 2001] Yamashita, Y., Maki, A., and Kiozumi, H. Wavelength dependence of the precision of noninvasive optical measurement of oxy-, deoxy-, and total-hemoglobin concentration. *Med. Phys.*, 28:1108-1114, 2001. 70

[Zhang *et al.*, 2005a] Zhang, Y., Ghodrati, A., and Brooks, D.H. An analytical comparison of three spatio-temporal regularization methods for dynamic linear inverse problems in a common statistical framework. *Inverse Probl.*, 21:357–382, 2005. 32

[Zhang *et al.*, 2005b] Zhang, Y., Brooks, D.H., Boas, D.A., A haemodynamic response function model in spatio-temporal diffuse optical tomography. *Phys Med Biol*, 50:4625-4644, 2005. 71

[xfit software] Source modeling software (xfit), Elekta-Neuromag Oy, Helsinki, Finland. 61, 72