Inter-Subject Alignment of MEG Datasets at the Neural Representational Space

Qiong Zhang[a,b], Jelmer P. Borst[c], Robert E. Kass[a,b,d], John R. Anderson[e]

[a] *Machine Learning Department, Carnegie Mellon University*
[b] *Center for the Neural Basis of Cognition, Pittsburgh*
[c] *Department of Artificial Intelligence, University of Groningen*
[d] *Department of Statistics, Carnegie Mellon University*
[e] *Department of Psychology, Carnegie Mellon University*

Abstract

Pooling neural imaging data across subjects requires aligning the recordings that come from different subjects. In magnetoencephalography (MEG) recordings, sensors across subjects are poorly correlated both because of differences in the exact location of the sensors, and structural and functional differences in the brains. It is possible to achieve alignment by assuming that the same regions of different brains correspond across subjects. However, this relies on both the assumption that brain anatomy and function are well correlated, and the strong assumptions that go into solving the inverse problem of source localization. In this paper, we investigated an alternative method that bypasses source-localization. Instead, it analyzes the sensor recordings themselves and aligns their temporal signatures across subjects. We used a multivariate approach, multi-set canonical correlation analysis (M-CCA), to transform individual subject data to a common neural representational space. We evaluated the robustness of this approach using a synthetic dataset where we had ground truth. We demonstrated that M-CCA performs better on an MEG dataset than a method that assumes perfect sensor correspondence and a method that applies source localization. Lastly, we described how the standard M-CCA algorithm could be further improved with a regularization term that incorporates spatial sensor information.

Keywords: Subject alignment, Magnetoencephalography, Neural representational space, Canonical correlation analysis

1. Introduction

In neuroimaging studies, data are frequently combined from many subjects to provide results that represent a central tendency across a population. To achieve this, it is necessary to find an alignment between brain activities recorded from different subjects. In MEG recordings, corresponding sensors across subjects can be poorly correlated both because of differences in the exact location of the sensors[1] (i.e. different head position in the helmet), and structural and functional differences in the brains.

It is possible to achieve alignment by assuming that the same regions of different brains correspond across subjects. Corresponding brain regions are found by first localizing brain sources for each subject. The most commonly used source localization method computes the minimum-norm current estimates (MNE) in an inverse modeling approach (Gramfort et al., 2014). On a separate second step, typically using individual MRIs, the sources for individual subjects are morphed onto sources in a common 'average' brain[2]. The resulting morphed sources are considered to correspond from subject to subject. The validity of this approach depends both on the assumption that brain anatomy and function are well aligned and on the strong assumptions that go into solving the inverse problem of source localization. These steps and assumptions are satisfactory as long as the errors in localization combined with the distortions in morphing, and the subject-specific deviations in timing, are small relative to the effects being investigated.

In this paper, we proposed the use of multi-set canonical correlation analysis (M-CCA) to transform individual subject data to a common neural representational space where different subjects align. The transformation is obtained by maximizing the consistency of different subjects' data at corresponding time points (Kettenring, 1971). Our approach utilizes the rich temporal information that MEG sensor data offers, and circumvents the need to find anatomical correspondence across subjects. This gives many advantages. Firstly, M-CCA does not rely on any assumptions that go into solving the source localization problem. Secondly, M-CCA focuses on capturing distributed patterns of activity that have functional significance, and establishes correspondence across subjects based on these patterns. Therefore, M-CCA

[1] Note that we use 'sensor' to indicate the magnetic coils outside of the head and 'source' to indicate the origin of the measured signal on the cortex.

[2] Typically the standard MNI305 brain is used.

does not rely on the assumption that brain anatomy and functions are well correlated. It simultaneously avoids three types of subject mis-alignment: 1) the mis-alignment given the structural differences of the brains; 2) the sensor location differences during MEG recordings across subjects; 3) the different structural-function mappings in different subjects.

The output of the M-CCA analysis can be used in multiple ways. For example, one might want to find the time windows over which the differences between conditions are detectable in the neural signal (Norman et al., 2006). M-CCA allows one to go from intra-subject classification to inter-subject classification where one group of subjects can be used to predict another. M-CCA may also be used to build a group model that studies the sequence of temporal mental stages in a task (Anderson et al., 2016). Although temporal data from individual subjects are too sparse to be analyzed separately, once the data from all subjects are aligned, they can be combined to improve model parameter estimation. Once important time windows of activity modulation have been identified in such a group model, it is possible to go back to individual subjects and observe the corresponding localized signals.

M-CCA has been applied previously in fMRI studies. Rustandi et al. (2009) demonstrated an application of the M-CCA method to achieve successful prediction across multiple fMRI subjects. Li et al. (2012) also reported an application of M-CCA in integrating multiple-subject datasets in a visuomotor fMRI study, where meaningful CCA components were recovered with high inter-subject consistency. Correa et al. (2010) provided a detailed review of the range of neuroimaging applications that would benefit from the use of M-CCA. This includes not only a group fMRI analysis pooling multiple subjects, but also fusion of data from different neuroimaging modalities (i.e. fMRI, sMRI and EEG). The generalization of M-CCA to MEG data is not straightforward given the very different temporal resolutions of the two imaging modalities. In this paper, we outlined the approach of applying M-CCA on MEG data, and then evaluated it using a synthetic dataset (where we had ground truth). We demonstrated that M-CCA performs better on an MEG dataset than a method that assumes perfect sensor correspondence and a method that applies source localization. We also showed how the standard M-CCA algorithm could be further modified to take into account the similarity of M-CCA sensor mappings of different subjects. A documented package has been made available to preprocess data and apply M-CCA to combine data from different MEG subjects (`https://github.com/21zhangqiong/MEG_Alignment`).

2. Materials and Methods

2.1. MEG Experiment

Twenty individuals from the Carnegie Mellon University community completed the experiment, which was originally reported in Borst et al. (2016). Two subjects were excluded from analysis (one fell asleep and one performed subpar). All were right-handed and none reported a history of neurological impairment. The experiment consisted of two phases: a training phase in which subjects learned word pairs and a test phase in which subjects completed an associative recognition task. The test phase was scheduled the day after the training phase and took place in the MEG scanner. During the test phase subjects had to distinguish between targets which were learned word pairs, and foils which were alternative pairings of the learned words. Therefore, subjects needed to learn both the words and the associative information of the words. Subjects were instructed to respond quickly and accurately. There were four binary experimental factors: probe type (targets or foils), word length (short or long), associative fan (one or two associates), and response hand (left or right hand). Subjects completed a total of 14 blocks (7 with left-handed responses, 7 right-handed), with 64 trials per block. MEG data were recorded with a 306-channel Elekta Neuromag (Elekta Oy) whole-head scanner, which was digitized at 1 kHz, and later down-sampled to 100 Hz. More details of the experiment can be found in the original report of this MEG dataset (Borst et al., 2016).

2.2. Alignment by Correspondence of Brain Sources

We compare M-CCA with two other methods. The first method assumes that the same sensors correspond across subjects. The second method is to perform source localization first, and then assume that the same sources correspond across subjects. In MEG recordings, the measured magnetic signal does not directly indicate the location and magnitude of cortical currents, which can be found by projecting the sensor data onto the cortical surface with minimum norm estimates (MNE). The MNE method attempts to find the distribution of currents on the cortical surface with the minimum overall power that can explain the MEG sensor data (Gramfort et al., 2014). This is done by first constructing 3D cortical surface models from the subjects' structural MRIs using FreeSurfer which are then manually co-registered with the MEG data (Dale et al., 1999; Fischl, 2012). A linear inverse operator is used to project sensor data onto the source dipoles placed on the cortical surface. These source estimates are then morphed onto the standard MNI brain using MNE's surface-based normalization procedure. Source estimates on the standard MNI brain are thought to correspond across subjects. More detailed for obtaining source localization with MNE over the current MEG dataset can be found in the original report (Borst et al., 2016).

2.3. Alignment at the Neural Representational Space

This section outlines the method this paper proposes to align MEG subjects using functional information. M-CCA is used to find the optimal transformation for each subject from the activity of 306 sensors to a common neural representational space, where the inter-subject correlations of the transformed data are maximized across subjects.

2.3.1. M-CCA

We first illustrate the simplest case where we look for correspondence over datasets from two subjects instead of many subjects. Let $X_1 \in \mathbb{R}^{T\times m_1}$ and $X_2 \in \mathbb{R}^{T\times m_2}$ be datasets from two subjects, with the same number of time points T, and data dimensions m_1 and m_2, respectively. Both X_1 and X_2 have mean 0 for each column. The objective in canonical correlation analysis (CCA) is to find two vectors $h_1 \in \mathbb{R}^{m_1\times 1}$ and $h_2 \in \mathbb{R}^{m_2\times 1}$ such that after the projection $y_1 = X_1 h_1$ and $y_2 = X_2 h_2$, y_1 and y_2 are maximally correlated. This is equivalent to:

$$\arg\max_{h_1,h_2} \rho = \frac{y_1^T y_2}{\|y_1\|\|y_2\|} = \frac{h_1^T R_{12} h_2}{\sqrt{h_1^T R_{11} h_1 h_2^T R_{22} h_2}} = h_1^T R_{12} h_2, \text{ where } R_{ij} = X_i^T X_j.$$

There are N solutions to $h^{(i)} = [h_1^T, h_2^T]^T$ obtained collectively in a generalized eigenvalue problem with $i = 1, \ldots, N$, subject to the constraints $h_1^T R_{11} h_1 = h_2^T R_{22} h_2 = 1$ (Borga, 1998). This results in N dimensions (each referred as a CCA 'component', similar to a 'component' in principal component analysis) in the common neural representational space with the transformed data $Y_1 = [y_1^{(1)}, y_1^{(2)}, \ldots, y_1^{(N)}]$ and $Y_2 = [y_2^{(1)}, y_2^{(2)}, \ldots, y_2^{(N)}]$. The value of N does not exceed the smaller of m_1 and m_2. The resulting CCA components in the common representational space are ranked in a decreasing order of the between-subject correlations. The earlier CCA components are the more important ones and the later components can be removed. In other words, canonical correlation analysis finds the shared low-dimensional representation of data from different subjects.

M-CCA is an extension of CCA which considers more than 2 subjects. The objective is similar to before, but now it needs to maximize the correlations between every pair of subjects (i.e. inter-subject correlations) simultaneously. Let $X_k \in \mathbb{R}^{T \times m_k}$ with $k = 1, \ldots, M$ be datasets from M subjects ($M > 2$), each with mean 0 for all columns. The objective in M-CCA is to find M vectors $h_k \in \mathbb{R}^{m_k \times 1}$, where $k = 1, \ldots, M$, such that after the projection $y_k = X_k h_k$, the canonical variates y_k are maximally pairwise-correlated. The objective function to maximize is formulated as:

$$\begin{aligned} \arg \max_{h_1,\ldots,h_M} \rho &= \frac{1}{M(M-1)} \sum_{k,l=1,k\neq l}^{M} y_k^T y_l \\ &= \frac{1}{M(M-1)} \sum_{k,l=1,k\neq l}^{M} h_k^T R_{kl} h_l, \end{aligned}$$

where $R_{kl} = X_k^T X_l$, and $\frac{1}{M}\sum_{k=1}^{M} h_k^T R_{kk} h_k = 1$. The solution is given by solving a generalized eigenvalue problem (Vía et al., 2007). This formulation is an approximation but not an exact maximization of the pairwise correlations, given the complexity of the problem when $M > 2$. It is equivalent to the Maximum Variance (MAXVAR) generalization of CCA proposed by Kettenring (1971). See the proof of this equivalence in (Vía et al., 2005). Other ways of formulating the objective function in M-CCA yield similar results (Li et al., 2009).

2.3.2. Using the Temporal Resolution of MEG

fMRI studies have limited temporal resolution within a single trial. To have enough temporal information to align multiple datasets, M-CCA is typically applied to either a long stream of continuous scans or a series of trials where there is unique temporal information for each trial (Li et al., 2012; Rustandi et al., 2009). In contrast to fMRI, MEG has rich temporal information with millisecond resolution. This enables the application M-CCA over a much briefer period of time, as inter-subject correlations can be calculated based on temporal changes within a trial. When multiple trials are available, as is typically the case for MEG experiments, one can average multiple trials

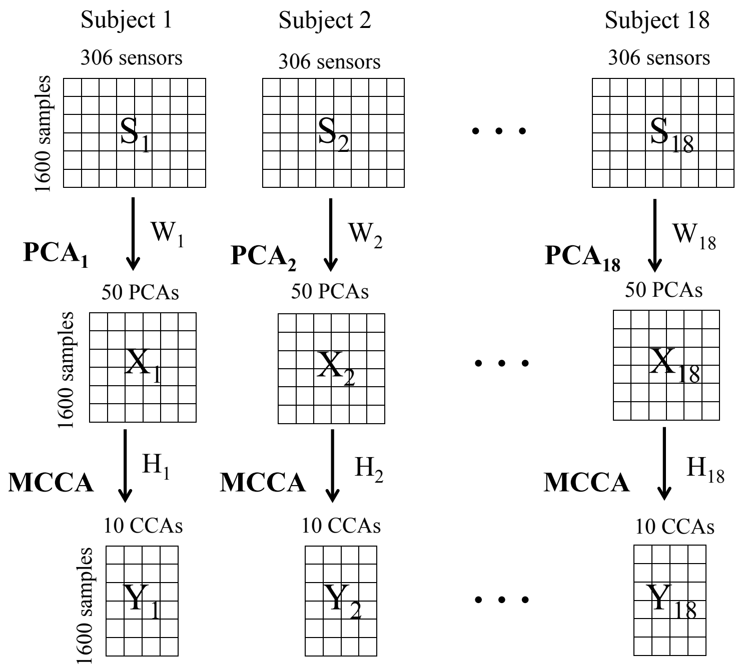

Figure 1: An illustration to demonstrate an application of the M-CCA procedure to an MEG dataset of 18 subjects. S_k is the averaged data (across all trials for each condition) from 306 sensors for subject k, each with 1600 time points; X_k has 50 PCA components for subject k, each with 1600 time points; Y_k has 10 CCA components, each with 1600 time points; $W_1, W_2, \ldots, W_{18}$ are PCA weights obtained for each subject independently; and $H_1, H_2, \ldots, H_{18}$ are CCA weights obtained jointly from all subjects by maximizing all of the inter-subject correlations.

to obtain a highly reliable representation of the change in sensor activity. This is similar to obtaining event-related potential waveforms in the EEG literature (Picton et al., 2000). However, trials are also quite variable in their duration, and temporal alignment is lost when a time sample is further away from stimulus presentation or response emission. Therefore, we only average the samples from the first half second (50 samples given the sampling rate) and the last half second (another 50 samples) of a trial in our application of M-CCA. This averaging process is repeated for each condition, as we potentially have different latent components for different conditions after averaging. With 4 binary factors (probe type, word length, associative fan, response hand), we have 16 conditions. Essentially, over a MEG dataset, the transformation of subject data in M-CCA is obtained by maximizing the consistency of different subjects both in response to different temporal points in a trial, and in response to different experimental conditions. With 100 samples per condition after averaging over trials, we have a 1600x306 matrix S_k from 306 sensors for each of 18 subjects. To reduce dimensionality and remove subject-specific noise, we perform a spatial PCA to obtain the top 50 components on these matrices for each subject first, instead of applying M-CCA directly to the sensor data S_k. This results in 18 matrices of dimension 1600 x 50, which are the inputs X_k our M-CCA analysis for subjects $k = 1, 2..18$. As is illustrated in Figure 1, W_k are PCA weights for subject k which are obtained independently for each subject. H_k are CCA weights for subject k which are obtained jointly from all subjects resulting in common dimensions $Y_k = X_k H_k$ that are maximally correlated across subjects. 10 CCA components are retained for each subject (see results session for discussion of how to choose the number of CCA components).

2.3.3. M-CCA with a Regularization Term

To further improve the application of M-CCA on subject alignment, we also take into consideration the spatial sensor information across different subjects. Projection weights that map sensor data to a given CCA component should be similar across different subjects, given that the misalignment potentially resulted from small spatial shifts during MEG recordings or anatomical variation across subjects. We enhance this similarity (captured as the correlations of weights across subjects) by adding a regularization term in the M-CCA algorithm. This term is useful in situations where there are multiple projection weights that give rise to similar results at a given CCA dimension due to the highly correlated sensor activities. Adding a regularization term makes sensor-to-CCA projection weights similar across subjects while simultaneously maximizing the inter-subject correlations of transformed data. This leads to more interpretable and unique projection weights, and potentially improves the resulted CCA dimensions under certain scenarios. Let $X_k \in \mathbb{R}^{T \times m_k}$, where $k = 1, \ldots, M$, be PCA components obtained independently for each of the M subjects ($M > 2$), T be the number of time points, and m_k be the number of PCA components. Let $W_k \in \mathbb{R}^{S \times m_k}$, where $k = 1, \ldots, M$, be the subject-specific PCA weights (sensor-to-PCA projection) from M subjects, and S be the number of MEG sensors. The modified

M-CCA is formulated as:

$$\arg\max_{h_1,\ldots,h_M} \rho = \frac{1}{M(M-1)} \sum_{k,l=1,k\neq l}^{M} (h_k^T R_{kl} h_l + \lambda h_k^T R'_{kl} h_l)$$
$$= \frac{1}{M(M-1)} \sum_{k,l=1,k\neq l}^{M} h_k^T (R_{kl} + \lambda R'_{kl}) h_l,$$

where $R_{kl} = X_k^T X_l$, $R'_{kl} = W_k^T W_l$, and $\frac{1}{M}\sum_{k=1}^{M} h_k^T (R_{kk} + \lambda R'_{kk}) h_k = 1$. Essentially, the inter-subject correlations of the transformed data $X_k h_k$ are over T time points, whereas the inter-subject correlations of the sensor-to-CCA mapping $W_k h_k$ are over S sensors. That explains why the maximization of the two components in the objective function takes a very similar format, with only the weights differing. One can vary the emphasis given to the regularization term via the weight λ. $\lambda = 0$ corresponds to the M-CCA algorithm without regularization, and very large values of λ correspond to using the same weight maps for all subjects. The solution to the M-CCA problem with regularization can be obtained in the same way as the standard M-CCA by solving a generalized eigenvalue problem.

3. Results

3.1. Synthetic Dataset

How well M-CCA recovers the common dimensions across subjects depends on how noisy the data are and how consistent the individual subjects are. Therefore, we test the robustness of the M-CCA algorithm by varying the noise level and individual differences in a synthetic dataset. We also test the improvement of the M-CCA algorithm when regularization is added. The constructed synthetic dataset can be described as follows: assume that all subjects share the same set of common underlying brain sources $S_0 \in \mathbb{R}^{T\times n_0}$, where T is the number of time points and n_0 is the number of common sources. Each subject also has a set of n_1 sources $S_1 \in \mathbb{R}^{T\times n_1}$ that is unique to each individual. A concatenation of S_0 and S_1 gives the overall underlying brain sources S' for each subject. Note that the number of common sources n_0 is the same for every subject, but n_1 can be set differently for each subject. The diagram below describes the sequence of generating sensor data from these sources and then applying M-CCA:

$$S' \xrightarrow{V_k} sensor \xrightarrow{W_k} PCAs \xrightarrow{H_k} CCAs$$

There is a unique mapping $V_k \in \mathbb{R}^{(n_0+n_1)\times n_2}$ for activity from source level to sensor level for each subject k which introduces the dimension misalignment across subjects, where (n_0+n_1) is the number of overall sources and n_2 is the number of sensors. Principal component analysis is done within each subject k, which gives a unique weight matrix $W_k \in \mathbb{R}^{n_2\times n_3}$ for each subject. This weight matrix transforms data from n_2 sensors ($S'V_k$) to n_3 PCAs ($S'V_kW_k$). The subsequent step is M-CCA after which there is a unique weight matrix $H_k \in \mathbb{R}^{n_3\times n_4}$ for each subject that transforms data from n_3 PCAs ($S'V_kW_k$) to n_4 CCAs ($S'V_kW_kH_k$). The obtained CCAs are not an exact recovery of S_0, but a linear transformation of S_0. In this section, we measure the performance of subject alignment by how well the obtained CCA components correlate across subjects, and the performance of source recovery by how well S_0 can be expressed as a linear combination of the obtained CCA components.

3.1.1. Noiseless Scenarios

In this section, we demonstrate that M-CCA aligns data from individual subjects in a simple and noiseless scenario. To investigate M-CCA in a situation like the MEG dataset, we use the average of the first 10 CCA components from the real MEG dataset as the common sources S_0. These common sources then map onto 306 sensors, using weight maps V_k that are smooth, and similar from subject to subject. Smoothness is created by randomly selecting one of the sensors (called the origin sensor) to have the maximum weight (i.e. 1) for a particular source. Then we assign sensors that are further away from the origin sensor with smaller weights (given by a Gaussian distance function). Similarity is created by having the origin sensor in each subject be a small shift from those of other subjects. Figure 2 shows representative weight maps V_k from the first five subjects ($k = 1, 2, \ldots, 5$) for the first common source in S_0. There are 18 subjects, 1600 time points ($T = 1600$), 10 common sources ($n_0 = 10$), 0 individual sources ($n_1 = 0$), 306 sensors ($n_2 = 306$), 10 PCA components ($n_3 = 10$) and 10 CCA components ($n_4 = 10$).

We can compare the averaged inter-subject correlations over the top 10 PCA components with the averaged inter-subject correlations over the top 10 CCA components. After the M-CCA step, the brain activity from different

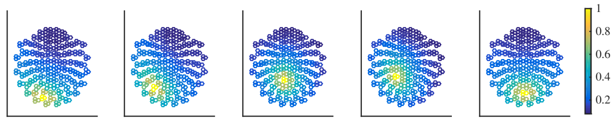

Figure 2: Weight maps V_k that map from the first common source in S_0 to sensors in the first five subjects $k = 1, 2, \ldots, 5$.

subjects over CCA dimensions (0.9598) are more correlated across subjects than the brain activity over PCA dimensions (0.4078) prior to the M-CCA step. Obtained CCA components do not correspond exactly to the original common sources S_0, but we can examine how well the original common sources are expressed as a linear combinations of the CCA components. A least-squares solution to Q is obtained from the system of equations $YQ = S_0$, where Q transforms the obtained CCA components from half of the subjects Y to the original common sources S_0 (the other half Y' is used to test for overfitting). The correlation matrix between $Y'Q$ and S_0 is calculated, with the closest correspondence for each of the 10 components in S_0 being the component in $Y'Q$ with the largest absolute correlation. The mean of these absolute correlations is 0.9980.

3.1.2. Sensor Noise and Individual Differences

The second scenario is built up upon the first one, but with sensor noise and individual differences added. We look at how the performance of M-CCA in aligning subjects is affected by sensor noise and individual differences. To model trial-to-trial noise, we generate 50 trials for each subject (in the MEG dataset around 50 trials are averaged per condition per subject prior to M-CCA), and add white Gaussian noise to each sensor at each time point with a signal-to-noise ratio (SNR) of -1dB, which is very conservative compared with values in the literature (Gonzalez-Moreno et al., 2014). Individual differences are introduced by having additional brain sources, S_1, that are unique to each subject besides the common sources S_0 (which, again, are the obtained CCA components from the real MEG dataset). The unique sources, S_1, are each modeled as a sum of 10 sinusoidal waves with their frequencies and powers sampled randomly. Weight maps V_k are generated to be smooth, and similar from subject to subject. There are 18 subjects, 1600 time points ($T = 1600$), 10 common sources ($n_0 = 10$), 40 individual sources per subject

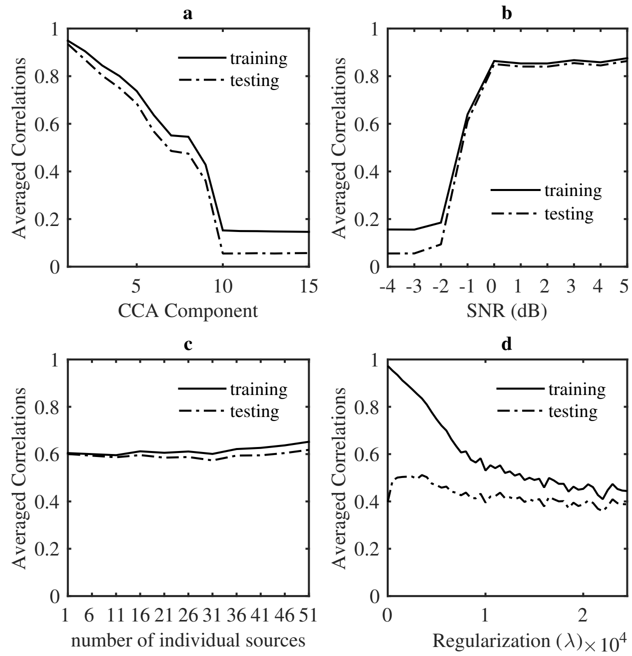

Figure 3: Averaged inter-subject correlations for each of the 15 CCA components (a); averaged inter-subject correlations averaged over the first 10 CCA components while increasing levels of SNR (b), the number of individual brain sources n_0 (c), or when $T = 60$ the amount of regularization (d). Whenever not specified, $T = 1600$. SNR $= -1.30$dB. $n_0 = 10$. $n_1 = 40$. $n_2 = 306$. $n_3 = 50$. $n_4 = 15$. Each data point was averaged over 10 simulations.

($n_1 = 40$), 306 sensors ($n_2 = 306$), 50 PCA components ($n_3 = 50$) and 15 CCA components ($n_4 = 15$). Figure 3a shows that, with noise and individual differences added, reasonable inter-subject correlations can be recovered up to the 9$^{\text{th}}$ CCA component. Additionally, inter-subject correlations beyond 10 CCA components are no longer significant, which reflects the fact that there are only 10 underlying common sources simulated in the first place.

The M-CCA algorithm is designed to recover only the common sources across subjects, and leaves the remaining signal that is unique to each subject in the late CCA components with very low inter-subject correlations. We can again examine how well the original common sources are expressed as a linear combinations of the CCA components. A least-squares solution to Q is obtained from the system of equations $YQ = S_0$, where Y stores the

CCA components averaged across half of the subjects (the other half Y' is used to test for overfitting). The correlation matrix between $Y'Q$ and S_0 is calculated, with the closest correspondence for each of the 10 components in S_0 being the component in $Y'Q$ with the largest absolute correlation. The mean of these absolute correlations is 0.8767, compared with 0.9980 obtained previously without sensor noises nor individual differences.

To further test the robustness of the M-CCA algorithm, we vary the level of SNR while fixing the rest of the parameters (Figure 3b), and vary the number of additional individual sources n_1 while fixing the rest of the parameters (Figure 3c). As SNR increases, the averaged inter-subject correlations of the first 10 CCA components also increases. The SNR threshold beyond which CCA components can be reasonably recovered falls into the range -2dB to -1dB. On the other hand, the addition of brain sources that are unique to individuals does not affect the recovery of the shared components during M-CCA. It can be observed that as the number of the additional individual sources grows, the averaged inter-subject correlations over the CCA components stay the same on the testing data, even though there is slightly more overfitting over the training data.

3.1.3. The effect of regularization

The synthetic data considers the case where subjects share similar weight maps. Under this assumption, we test if adding regularization to M-CCA improves its performance. In particular, we are interested in the situations where data is limited to 60 samples ($T = 60$) with small signal-to-noise ratio (SNR $= -1.30$dB), and where M-CCA has difficulty recovering unique weight maps. While the inter-subject correlations decrease with increasing regularization for the training data, this is not the case for the testing data. Rather, it increases first before decreasing (Figure 3d). Putting some emphasis to consider the similarity between sensor weight maps helps improve the inter-subject correlations of the resulted CCA components.

3.2. MEG Dataset

3.2.1. Application of M-CCA

The M-CCA procedure described earlier (Figure 1) is applied to the MEG dataset. To test for overfitting, M-CCA is applied to half of the data (even trials for each subject) as the training data to obtain the projection weights H_k for each subject k. The same projection weights are then applied to the other half of the data (odd trials for each subject) as the testing data. Inter-subject correlations in the testing data reflect how well the CCA components truly capture the underlying data. Figure 4 shows the averaged inter-subject correlations for each of the first 20 CCA components in training data (solid) and testing data (dashed). The obtained CCA weights generalize well from the training data to the testing data up to around 10 CCA components. As a result, 10 CCA components are retained for the rest of the analysis.

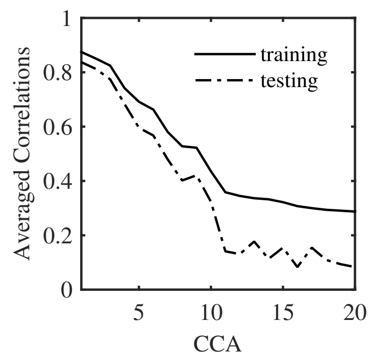

Figure 4: Averaged inter-subject correlations of the first 10 CCA components over training data (solid) and testing data (dashed) for the MEG dataset.

Figure 5 plots the first 10 CCA components over the 100 samples (the first 50 are stimulus-locked, and the last 50 are response-locked) for one condition. There is a considerable match between the CCA components over the training data (blue) and testing data (magenta). Since the obtained CCA components are in the order of decreasing averaged inter-subject correlations, the earlier CCA components are more important than the later ones. Intuitively, the CCA components can be taken as the basis functions of which every sensor is a linear combination. The more important CCA components capture low-frequency temporal information, and the later CCA components capture patterns of high-frequency oscillations (mostly in the first 50 samples).

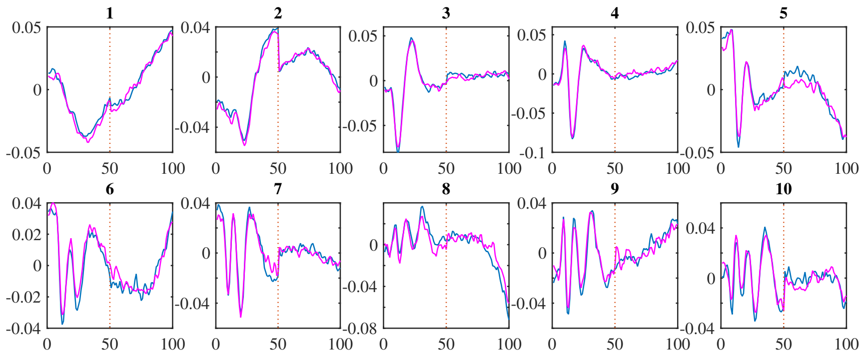

Figure 5: The first 10 CCA components over the first 100 samples (1 sec) averaged across all 18 subjects, with the first 50 stimulus-locked and the last 50 response-locked (separated by the red dashed line) over the training data (blue) and testing data (magenta)

3.2.2. Evaluation by Inter-subject Classification

To further evaluate subject alignment using M-CCA, we investigate how well we could use the transformed data to classify different experimental conditions of the data from one subject given the data from other subjects. Such inter-subject classification typically has lower accuracy than intra-subject classification, which is to use a classifier trained from a subset of data from the same subject. There are at least two factors that contribute to this. The first is mis-alignment of sensors across subjects and individual differences in brain anatomy. The second is functional differences in how subjects perform the task, including differences in the brain regions that are involved in the task, and the speed with which they perform the task. The goal of M-CCA is to correct the first issue. M-CCA is also robust to functional differences in cases where there are additional unique brain regions involved for each individual subject, as demonstrated in the synthetic dataset. The effect of differences across subjects in speed is minimized by using only the very early/late portions of the trial in M-CCA, which are associated with encoding and motor response, respectively.

We compare **inter-subject classification** using M-CCA with three other alternatives. The first alternative is **intra-subject classification** using a PCA on the subjects' sensor data, which is similar to the method used in Borst et al. (2016). This method just forgoes the challenge of finding a correspondence across subjects. Second, we consider **inter-subject classification** using a PCA on the subjects' sensor data. Third, we consider

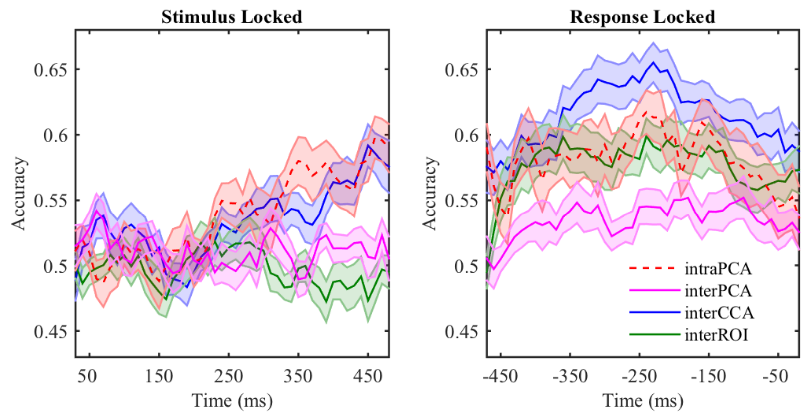

Figure 6: Intra-subject classification results of fan condition over PCA components of sensor data (red). Inter-subject classification results of fan condition over PCA components of sensor data (magenta), CCA components (blue), and PCA components of ROIs (red). Times on the x axis are relative to the stimulus (left) or the response (right). SEMs are shown in shaded error bars with $n = 18$.

inter-subject classification using a PCA performed on source data, which has been localized using MNE and aligned based on the subjects' anatomy. For each classification, we perform M-CCA on 100 time points for each combination of conditions excluding the condition to classify (i.e. 800 x 306 matrices instead of the 1600 x 306 matrices in Figure 1). Averaging over the dimension to be classified in M-CCA makes sure that the obtained CCA dimensions are only for maximizing the alignment across subjects, but have not learned any specific representations about the condition to classify. Classification of fan and word length are considered in the evaluation[3].

Figure 6 and 7 show the intra-subject classification accuracy over 10 PCA components of sensor data (intraPCA: red), the inter-subject classification accuracy over 10 PCA components of the sensor data (interPCA: magenta), the inter-subject classification accuracy over 10 CCA components (interCCA: blue), and the inter-subject classification accuracy over 10 PCA components of the source data (interROI: green). Linear discriminant analysis is used for

[3]Probe type condition is not included due to inferior intra-subject classification performance. Trials responded by left hand and right hand are so different that they have to be distinguished in the CCA step in order to have reasonable functional alignment. As a result, left and right hand condition cannot be taken as the condition to classify.

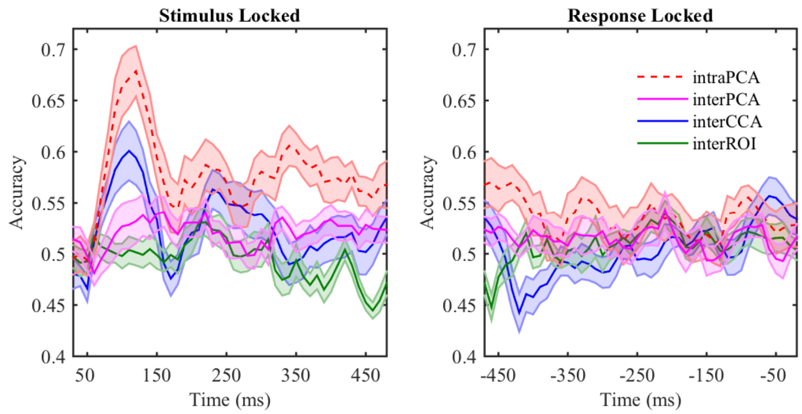

Figure 7: Intra-subject classification results of word length condition over PCA components of sensor data (red). Inter-subject classification results of word length condition over PCA components of sensor data (magenta), CCA components (blue), and PCA components of ROIs (green). Times on the x axis are relative to the stimulus (left) or the response (right). SEMs are shown in shaded error bars with $n = 18$.

classifying data averaged over a sliding window of 5 samples, over instances each averaged over 10 trials[4]. In both Figure 6 and 7, application of M-CCA improves inter-subject classification. In fan condition (Figure 6), using CCA components has comparable classification accuracy to that of intra-subject classification, both over the stimulus-locked data and the response-locked data. Inter-subject classification accuracy using CCA components is also consistently superior to that of PCA components over the sensor data or source data. In the stimulus locked data, accuracy is above chance only after 350 ms when encoding of the words has been completed. In the response-locked data, accuracy is above chance throughout but reaches a peak about 200 ms before response generation. In word length condition (Figure 7), using CCA components is the only inter-subject method to achieve above-chance accuracy even if it is inferior to intra-subject PCA. Both the intra-subject PCA and the inter-subject M-CCA find the period of high classifiability to be early when the words are being encoded.

[4]None of the classifiers have satisfactory performance over single trials, thus an effective comparison among the classifiers is not possible.

3.2.3. Regularization

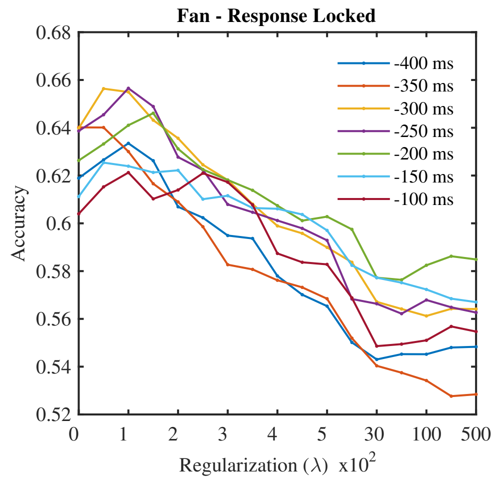

Figure 8: Inter-subject classification accuracies of fan condition over selected time windows of response-locked data with increasing λ values.

A regularization term in the M-CCA algorithm incorporates spatial sensor information by assuming similar projection weights across different subjects. Increasing λ corresponds to putting more emphasis on obtaining similar weight maps, compared with obtaining more correlated CCA components across subjects. Regularization has the potential to further improve subject alignment. We evaluate this by examining the performance of inter-subject classification when λ is increasing. Focusing on the most classifiable period, Figure 8 shows the effect of the λ on inter-subject classification of fan condition during response-locked period. Each curve represents classification accuracy in a 50 millisecond time window centered around different times (e.g. the line marked as -100 ms refers to the inter-subject classification results during the window -125 to -75 ms). On each curve, $\lambda = 0$ corresponds to the same classification performance using CCA components as in Figure 6. With increasing regularization, accuracy does not decrease right away and even shows a slight improvement.

The best λ value is one that is as large as possible, but does not decrease the classification accuracy much. Large λ values will give rise to more interpretable sensor weight maps. From Figure 8 we can tell that the best range of λ values is around 100–200. Figure 9 shows the sensor weights that map

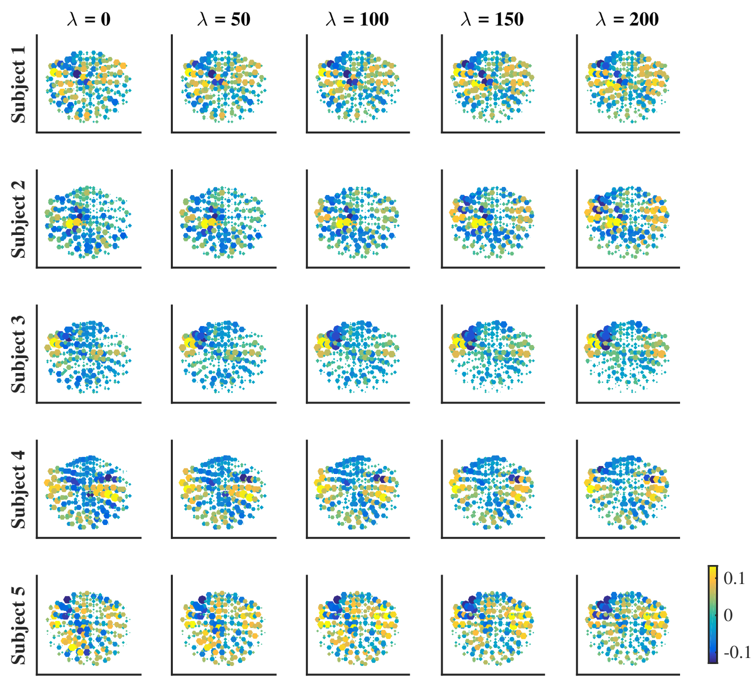

Figure 9: Projection weights from sensors to the first CCA component for the first five subjects with different degrees of regularization ($\lambda = 0, 50, 100, 150, 200$).

sensor data of the first 5 subjects to the first CCA component. Different columns correspond to different λ values ranging from 0–200. With increasing regularization, projection weight maps become more consistent across subjects.

4. Discussion

We evaluated M-CCA as a method for pooling MEG data from different subjects together. It successfully produced dimensions in the common representational space over which brain activities from different subjects were well correlated, the patterns of which also generalized well over the unseen half of the data.

Subject alignment was evaluated in an inter-subject classification task, where different conditions of the data in one subject were classified based on a classifier trained on the rest of the subjects. Inter-subject classification performance using M-CCA was close to that of intra-subject classification performance over sensor data, supporting the conclusion that MCCA succeeds in finding meaningful common dimensions. Inter-subject classification based directly on the sensors never did much better than chance even though intra-subject classification based on sensors did well. This is in accordance with our knowledge that sensors do not align identically with respect to the recording device in MEG data. Inter-subject classification based on source activity also performed poorly. It is likely that the quality of subject alignment using source localization was compromised by the strong assumptions that went into solving the inverse problem of source localization. M-CCA eliminates the need to make those assumptions by directly producing dimensions that align across subjects.

We also examined the performance of M-CCA in aligning data from multiple subjects in a synthetic dataset. Reasonable inter-subject correlations were obtained when the SNR was larger than -1 dB. Individual differences were introduced by adding additional unique brain sources for each subject in the task. M-CCA successfully recovered dimensions that were shared across subjects. As a result, the addition of brain sources that were unique to individuals did not affect the recovery of the shared components during M-CCA.

In addition to the standard M-CCA algorithm, we added a regularization term to give better alignment and more interpretable results. M-CCA involves finding a unique mapping for each subject from the sensors to the common dimensions. The regularization term introduces a spatial constraint to impose inter-subject similarity on these sensor weight maps. This adds an appropriate constraint if the subject mis-alignment is a result of small amount of spatial shifts in either sensor positions or anatomical brain regions from subject to subject. In a synthetic dataset where this kind of misalignment was present, we showed that regularization improved the recovery of the underlying sources. Over the real MEG dataset, adding the regularization term also improved the inter-subject classification performance, and produced more consistent and interpretable sensor weight maps across subjects.

The use of M-CCA in pooling data from different subjects is not restricted to only one neural imaging modality. We demonstrated in this study how M-CCA can be applied to align subjects in MEG that comes in at a fine temporal grain size. In this aspect, EEG (Electroencephalogram) and ECoG

(Electrocorticography) datasets share very similar temporal characteristics as MEG, and can utilize M-CCA to pool subjects together in the same way. In addition, it is possible to pool subjects recorded from different imaging modalities, as long as there are corresponding time points across subjects and that brain responses over that period of the time are considered consistent across subjects.

5. Acknowledgement

This research was supported by the National Science Foundation grant 1420009, the James S. McDonnell Foundation Scholar Award 220020162, and the Office of Naval Research Grant N00014-15-1-2151 to J.R.A. We thank Matthew M. Walsh and Aryn Pyke for their helpful comments on a draft of the paper.

References

Anderson, J. R., Zhang, Q., Borst, J. P., & Walsh, M. M. (2016). The discovery of processing stages: Extension of sternberg's method. *Psychological review*, *123*, 481–509.

Borga, M. (1998). *Learning Multidimensional Signal Processing*. Ph.D. thesis Linkoping University.

Borst, J. P., Ghuman, A. S., & Anderson, J. R. (2016). Tracking cognitive processing stages with meg: A spatio-temporal model of associative recognition in the brain. *NeuroImage*, *141*, 416 – 430.

Correa, N. M., Adali, T., Li, Y.-O., & Calhoun, V. D. (2010). Canonical correlation analysis for data fusion and group inferences: Examining applications of medical imaging data. *IEEE signal processing magazine*, *27*, 39–50.

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. segmentation and surface reconstruction. *NeuroImage*, *9*, 179 – 194.

Fischl, B. (2012). Freesurfer. *NeuroImage*, *62*, 774 – 781.

Gonzalez-Moreno, A., Aurtenetxe, S., Lopez-Garcia, M.-E., del Pozo, F., Maestu, F., & Nevado, A. (2014). Signal-to-noise ratio of the MEG signal after preprocessing. *Journal of Neuroscience Methods*, *222*, 56 – 61.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Parkkonen, L., & Hämäläinen, M. S. (2014). MNE software for processing MEG and EEG data. *NeuroImage*, *86*, 446 – 460.

Kettenring, J. R. (1971). Canonical analysis of several sets of variables. *Biometrika*, *58*, 433–451.

Li, Y.-O., AdalÄś, T., Wang, W., & Calhoun, V. D. (2009). Joint blind source separation by Multi-set Canonical Correlation analysis. *IEEE transactions on signal processing : a publication of the IEEE Signal Processing Society*, *57*, 3918–3929.

Li, Y.-O., Eichele, T., Calhoun, V. D., & Adali, T. (2012). Group study of simulated driving fMRI data by multiset canonical correlation analysis. *Journal of Signal Processing Systems*, *68*, 31–48.

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fmri data. *Trends in Cognitive Sciences*, *10*, 424 – 430.

Picton, T., Bentin, S., Berg, P., Donchin, E., Hillyard, S., Johnson, R., Miller, G., Ritter, W., Ruchkin, D., Rugg, M., & Taylor, M. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, *37*, 127–152.

Rustandi, I., A., J. M., & Mitchell, T. M. (2009). Integrating multiple-study multiple-subject fMRI datasets using canonical correlation analysis.

Vía, J., Santamaría, I., & Pérez, J. (2005). Canonical correlation analysis (CCA) algorithms for multiple data sets: Application to blind simo equalization.

Vía, J., Santamaría, I., & Pérez, J. (2007). A learning algorithm for adaptive canonical correlation analysis of several data sets. *Neural Networks*, *20*, 139 – 152.