Decoding Brain States Based on Magnetoencephalography From Prespecified Cortical Regions

Jinyin Zhang, Xin Li*, *Senior Member, IEEE*, Stephen T. Foldes, Wei Wang, Jennifer L. Collinger, Douglas J. Weber, *Senior Member, IEEE*, and Anto Bagić

Abstract—**Brain state decoding based on whole-head MEG has been extensively studied over the past decade. Recent MEG applications pose an emerging need of decoding brain states based on MEG signals originating from prespecified cortical regions. Toward this goal, we propose a novel region-of-interest-constrained discriminant analysis algorithm (RDA) in this paper. RDA integrates linear classification and beamspace transformation into a unified framework by formulating a constrained optimization problem. Our experimental results based on human subjects demonstrate that RDA can efficiently extract the discriminant pattern from prespecified cortical regions to accurately distinguish different brain states.**

Index Terms—**Decoding, discriminant analysis, magnetoencephalography (MEG), regions of interest (ROI).**

I. INTRODUCTION

Magnetoencephalography (MEG) is a noninvasive neuroimaging method that measures magnetic fields generated by electrical neural activity with high temporal resolution [1]. It is a valuable technique complementary to other noninvasive recording modalities such as electroencephalography and functional magnetic resonance imaging [2], [3].

Decoding brain states based on MEG has been extensively studied over the past decade [4]–[8]. A number of machine learning techniques, such as linear discriminant analysis (LDA) [5] and support vector machine [6], have been applied for MEG decoding. Most of these decoding algorithms aim to identify the optimal discriminant pattern that distinguishes different brain states with maximal accuracy. This optimal discriminant pattern is often constructed by brain signals arising from widely distributed cortical regions.

In this paper, we focus on a different decoding problem that has not been extensively explored in the literature. Our objective is to decode different brain states by using MEG signals generated from prespecified cortical regions [9], [10]. Such a

decoding problem is not trivial, because it is difficult, if not impossible, to force only the target cortical region to generate discriminant signals in most experiments. Since human actions are controlled by distributed cortical networks, the discriminant brain signals often come from multiple cortical regions [11]. For example, in a visually-guided motor study, a subject is asked to move the left or right hand by following a visual cue. The visual cue for different hands will activate discriminant signals in visual cortex, although the intention is to activate discriminant signals in motor cortex.

For this reason, we need to develop a new MEG decoding tool to extract the discriminant information from prespecified cortical regions to distinguish different brain states while suppressing the brain signals from other cortical regions. In this paper, the prespecified cortical regions are referred to as the *regions of interest* (ROI), other cortical regions outside ROI are referred to as the *regions of noninterest* (RON), and the aforementioned decoding problem is referred to as *region-of-interest-constrained decoding* (ROI decoding).

One possible approach for ROI decoding is to first extract the MEG signals from the ROI by source localization, and then construct a decoder to distinguish different brain states based on these signals from the ROI [12]–[16]. Such a two-step approach heavily depends on the results of the first step that may not perfectly capture all signals originating from the ROI and simultaneously remove all signals originating from the RON. It is well known that the source localization problem is profoundly underdetermined due to the limited number of MEG channels compared to the number of possible electrical sources [2]. Source localization algorithms choose a solution out of an infinite set of candidates based on various prior assumptions. In most cases, these algorithms can only capture the "dominant" sources, but not necessarily the "discriminant" sources inside the ROI. If the discriminant sources within the ROI are not accurately estimated, we cannot achieve high accuracy for the following decoding process.

On the other hand, it has been reported in the literature that ROI decoding may be implemented by using the channels on top of the ROI only [5], [7], [8]. An MEG channel often captures the brain signals originating from multiple cortical regions. Therefore, constraining the channel location does not necessarily constrain the ROI location on the cortex. For this reason, a simple channel selection approach cannot efficiently extract the MEG signals from the ROI and, hence, does not offer an optimal solution for ROI decoding.

In this paper, we propose a novel *ROI-constrained discriminant analysis* (RDA) algorithm to address the aforementioned ROI decoding problem. RDA formulates a constrained optimization to find the optimal spatial filter that can accurately distinguish different brain states based on the MEG signals generated from the ROI. In our proposed optimization formulation, the cost function is derived from the theory of linear classification. It aims to maximize the discriminant information carried by the output signal of the spatial filter so that different brain states can be accurately separated. The constraint is designed by borrowing the idea of beamspace transformation [14]–[16]. It guarantees that the spatial filter is able to pass the signals generated by the ROI while attenuating the signals generated by the RON. By simultaneously considering the decoding accuracy and the ROI constraint, RDA identifies the optimal discriminant pattern associated with the ROI and, hence, efficiently solves the ROI decoding problem. In addition, a numerical solver is developed in this paper to solve the nonconvex optimization problem posed by RDA with guaranteed global optimum.

The remainder of this paper is organized as follows. We derive the proposed RDA algorithm in Section II. Next, the efficiency of RDA is demonstrated by a number of simulation and experimental examples in Sections III and IV, followed by the discussions for future work in Section V. Finally, we conclude in Section VI and describe the numerical solver proposed for RDA in Appendix.

II. ROI-CONSTRAINED DISCRIMINANT ANALYSIS

Without loss of generality, we consider two brain states that are labeled as "State-1" and "State-2," respectively. For example, these two states may correspond to the brain activity when a subject performs two different movement tasks. We further assume that MEG data are recorded from M channels. We use a vector $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_M]^T \in \Re^M$ to represent the MEG features corresponding to these M channels, where x_m denotes the mth feature which is associated with the mth MEG channel. The feature vector $\mathbf{x}$ can be a collection of MEG signals from M different channels at a particular time t. Alternatively, if a linear transformation (e.g., short-time Fourier transform [17], wavelet transform [18], etc.) is applied, $\mathbf{x}$ can be a set of transformed signals in frequency or wavelet domain. To simplify our notation, we assume that a single feature is extracted from each MEG channel. However, the proposed RDA algorithm can be extended to accommodate multiple features per channel, as will be discussed in Section V.

RDA seeks for a linear combination of all features as the discriminant function for decoding

$$y = \mathbf{w}^T \cdot \mathbf{x}, \qquad (1)$$

where $\mathbf{w} \in \Re^M$ defines the weight values of all MEG features. Since $\mathbf{x}$ is the collection of features associated with the MEG channels at different spatial locations, the linear combination in (1) can be conceptually viewed as a spatial filter. For ROI decoding, the spatial filter must be optimally designed to meet the following two criteria:

1) *ROI Selectivity*: The spatial filter should pass the MEG signals generated by the neural sources within the ROI and attenuate the MEG signals generated by the RON.

2) *Decoding Accuracy*: The output of the spatial filter must carry sufficient discriminant information to accurately distinguish different brain states.

In what follows, we will formulate a constrained optimization to determine the vector $\mathbf{w}$ in (1) that simultaneously satisfies the aforementioned two criteria.

Consider two sets of training data $\{\mathbf{x}_{n,1}; n = 1, 2, \dots, N_1\}$ and $\{\mathbf{x}_{n,2}; n = 1, 2, \dots, N_2\}$ where $\mathbf{x}_{n,1}$ and $\mathbf{x}_{n,2}$ are the feature vectors of the nth trial for the brain states "State-1" and "State-2," respectively. The symbols N_1 and N_2 represent the numbers of training trials for these two brain states. In order to distinguish the two brain states, we follow the idea of LDA [5], [8], [19] to maximize the ratio of the between-class scatter over the within-class scatter for the output y of the spatial filter in (1)

$$\underset{\mathbf{w}}{\text{maximize}} \quad \frac{\mathbf{w}^T \cdot \mathbf{S}_B \cdot \mathbf{w}}{\mathbf{w}^T \cdot \mathbf{S}_W \cdot \mathbf{w}}. \qquad (2)$$

The between-class scatter matrix $\mathbf{S}_B \in \Re^{M \times M}$ and the within-class scatter matrix $\mathbf{S}_W \in \Re^{M \times M}$ are defined by

$$\mathbf{S}_B = N_1 \cdot (\boldsymbol{\mu}_1 - \boldsymbol{\mu}) \cdot (\boldsymbol{\mu}_1 - \boldsymbol{\mu})^T$$
$$+ N_2 \cdot (\boldsymbol{\mu}_2 - \boldsymbol{\mu}) \cdot (\boldsymbol{\mu}_2 - \boldsymbol{\mu})^T \qquad (3)$$

$$\mathbf{S}_W = \sum_{n=1}^{N_1} (\mathbf{x}_{n,1} - \boldsymbol{\mu}_1) \cdot (\mathbf{x}_{n,1} - \boldsymbol{\mu}_1)^T$$
$$+ \sum_{n=1}^{N_2} (\mathbf{x}_{n,2} - \boldsymbol{\mu}_2) \cdot (\mathbf{x}_{n,2} - \boldsymbol{\mu}_2)^T, \qquad (4)$$

where $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ stand for the mean values of $\{\mathbf{x}_{n,1}; n = 1, 2, \dots, N_1\}$ and $\{\mathbf{x}_{n,2}; n = 1, 2, \dots, N_2\}$, respectively, and $\boldsymbol{\mu}$ is the mean of $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$. Note that the matrix $\mathbf{S}_B$ is positive semi-definite and the matrix $\mathbf{S}_W$ is positive definite.

In addition to the cost function in (2), the spatial filter $y = \mathbf{w}^T \cdot \mathbf{x}$ must appropriately select the MEG signals generated by the ROI. Toward this goal, we adopt a vertex-based model that has been widely used by many MEG source localization algorithms [2], [20]. Specifically, we assume that the cortex is partitioned into a number of vertices. Each vertex contains a current dipole that models the electrical source within the vertex. Based on this model, the MEG feature vector $\mathbf{x}$ can be represented as

$$\mathbf{x} = \sum_{n=1}^{N} \mathbf{A}_n \cdot \mathbf{q}_n, \qquad (5)$$

where $\mathbf{A}_n \in \Re^{M \times 3}$ is the leadfield matrix of the nth dipole, $\mathbf{q}_n \in \Re^3$ denotes the moment of the nth dipole, and N is the total number of dipoles. The leadfield matrix $\mathbf{A}_n$ models the influence of the nth dipole on the measured MEG signals. It can be calculated according to the geometrical structure and the conducting medium of the human head [2].

Substituting (5) into (1), the output of the spatial filter $y = \mathbf{w}^T \cdot \mathbf{x}$ can be represented as

$$y = \mathbf{w}^T \cdot \sum_{n=1}^{N} \mathbf{A}_n \cdot \mathbf{q}_n. \tag{6}$$

We further partition y into two nonoverlap components:

$$y_\Phi = \mathbf{w}^T \cdot \sum_{n \in \Phi} \mathbf{A}_n \cdot \mathbf{q}_n \tag{7}$$

$$y_{\tilde{\Phi}} = \mathbf{w}^T \cdot \sum_{n \notin \Phi} \mathbf{A}_n \cdot \mathbf{q}_n \tag{8}$$

$$y = y_\Phi + y_{\tilde{\Phi}}, \tag{9}$$

where Φ represents the set containing all current dipoles inside the ROI. In other words, y_Φ in (7) represents the filter output generated by the sources within the ROI and $y_{\tilde{\Phi}}$ in (8) denotes the filter output generated by the sources inside the RON. The power of y_Φ and $y_{\tilde{\Phi}}$ can be calculated as

$$p_\Phi = y_\Phi y_\Phi^T = \mathbf{w}^T \cdot \sum_{m \in \Phi} \sum_{n \in \Phi} \mathbf{A}_m \cdot \mathbf{q}_m \cdot \mathbf{q}_n^T \cdot \mathbf{A}_n^T \cdot \mathbf{w} \tag{10}$$

$$p_{\tilde{\Phi}} = y_{\tilde{\Phi}} y_{\tilde{\Phi}}^T = \mathbf{w}^T \cdot \sum_{m \notin \Phi} \sum_{n \notin \Phi} \mathbf{A}_m \cdot \mathbf{q}_m \cdot \mathbf{q}_n^T \cdot \mathbf{A}_n^T \cdot \mathbf{w}. \tag{11}$$

We use the ratio between p_Φ and $p_{\tilde{\Phi}}$ to quantitatively measure the ROI selectivity of the spatial filter:

$$\frac{p_\Phi}{p_{\tilde{\Phi}}} = \frac{\mathbf{w}^T \cdot \sum_{m \in \Phi} \sum_{n \in \Phi} \mathbf{A}_m \cdot \mathbf{q}_m \cdot \mathbf{q}_n^T \cdot \mathbf{A}_n^T \cdot \mathbf{w}}{\mathbf{w}^T \cdot \sum_{m \notin \Phi} \sum_{n \notin \Phi} \mathbf{A}_m \cdot \mathbf{q}_m \cdot \mathbf{q}_n^T \cdot \mathbf{A}_n^T \cdot \mathbf{w}}. \tag{12}$$

In practice, the dipole moments $\{\mathbf{q}_n; n = 1, 2, \ldots, N\}$ are not known. Hence, the ratio in (12) must be estimated with a number of assumptions. For instance, if there is no prior information about the dipole moments, the traditional beamspace transformation method calculates the ratio $p_\Phi / p_{\tilde{\Phi}}$ by assuming that all dipoles are uniformly distributed and mutually independent [15]. With these assumptions, the ratio in (12) can be simplified as

$$\frac{p_\Phi}{p_{\tilde{\Phi}}} = \frac{\mathbf{w}^T \cdot \sum_{n \in \Phi} \mathbf{A}_n \cdot \mathbf{A}_n^T \cdot \mathbf{w}}{\mathbf{w}^T \cdot \sum_{n \notin \Phi} \mathbf{A}_n \cdot \mathbf{A}_n^T \cdot \mathbf{w}} = \frac{\mathbf{w}^T \cdot \mathbf{G}_\Phi \cdot \mathbf{w}}{\mathbf{w}^T \cdot \mathbf{G}_{\tilde{\Phi}} \cdot \mathbf{w}}, \tag{13}$$

where

$$\mathbf{G}_\Phi = \sum_{n \in \Phi} \mathbf{A}_n \cdot \mathbf{A}_n^T \tag{14}$$

$$\mathbf{G}_{\tilde{\Phi}} = \sum_{n \notin \Phi} \mathbf{A}_n \cdot \mathbf{A}_n^T. \tag{15}$$

The matrices $\mathbf{G}_\Phi$ and $\mathbf{G}_{\tilde{\Phi}}$ in (14) and (15) are both positive definite [15]. Once the leadfield matrices $\{\mathbf{A}_n; n = 1, 2, \ldots, N\}$ are known, $p_\Phi / p_{\tilde{\Phi}}$ in (13) is represented as a function of the vector $\mathbf{w}$.

Combining (2) and (13) yields the following constrained optimization problem:

$$\begin{aligned} \underset{\mathbf{w}}{\text{maximize}} \quad & \frac{\mathbf{w}^T \cdot \mathbf{S}_B \cdot \mathbf{w}}{\mathbf{w}^T \cdot \mathbf{S}_W \cdot \mathbf{w}} \\ \text{subject to} \quad & \frac{\mathbf{w}^T \cdot \mathbf{G}_\Phi \cdot \mathbf{w}}{\mathbf{w}^T \cdot \mathbf{G}_{\tilde{\Phi}} \cdot \mathbf{w}} \geq \lambda, \end{aligned} \tag{16}$$

where λ is a user-defined parameter. The optimization in (16) aims to find the optimal spatial filter $y = \mathbf{w}^T \cdot \mathbf{x}$ to distinguish two brain states, while simultaneously constraining the power ratio $p_\Phi / p_{\tilde{\Phi}}$ to be no less than λ.

Studying the optimization formulation in (16) reveals several important properties. First, since the left side of the constraint function is in the form of a generalized Rayleigh quotient, its lower bound and upper bound are equal to [21]

$$\inf_{\mathbf{w}} \frac{\mathbf{w}^T \cdot \mathbf{G}_\Phi \cdot \mathbf{w}}{\mathbf{w}^T \cdot \mathbf{G}_{\tilde{\Phi}} \cdot \mathbf{w}} = \eta \left(\mathbf{G}_{\tilde{\Phi}}^{-1} \cdot \mathbf{G}_\Phi \right) \tag{17}$$

$$\sup_{\mathbf{w}} \frac{\mathbf{w}^T \cdot \mathbf{G}_\Phi \cdot \mathbf{w}}{\mathbf{w}^T \cdot \mathbf{G}_{\tilde{\Phi}} \cdot \mathbf{w}} = \rho \left(\mathbf{G}_{\tilde{\Phi}}^{-1} \cdot \mathbf{G}_\Phi \right) \tag{18}$$

where inf($\bullet$) and sup ($\bullet$) denote the infimum (i.e., the greatest lower bound) and the supremum (i.e., the least upper bound) of a function, and $\eta(\bullet)$ and $\rho(\bullet)$ stand for the minimum and maximum eigenvalues of a matrix.

Second, adjusting the value of λ allows us to explore the trade-off between the discriminant power and the ROI selectivity of the spatial filter. In one extreme case, if λ is set to the minimum possible value (i.e., $\eta(\mathbf{G}_{\tilde{\Phi}}^{-1} \cdot \mathbf{G}_\Phi)$), the constraint in (16) is satisfied for any given $\mathbf{w}$ and, hence, the ROI decoding in (16) is equivalent to the traditional LDA in (2) where the discriminant information is maximized and the ROI constraint is completely ignored. In the other extreme case, if λ is set to the maximum possible value (i.e., $\rho(\mathbf{G}_{\tilde{\Phi}}^{-1} \cdot \mathbf{G}_\Phi)$), there is only one possible direction of $\mathbf{w}$ to satisfy the constraint in (16) and the optimal solution of (16) is independent of the cost function. In this case, the filter is designed to maximize the total signal power originating from the ROI, instead of the discriminant information generated by the ROI. In other words, it maximizes the ROI selectivity only and, hence, is unlikely to achieve high decoding accuracy. When λ varies between these two extreme cases, the optimization in (16) explores the tradeoff between its cost function (related to the decoding accuracy) and the constraint function (related to the ROI selectivity).

Third, neither the cost function nor the constraint set in (16) is convex. In this case, a general-purpose nonlinear optimizer is likely to settle at a local optimum and, hence, cannot efficiently find the global optimum. To the best of our knowledge, there is no existing optimization algorithm that guarantees to find the global optimum of (16). Motivated by this observation, we propose a novel numerical algorithm to solve (16) both efficiently (i.e., with low computational cost) and robustly (i.e., with guaranteed global optimum). The details of our proposed numerical solver are described in the Appendix. Algorithm 1 summarizes the major steps of the proposed RDA algorithm.

III. SIMULATION STUDIES

In this section, we will demonstrate how RDA controls the ROI selectivity by varying the value of λ based on simulation experiments. In addition, we will make a full comparison between RDA and other traditional methods for ROI decoding.

Algorithm 1: Region-of-Interest-Constrained Discriminant Analysis (RDA)

1. Start from the training data $\{\mathbf{x}_{n,1}; n = 1, 2, \ldots, N_1\}$ and $\{\mathbf{x}_{n,2}; n = 1, 2, \ldots, N_2\}$ corresponding to two different brain states, the user-specified ROI, and a prespecified threshold λ.
2. Partition the structural brain image into N vertices and define the set Φ containing all current dipoles inside the ROI.
3. Calculate the leadfield matrices $\{\mathbf{A}_n; n = 1, 2, \ldots, N\}$ for the current dipoles associated with the vertex-based model.
4. Calculate the matrices $\mathbf{S}_B$ and $\mathbf{S}_W$ based on (3), (4).
5. Calculate the matrices $\mathbf{G}_\Phi$ and $\mathbf{G}_{\tilde{\Phi}}$ based on (14), (15).
6. Formulate the constrained nonlinear optimization problem in (16).
7. Apply Algorithm 2 in the Appendix to find the optimal solution $\mathbf{w}^*$ of (16) to form the spatial filter in (1).
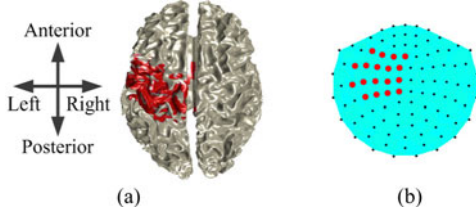

Fig. 1. (a) Red color indicates the spatial location of the ROI (i.e., the left precentral and postcentral gyri). (b) Shown in red are the 36 channels on top of the ROI where each dot represents a pair of gradiometers.

A. Simulation Setup

In the simulation experiments, we set the left sensorimotor area (i.e., the left precentral and the left postcentral gyri) as the ROI, as shown in Fig. 1(a). Two different brain states, "State-1" and "State-2," are considered. We assume that the discriminant information between "State-1" and "State-2" comes from the cortical regions both inside and outside the ROI. To generate the simulation data, the MRI image of a human subject is used to define the source space. A fixed grid with 2-mm spacing is generated from the gray/white matter boundary of the MRI image by using FreeSurfer [20]. It results in 172 902 vertices in total. Among them, 9481 vertices are inside the ROI and 163 421 vertices are inside the RON. We use one current dipole at each vertex of the grid to model the source within the brain [22]. In addition, we automatically parcellate and label the cortical surface as different cortical areas by using FreeSurfer [22].

Several different current dipoles and noise components are considered in our simulation setup:

1) 20 current dipoles are used to model the discriminant sources between "State-1" and "State-2." A half of these dipoles $\{\mathbf{q}_{ROI,n}; n = 1, 2, \ldots, 10\}$ are randomly distributed inside the ROI, and the other half $\{\mathbf{q}_{RON,n}; n = 1, 2, \ldots, 10\}$ are randomly distributed inside the RON. The magnitude of these dipoles is randomly generated from two Gaussian distributions: $N(0, 1)$ for "State-1" and $N(6, 1)$ for "State-2". The different

mean values of these two Gaussian distributions model the discriminant information of the two brain states. The orientation of these dipoles is uniformly distributed over all possible directions.

2) 200 current dipoles $\{\mathbf{q}_{Noi,n}; n = 1, 2, \ldots, 200\}$ are used to model the noisy sources that are independent of the brain states. These dipole locations are randomly distributed all over the cortical surface, including both the ROI and the RON. The magnitude of these dipoles is randomly generated from a Gaussian distribution $N(0, 2^2)$ and their orientation is uniformly distributed over all possible directions.

3) In addition to the current dipoles, a vector $\mathbf{n}$ is used to model the MEG measurement noise of all channels. In our simulation setup, we assume that the noise is statistically independent between different channels and it follows a Gaussian distribution $N(0, 0.001^2)$.

It is important to mention that there are several possible methods to set up the current dipoles to model neural sources. For instance, the orientation of a dipole may be constrained to be normal to the cortical surface or it may not be fixed to a specific direction [2]. In this study, we adopt the latter approach where the dipole orientation is not necessarily normal to the cortical surface and it can possibly take any direction.

Once the dipole model is set up, the leadfield of each dipole is calculated for a whole-head MEG system (Elekta Neuromag) with 204 gradiometers. These 204 gradiometers are distributed over 102 locations where each location contains two gradiometers. The MEG feature vector $\mathbf{x}$ is computed by

$$\mathbf{x} = \sum_{n=1}^{10} \mathbf{A}_{ROI,n} \cdot \mathbf{q}_{ROI,n} + \sum_{n=1}^{10} \mathbf{A}_{RON,n} \cdot \mathbf{q}_{RON,n}$$
$$+ \sum_{n=1}^{200} \mathbf{A}_{Noi,n} \cdot \mathbf{q}_{Noi,n} + \mathbf{n}, \qquad (19)$$

where $\mathbf{A}_{ROI,n}$, $\mathbf{A}_{RON,n}$ and $\mathbf{A}_{Noi,n}$ represent the corresponding leadfield matrices, respectively. According to (19), we generate 500 feature vectors for each brain state to construct the training set.

For the aforementioned training set, the discriminant information originates from both the ROI and the RON. However, ROI decoding aims to extract the optimal discriminant pattern originating from the ROI only to distinguish different brain states. In order to evaluate the performance of an ROI decoding algorithm, we consider two test sets:

1) A test set with discriminant signals generated by the dipoles outside the ROI only (TestSet-RON): The MEG feature vector is computed by

$$\mathbf{x} = \sum_{n=1}^{10} \mathbf{A}_{RON,n} \cdot \mathbf{q}_{RON,n} + \sum_{n=1}^{200} \mathbf{A}_{Noi,n} \cdot \mathbf{q}_{Noi,n} + \mathbf{n}. \qquad (20)$$

2) A test set with discriminant signals generated by the dipoles inside the ROI only (TestSet-ROI): The MEG feature vector is

computed by

$$\mathbf{x} = \sum_{n=1}^{10} \mathbf{A}_{ROI,n} \cdot \mathbf{q}_{ROI,n} + \sum_{n=1}^{200} \mathbf{A}_{Noi,n} \cdot \mathbf{q}_{Noi,n} + \mathbf{n}. \quad (21)$$

For each of these test sets, we simulate 5000 trials for each of the two brain states.

Ideally, if a decoding algorithm does not select the discriminant information originating from the RON, its decoding accuracy on TestSet-RON should be at the chance level. On the other hand, if a decoding algorithm is efficient at extracting the discriminant information from the ROI, its decoding accuracy on TestSet-ROI should be sufficiently high. In other words, a good ROI decoding algorithm should yield high decoding accuracy on TestSet-ROI but low decoding accuracy on TestSet-RON.

B. Simulation Results

For testing and comparison purposes, five different decoding algorithms are implemented:

1) LDA based on whole-head MEG (LDA-W): The MEG data recorded from all gradiometers are used for decoding.
2) LDA based on selected channels (LDA-S): The MEG data recorded from 36 gradiometers located on top of the ROI are used for decoding. The locations of the selected channels are shown in Fig. 1(b).
3) Source-space discriminant analysis (SDA): MNE [23] is applied to estimate the sources for each trial. Thereafter, only the sources within the ROI are used for decoding. Since thousands of sources within the ROI are found by MNE, Fisher criterion (FC) is applied to select the important sources as features [24], where the number of selected features is determined by leave-one-out cross validation [19] based on the training set. With these selected features, LDA is applied for classification.
4) Beamspace discriminant analysis (BDA): A linear subspace is first determined by the beamspace transformation [14]–[16] to extract the signals originating from the ROI. Next, LDA is applied to this linear subspace for decoding.
5) ROI-constrained discriminant analysis (RDA): The proposed RDA algorithm is applied for decoding.

Fig. 2(a) and (b) shows the decoding accuracy of five different algorithms based on TestSet-ROI and TestSet-RON, respectively. Here, the decoding accuracy is defined as the number of correctly classified trials divided by the total number of trials. In Fig. 2, the decoding accuracy is plotted as a function of λ (i.e., the power ratio $p_\Phi / p_{\bar{\Phi}}$ defined in (13)) for both BDA and RDA. As discussed in Section II, the parameter λ varies between the lower bound and the upper bound of the constraint function in (16).

A number of important observations can be made from Fig. 2. First, LDA-W achieves similar decoding accuracy for TestSet-ROI and TestSet-RON. It implies that the discriminant information used by LDA-W originates from both the ROI and the RON. Namely, without taking into account the ROI constraint, LDA-W attempts to extract the optimal discriminant pattern from all cortical regions.

By using the MEG channels on top of the ROI only, LDA-S is able to select the discriminant information originating from the ROI. Compared to the decoding accuracy of LDA-W, the
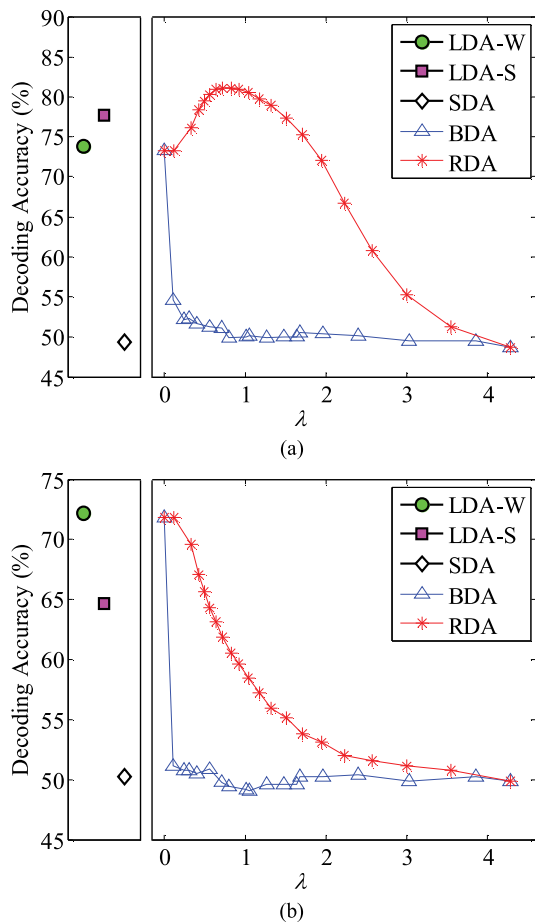

Fig. 2. Decoding accuracy (i.e., the number of correctly classified trials divided by the total number of trials) is shown for different algorithms based on: (a) TestSet-ROI and (b) TestSet-RON. The accuracy of RDA and BDA is shown as a function of λ (i.e., the power ratio $p_\Phi / p_{\bar{\Phi}}$ defined in (13)).

accuracy of LDA-S increases for TestSet-ROI and decreases for TestSet-RON. However, the decoding accuracy of LDA-S remains high for TestSet-RON, implying that LDA-S still takes advantage of the discriminant information from the RON.

SDA maps the MEG signals into the source space, and then uses the sources within the ROI for decoding. As shown in Fig. 2, the decoding accuracy of SDA is extremely low in our simulation experiments due to two reasons. First, it is difficult to accurately estimate the sources inside the ROI for each single trial. It is well known that the problem of source localization is profoundly underdetermined. Most source localization algorithms (e.g., MNE [23]) can only capture the "dominant" sources, but not necessarily the "discriminant" sources inside the ROI. If the discriminant sources within the ROI are not accurately estimated, we cannot achieve high decoding accuracy by applying SDA. Second, source localization often results in thousands of sources that are considered as the input features for decoding, while our training data are limited. In this case, it is extremely difficult to prevent the decoder from overfitting the training data, even though we have carefully applied feature selection in the experiments. For these reasons, SDA cannot efficiently extract the discriminant pattern inside the ROI to achieve high decoding accuracy.

Next, we consider our proposed RDA algorithm. As shown in Fig. 2, when λ reaches the lower bound, the decoding accuracy of RDA is 73.7% and 72.1% for TestSet-ROI and TestSet-RON, respectively. It implies that when the ROI constraint is completely ignored, the discriminant pattern extracted by RDA contains the discriminant information originating from both the ROI and the RON. In this extreme case, RDA is exactly equivalent to LDA-W. As a result, the decoding accuracy is almost identical for TestSet-ROI and TestSet-RON. As λ increases, the decoding accuracy of TestSet-RON decreases, indicating that a reduced amount of discriminant information originating from the RON is used for decoding. In other words, RDA can smoothly control the ROI selectivity of the decoder by varying the value of λ. It is important to note that as λ increases, the extracted discriminant information from the ROI does not monotonically increase. Here, the decoding accuracy for TestSet-ROI is not a monotonic function of λ, because RDA cannot perfectly remove the discriminant information originating from the RON without distorting the information originating from the ROI.

Finally, we study the results of BDA. BDA uses beamspace transformation [14]–[16] to extract the signals generated by the ROI. Similar to RDA, beamspace transformation uses the power ratio $p_\Phi/p_{\tilde{\Phi}}$ in (13) to control the ROI selectivity. However, unlike RDA that formulates a constrained optimization problem in (16), beamspace transformation cannot directly set the ratio $p_\Phi/p_{\tilde{\Phi}}$ to a specific value. Let $\{\varphi_i; i = 1, 2, \ldots, M\}$ denote the eigenvalues of the matrix $\mathbf{G}_{\tilde{\Phi}}^{-1} \cdot \mathbf{G}_\Phi$ arranged in descending order. Beamspace transformation selects the K dominant eigenvectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_K\}$ associated with the K largest eigenvalues $\{\varphi_1, \varphi_2, \ldots, \varphi_K\}$. Next, the ROI constraint $p_\Phi/p_{\tilde{\Phi}}$ is approximated as a linear subspace spanned by these K dominant eigenvectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_K\}$. In this case, the exact value of $p_\Phi/p_{\tilde{\Phi}}$ is unknown but its lower bound is set by φ_K (i.e., the Kth largest eigenvalue). As K increases, the lower bound of $p_\Phi/p_{\tilde{\Phi}}$ decreases. In other words, BDA increases the ROI selectivity by reducing the number of selected dominant eigenvectors. By varying K from M to 1, we plot the decoding accuracy of BDA as a function of λ in Fig. 2. Here, λ is the lower bound of the power ratio $p_\Phi/p_{\tilde{\Phi}}$, instead of the exact value of $p_\Phi/p_{\tilde{\Phi}}$.

Similar to RDA, the decoding accuracy of BDA quickly decreases for TestSet-RON as λ increases. However, comparing the decoding accuracy between RDA and BDA for TestSet-ROI, we notice that RDA offers superior accuracy over BDA in this example. Note that the constraint set of (16) is nonconvex, instead of a simple linear subspace. When BDA is applied, it approximates the nonconvex constraint set of (16) by a linear subspace. Such an approximation does not fully cover the actual constraint set of (16). In other words, BDA attempts to decode the brain states by searching a smaller constraint set than RDA. It, in turn, results in a suboptimal decoder and is less accurate than RDA.

IV. EXPERIMENTAL STUDIES

In this section, we will present the MEG data collected from two human subjects performing both overt and imagined wrist movements to compare the proposed RDA algorithm with sev-

TABLE I
NUMBER OF SUCCESSFUL TRIALS PER MOVEMENT DIRECTION FOR EACH HUMAN SUBJECT

Subject ID	Sub1	Sub2
Overt	84	150
Imagined	174	126

eral other traditional methods. All procedures are approved by the Institutional Review Boards of Carnegie Mellon University and the University of Pittsburgh. All experiments are performed in accordance with the approved protocol.

A. Experimental Setup and Data Preprocessing

In our experiments, two healthy subjects perform a center-out movement task with their right wrist while holding an MEG-compatible joystick [8]. Target images are generated by a computer and projected to a nonmagnetic screen in front of the subjects to prompt the movement direction. Two different experimental paradigms are designed: 1) overt movement and 2) imagined movement. During the overt movement trials, the subjects are instructed to move a cursor from the center of the screen to one of the four locations (i.e., up, down, left or right) by making wrist movements (i.e., radial deviation, ulnar deviation, flexion and extension), while keeping the rest of the body relaxed. During the imagined movement trials, the subjects are instructed to imagine making the wrist movements to the target displayed on the screen, while the cursor moves from the center to the target automatically. For both overt and imagined movements, the subjects are instructed to keep their gaze at the center of the screen and only attend to the targets using their peripheral vision.

Several studies in the literature suggest that a significant amount of information regarding the wrist movement direction can be found from the contralateral sensorimotor cortex [25]–[27]. Hence, we aim to decode the movement direction of the right wrist from left sensorimotor cortex in this study. On the other hand, our experimental paradigm involves visual stimuli. Due to this reason, the visual cortex may generate neural signals that can distinguish different movement directions. To focus on the left sensorimotor cortex and remove the discriminant signals from other cortical regions, the left sensorimotor area (i.e., the left precentral gyrus and the left postcentral gyrus) is set as the ROI and the other regions are set as the RON, as shown in Fig. 1(a).

MEG data are acquired by using a 306-channel whole-head MEG system (Elekta Neuromag) with 1-kHz sampling frequency. The signals recorded by 204 gradiometers are used for movement decoding. Other signals recorded by 102 magnetometers are not used due to their low signal-to-noise ratio. In addition, electrooculography (EOG) is used to monitor eye blinks and eye movements. Electromyography (EMG) of wrist flexor and extensor muscles is recorded to make sure that no movement happens during the imagined sessions. All trials with EOG or EMG contamination are rejected. Table I summarizes
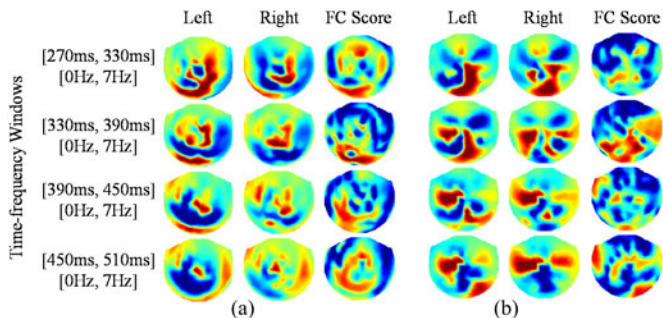
Fig. 3.　Spatial distributions of wavelet coefficients are shown for (a) first set of 102 gradiometers and (b) second set of 102 gradiometers. In both plots, the color maps in the first two columns are calculated by averaging the wavelet coefficients over all trials of the left and right movement directions, respectively, and the color maps in the last column show the scores calculated by Fisher criterion (FC). Each row corresponds to the wavelet coefficients associated with the same time-frequency window. Red color indicates large value and blue color indicates small value. All plots of wavelet coefficients share the same color scale, and all plots of FC scores share the same color scale.

the number of successful trials per movement direction for each subject.

We consider the ROI decoding problem for the left and right directions only, since the RDA algorithm is formulated for two-class classification in this paper. It has been demonstrated in the literature that significant power modulation related to movement directions can be observed in the low-frequency band (≤ 7 Hz) during a short time window [5]. For these reasons, we only consider the low-frequency band (≤ 7 Hz) for the time window $t \in$ [270 ms, 510 ms], where $t = 0$ ms represents target onset. We apply discrete wavelet transform (DWT) with second-order Symlet wavelet function [28] to decompose the MEG signals from each channel and each trial to multiple levels. The DWT results in four wavelet coefficients within the selected frequency band and time window for each channel. These four wavelet coefficients correspond to four time windows $t \in$ [270 ms, 330 ms], [330 ms, 390 ms], [390 ms, 450 ms] and [450 ms, 510 ms] and cover the low-frequency band (≤ 7 Hz). Taking the overt case of Sub1 as an example, Fig. 3 shows the spatial distributions of the wavelet coefficients for two different classes over four time-frequency windows. In this section, we investigate the ROI decoding problem for each time-frequency window separately. In Section V, we will further discuss how to extend the proposed RDA algorithm to simultaneously use multiple time-frequency windows for decoding.

Since there are two subjects, two experimental paradigms (i.e., overt and imagined movements) for each subject and four time-frequency windows for each paradigm, we have 16 different datasets in total. To simplify the notation, the datasets associated with the ith ($i = 1, 2$) subject and the jth ($j = 1, 2, 3, 4$) time-frequency window are labeled as "Sub$i$ Oj" and "Sub$i$ Ij" for overt and imaged movements, respectively. For instance, "Sub1 O1" refers to the first subject and the first time-frequency window. For each dataset, the feature vector is 204-dimensional, corresponding to the wavelet coefficients calculated for 204 gradiometers.

During a separate visit, each subject is asked to perform a standard structural MRI scan of the head. The MRI data are

used to define the vertices and the associated current dipoles to formulate the ROI constraint. Toward this goal, the MRI data are coregistered with the MEG data. A fixed source grid with 2-mm spacing is generated from the gray/white matter boundary of the MRI image by using FreeSurfer [20]. In addition, the cortical surface is parcellated and labeled as different cortical regions by using FreeSurfer.

B. ROI Decoding Results

As discussed in Section II, the performance of ROI decoding should be evaluated according to two criteria: 1) decoding accuracy and 2) ROI selectivity. For the simulation experiments, we assess these two criteria based on two different test sets: TestSet-RON and TestSet-ROI. However, it is not possible to generate these test sets for the MEG measurement data. Hence, we evaluate the decoding accuracy by using six-fold cross-validation [19] and assess the ROI selectivity by using the discriminant pattern source localization (DPSL) method [29]. Considering the spatial filter $y = \mathbf{w}^T \cdot \mathbf{x}$ in (1), DPSL calculates the gain of the spatial filter for the neural source (i.e., a current dipole) at each spatial location to determine whether the MEG signals generated by this particular source can pass the spatial filter. The resulting gain is normalized between 0 and 1. A large gain implies that the MEG signals generated by the corresponding source can pass the spatial filter and, hence, strongly contribute to the decoding result. In the ideal case of ROI decoding, the neural sources with nonzero gain should sit inside the ROI only.

Fig. 4 shows the decoding accuracy of five different algorithms that we implement: 1) LDA-W, 2) LDA-S, 3) SDA, 4) BDA, and 5) RDA. The detailed descriptions for these five methods can be found from our simulation studies in Section III-B and the same algorithm setup is used for the experimental studies here. The decoding accuracy is defined as the number of correctly classified trials divided by the total number of trials.

In order to fully understand the proposed RDA algorithm and compare it with the traditional BDA method, the decoding accuracy is plotted as a function of λ (i.e., the power ratio $p_\Phi / p_{\bar{\Phi}}$ defined in (13)) for both BDA and RDA. Furthermore, the spatial locations of the discriminant sources (i.e., the sources that generate discriminant signals) are estimated by DPSL and are plotted for LDA-W, LDA-S and RDA in Fig. 5. Since SDA and BDA are expected to have good ROI selectivity, we do not show the DPSL results for SDA and BDA here.

Studying Figs. 4 and 5 reveals several important observations. First, while LDA-W achieves high decoding accuracy as shown in Fig. 4, it completely ignores the ROI constraint. Therefore, as shown in Fig. 5, the discriminant sources corresponding to LDA-W are widely distributed over several cortical regions. Such results are observed, because our experimental paradigm involves both visual stimulus and overt/imagined movements and, hence, multiple cortical regions are activated. Without considering the ROI constraint, LDA-W picks up the discriminant information from all these cortical regions for decoding.

By using the MEG channels located on top of the ROI, LDA-S is able to constrain the discriminant sources close to the ROI.
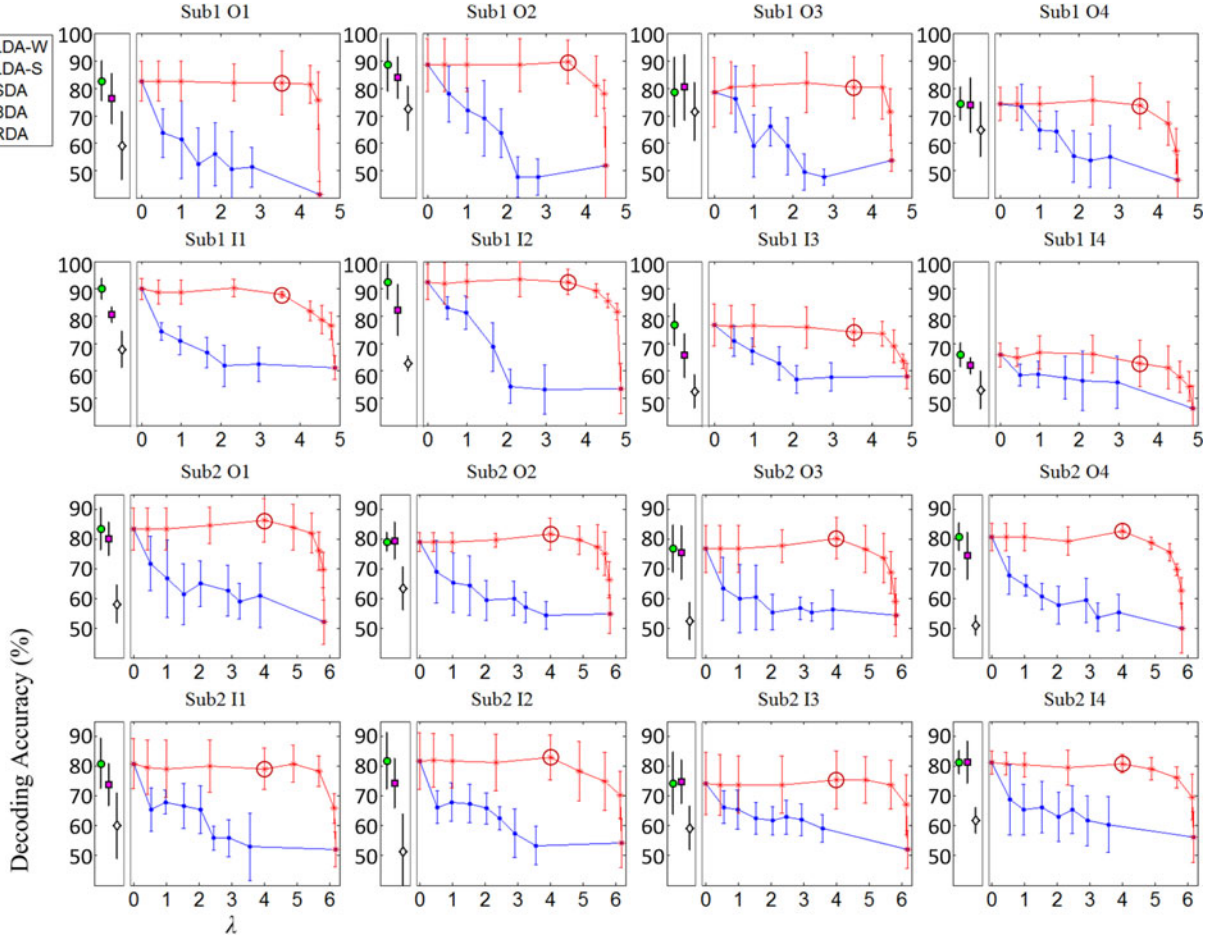
Fig. 4. Movement decoding accuracy (both mean and standard deviation) is plotted for LDA-W, LDA-S, SDA, BDA and RDA. The decoding accuracy is defined as the number of correctly classified trials divided by the total number of trials. For BDA and RDA, the accuracy is plotted as a function of λ (i.e., the power ratio $p_\Phi/p_{\bar\Phi}$ defined in (13)). The plots associated with the *i*th ($i = 1, 2$) subject and the *j*th ($j = 1, 2, 3, 4$) time-frequency window are labeled as "Sub*i* O*j*" and "Sub*i* I*j*" for overt and imaged movements, respectively. Each red circle denotes a reasonably large value of λ where the discriminant information from the RON is substantially removed and the decoding accuracy of RDA is sufficiently high.

However, as shown in Fig. 5, LDA-S fails to constrain the discriminant sources within the ROI in most cases. In addition, because LDA-S only uses a limited number of channels for decoding, its accuracy is not as high as that of LDA-W, as shown in Fig. 4. These results demonstrate that directly selecting a subset of MEG channels on top of the ROI is not an efficient approach to extract the discriminant information for ROI decoding.

As shown in Fig. 4, the decoding accuracy of SDA is extremely low. SDA fails to achieve high decoding accuracy due to several reasons. First, it is difficult to accurately estimate the sources inside the ROI for each single trial. Second, source localization often results in thousands of neural sources that are considered as the input features of decoding, while our training data are limited. In this case, it is extremely difficult to prevent the decoder from overfitting the training data. For these reasons, even though SDA can offer perfect ROI selectivity, it cannot efficiently extract the discriminant pattern inside the ROI to achieve high decoding accuracy.

Next, we consider our proposed RDA algorithm and compare it with other traditional methods. As shown in Fig. 4, the decoding accuracy of RDA is plotted as a function of λ (i.e., the power ratio $p_\Phi/p_{\bar\Phi}$ defined in (13)). If λ reaches the lower

bound, RDA and LDA-W are exactly equivalent. On the other hand, if λ reaches the upper bound, the decoding accuracy drops significantly. In this extreme case, the spatial filter $y = \mathbf{w}^T \cdot \mathbf{x}$ formed by RDA is determined by the dominant eigenvector of $\mathbf{G}_{\bar\Phi}^{-1} \cdot \mathbf{G}_\Phi$, which maximizes the total signal power originating from the ROI, instead of the discriminant information generated by the ROI. In other words, if λ is set to its upper bound, RDA maximizes the ROI selectivity only and, hence, is unlikely to achieve high decoding accuracy. When λ varies between these two extreme cases, RDA explores the tradeoff between ROI selectivity and decoding accuracy. As λ increases from the lower bound to the upper bound, the ROI selectivity is improved while the decoding accuracy is expected to decrease.

It is also important to note that the decoding accuracy of RDA is almost unchanged, as λ varies between the lower bound and a reasonably large value (e.g., the value denoted by the red circle in Fig. 4). For several datasets such as Sub2 O3 and Sub2 O4, the decoding accuracy even slightly increases as λ increases, due to the following two reasons. First, the discriminant information generated by multiple cortical regions may be redundant. Hence, high decoding accuracy can be achieved by extracting the discriminant information from the ROI, instead of other
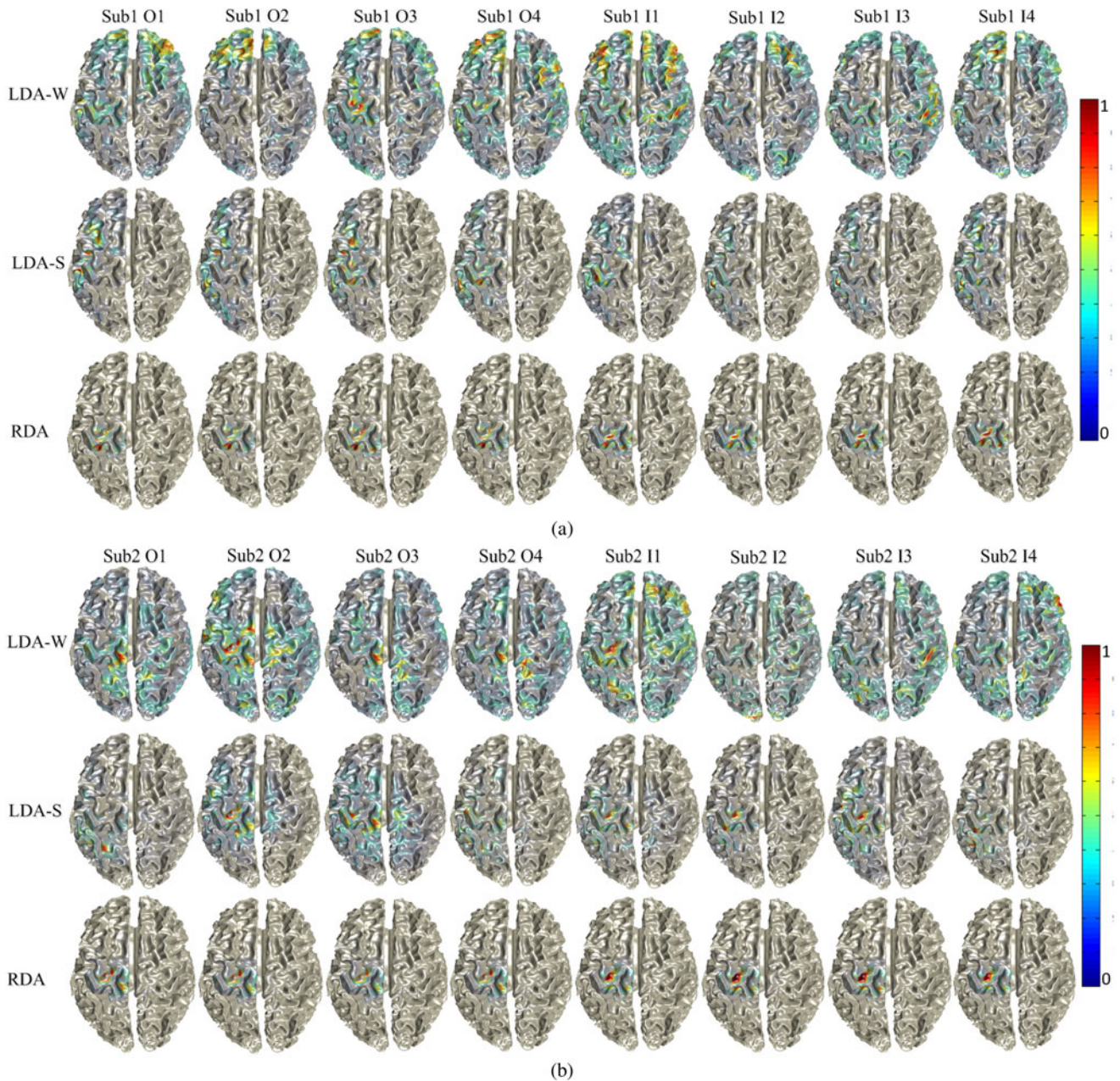
Fig. 5. The discriminant pattern source localization (DPSL) method [29] is applied to find the spatial locations of the discriminant sources for LDA-W, LDA-S and RDA: (a) Subject 1 and (b) Subject 2. The plots associated with the ith ($i = 1, 2$) subject and the jth ($j = 1, 2, 3, 4$) time-frequency window are labeled as "Sub$i$ Oj" and "Sub$i$ Ij" for overt and imaged movements, respectively. Red color indicates the spatial locations where the gain of the spatial filter $y = \mathbf{w}^T \cdot \mathbf{x}$ in (1) is large. In other words, the MEG signals generated by the sources at these locations can pass the spatial filter $y = \mathbf{w}^T \cdot \mathbf{x}$ and contribute to the decoding outcome.

cortical regions. Second, as only a small number of training trials are available from the dataset, the ROI constraint posed by RDA can act as an efficient regularization strategy to avoid overfitting.

The spatial locations of the discriminant sources captured by RDA are shown in Fig. 5, when λ is set to the values denoted by the red circles in Fig. 4. Based on Fig. 5, we observe that when λ is sufficiently large, the proposed RDA algorithm successfully forms a spatial filter to pass the MEG signals generated by the ROI only. In other words, unlike LDA-W and LDA-S where the discriminant sources are widely distributed over multiple

cortical regions, all sources corresponding to RDA appropriately sit inside the ROI, as shown in Fig. 5.

Finally, comparing the decoding accuracy between BDA and RDA in Fig. 4, we notice that these two methods result in the same decoding accuracy, when the value of λ reaches its lower bound or upper bound. Remember that the lower and upper bounds are determined by (17) and (18), respectively. When λ reaches its lower bound, the constraint in (16) is always satisfied for any $\mathbf{w} \in \Re^M$. On the other hand, when λ reaches its upper bound, the constraint set in (16) is a 1-D subspace defined by the dominant eigenvector of $\mathbf{G}_{\bar{\Phi}}^{-1} \cdot \mathbf{G}_{\Phi}$. In these two extreme cases,

BDA and RDA are exactly equivalent and, hence, their decoding results are identical. However, as the value of λ varies between these two bounds, RDA results in superior decoding accuracy over BDA. In these cases, the constraint set of (16) is nonconvex, instead of a simple linear subspace. When BDA is applied, it approximates the nonconvex constraint set of (16) by a linear subspace. Such an approximation does not fully cover the actual constraint set and, hence, results in a suboptimal decoder.

V. DISCUSSION AND FUTURE WORK

The objective of RDA is to extract the MEG signals from prespecified cortical regions for decoding. In the literature, various signal processing algorithms are applied to remove the artifacts that are introduced by nonneural activities (e.g., eye movement [30]). In this paper, we develop the RDA algorithm that aims to remove the neural signals corresponding to the RON. Namely, when we try to study one specific cortical region (i.e., our ROI), we consider the neural signals generated by other cortical regions (i.e., the RON) as artifacts and, hence, should not be used for decoding. For instance, consider the example of a visually-guided motor study where the subject is asked to move the left or right hand by following a visual cue. The visual cue for different hands will activate the visual cortex and the neural signals generated by the visual cortex should be removed, if our interest is to decode the movement information from the motor cortex.

While our simulation and experimental studies both demonstrate the superior performance of RDA over other traditional methods, there remain a number of open questions related to RDA. First, RDA provides a tool that can extract the optimal discriminant pattern for decoding based on the MEG signals from the ROI, where the ROI selectivity is controlled by a user-defined parameter λ. How to appropriately choose the parameter λ to satisfy the requirements for a practical application remains an open question. One possible solution is to plot the decoding accuracy as a function of λ based on the training data, as shown in Fig. 4. Next, we set the minimum decoding accuracy according to the application of interest, and then take the appropriate value of λ to meet the minimum decoding accuracy.

Second, RDA is currently limited to the case where only one feature is extracted from each MEG channel for decoding. However, it is possible to extend RDA to handle the general case with multiple features per MEG channel. For example, if we use the feature extraction method described in Section IV-A, multiple features of the same channel correspond to different time windows. In this case, the linear discriminant function $y = \mathbf{w}^T \cdot \mathbf{x}$ in (1) can be conceptually viewed as a spatio-temporal filter. We can use the same objective function shown in (2), but redefine p_Φ and $p_{\tilde{\Phi}}$ in (13) based on the output of this spatio-temporal filter.

Third, RDA is formulated for two-class classification in this paper. It is possible to extend RDA to multiple classes by integrating it with a multiclass classification algorithm developed by the machine learning community. For example, we can use the one-versus-rest approach, which attempts to build a K-class classifier by combining K two-class classifiers [19].

Finally, the possible applications of the proposed RDA algorithm should be further explored in both clinical and nonclinical domains. For instance, by setting ROI to different cortical regions, RDA may be used to study and compare the functional roles of different cortical regions when performing specific brain tasks. It, in turn, may serve as a powerful tool for presurgical localization of the cortical areas that are significantly modulated by the intended brain task [31].

VI. CONCLUSION

In this paper, we propose a novel RDA algorithm for ROI decoding. RDA aims to extract the optimal discriminant pattern to distinguish different brain states based on the MEG signals originating from prespecified cortical regions. Such a goal is accomplished by formulating a constrained nonconvex optimization problem based on linear classification and beamspace transformation. In addition, an efficient numerical solver is developed to find the solution of RDA with guaranteed global optimum. Compared to other traditional methods, RDA can efficiently extract the discriminant signals from the ROI and suppress the signals originating from the RON, as is demonstrated by the experimental results in this paper.

APPENDIX
NUMERICAL SOLVER FOR RDA

As discussed in Section II, the constraint function of (16) has an upper bound of $\rho(\mathbf{G}_{\tilde{\Phi}}^{-1} \cdot \mathbf{G}_\Phi)$ defined in (18). Therefore, the optimization in (16) is feasible, if and only if λ is less than or equal to $\rho(\mathbf{G}_{\tilde{\Phi}}^{-1} \cdot \mathbf{G}_\Phi)$. If λ equals $\rho(\mathbf{G}_{\tilde{\Phi}}^{-1} \cdot \mathbf{G}_\Phi)$, the globally optimal solution $\mathbf{w}^*$ should be the dominant eigenvector of $\mathbf{G}_{\tilde{\Phi}}^{-1} \cdot \mathbf{G}_\Phi$ corresponding to the eigenvalue λ. Otherwise, if λ is less than $\rho(\mathbf{G}_{\tilde{\Phi}}^{-1} \cdot \mathbf{G}_\Phi)$, the optimization in (16) is strictly feasible. In this section, we will develop a numerical algorithm to find the globally optimal solution $\mathbf{w}^*$ for this strictly feasible case.

Given the fact that $\mathbf{G}_{\tilde{\Phi}}$ is positive-definite, we rewrite (16) as

$$\underset{\mathbf{w}}{\text{minimize}} \quad -\frac{\mathbf{w}^T \cdot \mathbf{S}_B \cdot \mathbf{w}}{\mathbf{w}^T \cdot \mathbf{S}_W \cdot \mathbf{w}}$$
$$\text{subject to} \quad \mathbf{w}^T \cdot (\lambda \cdot \mathbf{G}_{\tilde{\Phi}} - \mathbf{G}_\Phi) \cdot \mathbf{w} \leq 0. \qquad (22)$$

Since the cost function in (22) is invariant with respect to the rescaling of $\mathbf{w}$, we can fix its denominator to an arbitrary constant and optimize the numerator only. It allows us to transform the optimization in (22) to an equivalent problem

$$\underset{\mathbf{w}}{\text{minimize}} \quad -\mathbf{w}^T \cdot \mathbf{S}_B \cdot \mathbf{w}$$
$$\text{subject to} \quad \mathbf{w}^T \cdot \mathbf{S}_W \cdot \mathbf{w} = 1$$
$$\mathbf{w}^T \cdot (\lambda \cdot \mathbf{G}_{\tilde{\Phi}} - \mathbf{G}_\Phi) \cdot \mathbf{w} \leq 0. \qquad (23)$$

As shown by the Proposition 2 of [32], strong duality holds for (23) and its dual problem is expressed as

$$\underset{\alpha, \beta}{\text{maximize}} \quad \alpha$$
$$\text{subject to} \quad -\mathbf{S}_B - \alpha \cdot \mathbf{S}_W + \beta \cdot (\lambda \cdot \mathbf{G}_{\tilde{\Phi}} - \mathbf{G}_\Phi) \geq 0$$
$$\beta \geq 0 \qquad (24)$$

where the constraint $-\mathbf{S}_B - \alpha \cdot \mathbf{S}_W + \beta \, (\lambda \cdot \mathbf{G}_{\tilde{\Phi}} - \mathbf{G}_\Phi) \geq 0$ means that the matrix $-\mathbf{S}_B - \alpha \cdot \mathbf{S}_W + \beta \, (\lambda \cdot \mathbf{G}_{\tilde{\Phi}} - \mathbf{G}_\Phi)$ must be positive semi-definite.

Due to strong duality, the optimal cost function value of (23) is equal to that of (24). In addition, since the two optimization problems (22) and (23) yield the same optimal cost function values, we can conclude that the optimal cost function values of (22)–(24) are all identical. We denote this optimal value as f^*. On the other hand, since the optimization in (24) is a convex semi-definite programming (SDP) problem [33], it can be solved both robustly (i.e., with guaranteed global optimum) and efficiently (i.e., with low computational cost) [33]. Once the optimization in (24) is solved, the optimal cost function value f^* is known and we need to further find the optimal solution $\mathbf{w}^*$ to fully determine the spatial filter in (1).

Given the optimal cost function value f^* of (22), the following inequality holds for any vector $\mathbf{w}$ within the feasible set of (22):

$$\frac{-\mathbf{w}^T \cdot \mathbf{S}_B \cdot \mathbf{w}}{\mathbf{w}^T \cdot \mathbf{S}_W \cdot \mathbf{w}} \geq f^*. \qquad (25)$$

Remember that the matrix $\mathbf{S}_W$ is positive definite and, hence, the scalar $\mathbf{w}^T \cdot \mathbf{S}_W \cdot \mathbf{w}$ is positive. (25) is equivalent to:

$$-\mathbf{w}^T \cdot \mathbf{S}_B \cdot \mathbf{w} - f^* \cdot \mathbf{w}^T \cdot \mathbf{S}_W \cdot \mathbf{w} \geq 0. \qquad (26)$$

The left-hand side of (26) reaches the minimum value 0 or, equivalently, the equality sign is reached for the inequality in (25), when $\mathbf{w}$ is equal to the optimal solution $\mathbf{w}^*$. Therefore, the solution $\mathbf{w}^*$ of (22) can be determined by solving the following optimization problem:

$$\underset{\mathbf{w}}{\text{minimize}} \;\; -\mathbf{w}^T \cdot (\mathbf{S}_B + f^* \cdot \mathbf{S}_W) \cdot \mathbf{w}$$

$$\text{subject to} \;\; \mathbf{w}^T \cdot (\lambda \cdot \mathbf{G}_{\tilde{\Phi}} - \mathbf{G}_\Phi) \cdot \mathbf{w} \leq 0. \qquad (27)$$

The optimal cost function value of (27) is zero and, hence, is invariant with respect to the rescaling of the vector $\mathbf{w}$. Based on this observation, we define

$$\mathbf{w} = \begin{bmatrix} w_1 & w_2 & \cdots & w_{M-1} & 1 \end{bmatrix}^T = \begin{bmatrix} \mathbf{v}^T & 1 \end{bmatrix}^T, \qquad (28)$$

where $\mathbf{v} = [w_1 \; w_2 \; ... \; w_{M-1}]^T$. In other words, since the optimization in (27) aims to find the direction of the vector $\mathbf{w}$, we can fix the last element of $\mathbf{w}$ to be any constant (e.g., 1). Next, we represent the M-by-M symmetric matrix $-(\mathbf{S}_B + f^* \cdot \mathbf{S}_W)$ in the form of four submatrices

$$-(\mathbf{S}_B + f^* \cdot \mathbf{S}_W) = \begin{bmatrix} \mathbf{A}_0 & \mathbf{b}_0 \\ \mathbf{b}_0^T & c_0 \end{bmatrix}, \qquad (29)$$

where $\mathbf{A}_0 \in \Re^{(M-1)\times(M-1)}$, $\mathbf{b}_0 \in \Re^{M-1}$, and $c_0 \in \Re$. Similarly, we represent the matrix $\lambda \cdot \mathbf{G}_{\tilde{\Phi}} - \mathbf{G}_\Phi$ as

$$\lambda \cdot \mathbf{G}_{\tilde{\Phi}} - \mathbf{G}_\Phi = \begin{bmatrix} \mathbf{A}_1 & \mathbf{b}_1 \\ \mathbf{b}_1^T & c_1 \end{bmatrix}, \qquad (30)$$

where $\mathbf{A}_1 \in \Re^{(M-1)\times(M-1)}$, $\mathbf{b}_1 \in \Re^{M-1}$, and $c_1 \in \Re$. Based on (28)–(30), we rewrite the optimization problem in (27) as:

$$\underset{\mathbf{y}}{\text{minimize}} \;\; \mathbf{v}^T \cdot \mathbf{A}_0 \cdot \mathbf{v} + 2 \cdot \mathbf{b}_0^T \cdot \mathbf{v} + c_0$$

$$\text{subject to} \;\; \mathbf{v}^T \cdot \mathbf{A}_1 \cdot \mathbf{v} + 2 \cdot \mathbf{b}_1^T \cdot \mathbf{v} + c_1 \leq 0. \qquad (31)$$

To solve (31), we write the corresponding dual problem [33]

$$\underset{r,\zeta}{\text{maximize}} \;\; r$$

$$\text{subject to} \;\; \zeta \geq 0$$

$$\begin{bmatrix} \mathbf{A}_0 + \zeta \cdot \mathbf{A}_1 & \mathbf{b}_0 + \zeta \cdot \mathbf{b}_1 \\ (\mathbf{b}_0 + \zeta \cdot \mathbf{b}_1)^T & c_o + \zeta \cdot c_1 - r \end{bmatrix} \geq 0 \qquad (32)$$

where $r \in \Re$ and $\zeta \in \Re$ are the optimization variables. The optimization in (32) is a convex SDP problem [33]. Solving it yields the optimal solution $\zeta = \zeta^*$. Once ζ^* is known, the optimal solution $\mathbf{v}^*$ of (31) is given by [33]

$$\mathbf{v}^* = -\left(\mathbf{A}_0 + \zeta^* \cdot \mathbf{A}_1\right)^{-1} \cdot \left(\mathbf{b}_0 + \zeta^* \cdot \mathbf{b}_1\right). \qquad (33)$$

Substituting $\mathbf{v} = \mathbf{v}^*$ into (28) yields the optimal solution $\mathbf{w}^*$ of (22), which is also the optimal solution of (16).

Algorithm 2 summarizes the major steps of our proposed algorithm to find the optimal solution $\mathbf{w}^*$ of (16) with guaranteed global optimum.

Algorithm 2: RDA Solver

1. Start from the optimization problem in (16) where $\mathbf{S}_B$, $\mathbf{S}_W$, $\mathbf{G}_\Phi$, $\mathbf{G}_T$ and λ are known.
2. Calculate $\rho(\mathbf{G}_{\tilde{\Phi}}^{-1} \cdot \mathbf{G}_\Phi)$ in (18).
3. If $\lambda > \rho(\mathbf{G}_{\tilde{\Phi}}^{-1} \cdot \mathbf{G}_\Phi)$, the optimization in (16) is infeasible. If $\lambda = \rho(\mathbf{G}_{\tilde{\Phi}}^{-1} \cdot \mathbf{G}_\Phi)$, the optimal solution $\mathbf{w}^*$ equals the dominant eigenvector of $\mathbf{G}_{\tilde{\Phi}}^{-1} \cdot \mathbf{G}_\Phi$. Otherwise, if $\lambda < \rho(\mathbf{G}_\Phi^{-1} \cdot \mathbf{G}_\Phi)$, follow Steps 4–8 to find the optimal solution $\mathbf{w}^*$.
4. Solve the SDP problem in (24) and get the optimal cost function value f^*.
5. Compute $\mathbf{A}_0$, $\mathbf{b}_0$ and c_0 using (29), and $\mathbf{A}_1$, $\mathbf{b}_1$ and c_1 using (30).
6. Solve the SDP problem in (32) and get the optimal solution ζ^*.
7. Calculate $\mathbf{v}^*$ using (33).
8. Substitute $\mathbf{v} = \mathbf{v}^*$ into (28) to calculate the optimal solution $\mathbf{w}^*$.

REFERENCES

[1] M. Hamalainen *et al.*, "Magnetoencephalography—Theory, instrumentation, and applications to noninvasive studies of the working human brain," *Rev. Mod. Phys.*, vol. 65, no. 2, pp. 413–497, Oct. 1993.

[2] S. Baillet *et al.*, "Electromagnetic brain mapping," *IEEE Signal Process. Mag.*, vol. 18, no. 6, pp. 14–30, Nov. 2001.

[3] B. He and Z. Liu, "Multimodal functional neuroimaging: Integrating functional MRI and EEG/MEG," *IEEE Rev. Biomed. Eng.*, vol. 1, no. 1, pp. 23–40, Nov. 2008.

[4] J. Wolpaw *et al.*, "Brain-computer interface technology: A review of the first international meeting," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 164–173, Jun. 2000.

[5] S. Waldert *et al.*, "Hand movement direction decoded from MEG and EEG," *J. Neurosci.*, vol. 28, no. 4, pp. 1000–1008, Jan. 2008.

[6] N. Montazeri *et al.*, "MEG based classification of wrist movement," in *Proc. IEEE 31st Annu. Int. Conf. Eng. Med. Biol. Soc.*, 2009, pp. 986–989.

[7] M. Gerven *et al.*, "Selecting features for BCI control based on a covert spatial attention paradigm," *Neural Netw.*, vol. 22, no. 9, pp. 1271–1277, Jun. 2009.

[8] J. Zhang *et al.*, "Clustering linear discriminant analysis for MEG-based brain computer interfaces," *IEEE Trans. Rehabil. Eng.*, vol. 19, no. 3, pp. 221–231, Jun. 2011.

[9] S. Foldes *et al.*, "Accessing and processing MEG signals in real-time: Emerging applications and enabling technologies," in *Magnetoencephalography*. Rijeka, Croatia: InTech, 2011.

[10] B. He *et al.*, "Grand challenges in mapping the human brain: NSF workshop report," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 11, pp. 2983–2992, Nov. 2013.

[11] D. Purves *et al.*, *Neuroscience*. Sunderland, MA, USA: Sinauer Associates, 2007.

[12] H. Yuan *et al.*, "Cortical imaging of event-related (de)synchronization during online control of brain-computer interface using minimum-norm estimates in frequency domain," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 16, no. 5, pp. 425–431, Oct. 2008.

[13] M. Ahn *et al.*, "Source space based brain computer interface," in *Proc. 17th Int. Conf. Biomagnetism Advances Biomagnetism*, 2010, pp. 366–369.

[14] J. Gross and A. Ioannides, "Linear transformations of data in space in MEG," *Phys. Med. Biol.*, vol. 44, no. 8, pp. 2081–2097, Aug. 1999.

[15] A. Rodriguez-Rivera *et al.*, "MEG and EEG source localization in beamspace," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 3, pp. 430–441, Mar. 2006.

[16] M. Grosse-Wentrup *et al.*, "Beamforming in noninvasive brain-computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 4, pp. 1209–1219, Apr. 2009.

[17] D. Coyle *et al.*, "A time-frequency approach to feature extraction for a brain-computer interface with a comparative analysis of performance measures," *EURASIP J. Appl. Signal Process.*, vol. 2005, no. 19, pp. 3141–3151, 2005.

[18] B. Graimann *et al.*, "Toward a direct brain interface based on human subdural recordings and wavelet-packet analysis," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 954–962, Jun. 2004.

[19] C. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.

[20] A. Dale *et al.*, "Cortical surface-based analysis: I. Segmentation and surface reconstruction," *Neuroimage*, vol. 9, no. 2, pp. 179–194, Feb. 1999.

[21] G. Golub and C. Loan, *Matrix Computations*. Baltimore, MD, USA: Johns Hopkins Univ. Press, 1996.

[22] R. Desikan *et al.*, "An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest," *Neuroimage*, vol. 31, no. 3, pp. 968–980, Jul. 2006.

[23] M. Hamalainen and R. Ilmoniemi, "Interpreting magnetic fields of the brain: Minimum norm estimates," *Med. Biol. Eng. Comput.*, vol. 32, pp. 35–42, Jan. 1994.

[24] K. Fukunaga, *Introduction to Statistical Pattern Recognition*. New York, NY, USA: Academic, 1990.

[25] D. Moran and A. Schwartz, "Motor cortical representation of speed and direction during reaching," *J. Neurophysiol.*, vol. 82, pp. 2676–2692, 1999.

[26] G. Santhanam *et al.*, "A high-performance brain-computer interface," *Nature*, vol. 442, pp. 195–198, 2006.

[27] W. Wang *et al.*, "Motor cortical representation of position and velocity during reaching," *J. Neurophysiol.*, vol. 97, pp. 4258–4270, 2007.

[28] D. Percival and A. Walden, *Wavelet Methods for Time Series Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2006.

[29] J. Zhang *et al.*, "Task-related MEG source localization via discriminant analysis," in *Proc. 33rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2011, pp. 2351–2354.

[30] J. Kelly *et al.*, "Fully automated reduction of ocular artifacts in high-dimensional neural data," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 3, pp. 598–606, Mar. 2011.

[31] E. Pataraia *et al.*, "Magnetoencephalography in presurgical epilepsy evaluation," *Neurosurg. Rev.*, vol. 25, pp. 141–159, 2002.

[32] A. Beck and M. Teboulle, "A convex optimization approach for minimizing the ratio of indefinite quadratic functions over an ellipsoid," *Math. Program*, vol. 118, no. 1, pp. 13–35, 2009.

[33] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

Jinyin Zhang received the B.S. and M.S. degrees in computer science from the Beijing University of Astronautics and Aeronautics, Beijing, China, in 2002 and 2005, respectively, and the Ph.D. degree in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, USA, in 2013.

Her research interests include signal processing, machine learning, and optimization with specific interests in their application to neural signal processing.

Xin Li (S'01–M'06–SM'10) received the Ph.D. degree in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, USA, in 2005.

From 2009 to 2012, he was the Assistant Director at the FCRP Focus Research Center for Circuit and System Solutions. He is currently an Associate Professor at the Department of Electrical and Computer Engineering, Carnegie Mellon University. His research interests include integrated circuit and signal processing.

Stephen T. Foldes received the Ph.D. degree in biomedical engineering from Case Western Reserve University, Cleveland, OH, USA, in 2011.

He is currently a Postdoctoral Fellow at the Department of Physical Medicine and Rehabilitation, University of Pittsburgh, Pittsburgh, PA, USA. His primary research interests include neurorehabilitation, brain–computer interfaces, neuroimaging, and plasticity after paralysis. He received both the NIH Individual Predoctoral Fellowship (NRSA-F31) in 2010 and the VA Predoctoral Associated Health Rehabilitation Research Fellowship in 2009.

Wei Wang received the medical degree from the Peking University Health Science Center, Beijing, China, in 1999, and the Ph.D. degree in biomedical engineering from Washington University, St. Louis, MO, USA, in 2006.

He is currently an Assistant Professor at the Department of Physical Medicine and Rehabilitation, Department of Bioengineering, and Clinical and Translational Science Institute, University of Pittsburgh, Pittsburgh, PA, USA. His research interests include motor system neurophysiology, brain–computer interface, and neurorehabilitation.

Jennifer L. Collinger received the B.S. and Ph.D. degrees in bioengineering from the University of Pittsburgh, Pittsburgh, PA, USA, in 2003 and 2009, respectively.

She is currently an Assistant Professor at the Department of Physical Medicine and Rehabilitation, University of Pittsburgh. She is also a Research Biomedical Engineer at the VA Pittsburgh Healthcare System at the Department of Veterans Affairs and an Investigator at the Human Engineering Research Laboratories. Her research interests include neurorehabilitation, the study of neuroplasticity, neuroprosthetics, and brain–computer interface technology.

Douglas J. Weber (SM'13) received the Ph.D. degree in biomedical engineering from the Arizona State University, Tempe, AZ, USA.

He is currently an Associate Professor at the Department of Bioengineering, University of Pittsburgh, Pittsburgh, PA, USA. He is also the Program Manager with the Biological Technologies Office at the Defense Advanced Research Projects Agency, Arlington, TX, USA. His research interests include neural engineering, including studies of functional electrical stimulation, activity-based neuromotor rehabilitation, neural coding, and neural control of prosthetic devices.

Anto Bagić received the M.D., M.Sc. and Ph.D. degrees from the University of Zagreb, Zagreb, Croatia, in 1987, 1995, and 2011, respectively.

He is currently an Associate Professor of neurology and neurosurgery at the University of Pittsburgh, Pittsburgh, PA, USA, where he serves as the Chief of Epilepsy Division, the Director of UPMC MEG Epilepsy Program, the Chief Scientific Advisor for MEG Research, the Director of UPMC PUH Epilepsy Monitoring Unit, and the Director of the University of Pittsburgh Comprehensive Epilepsy Center. His clinical practices include epilepsy and clinical applications of MEG, while his research interests include wide spectrum of MEG applications.