An F-ratio-based method for estimating the number of active sources in MEG

Amita Giri[1]*, John C. Mosher[2], Amir Adler[1,3] and
Dimitrios Pantazis[1]

[1]McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA, United States, [2]Department of Neurology, McGovern Medical School, Texas Institute for Restorative Neurotechnologies, UTHealth, Houston, TX, United States, [3]Department of Electrical Engineering, Braude College of Engineering, Karmiel, Israel

Introduction: Magnetoencephalography (MEG) is a powerful technique for studying the human brain function. However, accurately estimating the number of sources that contribute to the MEG recordings remains a challenging problem due to the low signal-to-noise ratio (SNR), the presence of correlated sources, inaccuracies in head modeling, and variations in individual anatomy.

Methods: To address these issues, our study introduces a robust method for accurately estimating the number of active sources in the brain based on the F-ratio statistical approach, which allows for a comparison between a full model with a higher number of sources and a reduced model with fewer sources. Using this approach, we developed a formal statistical procedure that sequentially increases the number of sources in the multiple dipole localization problem until all sources are found.

Results: Our results revealed that the selection of thresholds plays a critical role in determining the method's overall performance, and appropriate thresholds needed to be adjusted for the number of sources and SNR levels, while they remained largely invariant to different inter-source correlations, translational modeling inaccuracies, and different cortical anatomies. By identifying optimal thresholds and validating our F-ratio-based method in simulated, real phantom, and human MEG data, we demonstrated the superiority of our F-ratio-based method over existing state-of-the-art statistical approaches, such as the Akaike Information Criterion (AIC) and Minimum Description Length (MDL).

Discussion: Overall, when tuned for optimal selection of thresholds, our method offers researchers a precise tool to estimate the true number of active brain sources and accurately model brain function.

KEYWORDS

F-ratio, source localization, Alternating Projection (AP), source enumeration, MEG, AIC, MDL

1. Introduction

Magnetoencephalography (MEG) is a powerful non-invasive neuroimaging technique that offers high temporal resolution for studying human brain function (Hämäläinen et al., 1993; Baillet, 2017). Localization of MEG sources has garnered significant interest in recent years since it can reveal the origins of neural signals and offer valuable insights into the complex workings of the human brain. By identifying the sources of neural activity, researchers can study the underlying mechanisms of cognition, perception, and other brain functions (Ahveninen et al., 2006; Giorgetta et al., 2013; Klepp et al., 2015; Pancholi et al., 2023). Additionally, source localization can aid in diagnosing and studying neurological disorders and identifying abnormal brain activity (Oishi et al., 2002; Westlake et al., 2012; Wilkinson et al., 2020; Xu et al., 2021; Giri, 2022; Giri et al., 2022).

MEG source localization methods typically involve solving an inverse problem, which entails estimating current sources within the brain based on the measured MEG data. This is challenging because the measured signals are influenced by several factors, such as the geometry and conductivity of the head, sensor noise, and the ill-posed nature of the problem. Mathematically, the localization problem can be cast as finding the location and moment of the set of dipoles whose field best matches the M/EEG measurements (Mosher et al., 1992). Localization methods can be broadly categorized into *distributed* and *discrete* solutions.

Distributed source imaging approaches aim to estimate a density map of active dipoles across the entire cortex. Commonly used methods include minimum norm estimators (MNE) (Hämäläinen et al., 1993, 1994), dynamic statistical parametric mapping (dSPM) (Dale et al., 2020), and standardized low-resolution electromagnetic tomography (sLORETA) (Pascual-Marqui et al., 2002). However, these methods assume a significantly larger number of unknown sources in a discrete surface or volumetric grid compared to the number of MEG sensors. The ill-posed nature of the problem poses a significant challenge, especially in the presence of multiple active regions in the brain (Darvas et al., 2004). Non-linear source estimation methods, such as Mixed Norm Estimate (MxNE) (Strohmeier et al., 2016) and time-frequency mixed-norm estimates (TF-MxNE) (Gramfort et al., 2013), address this issue by incorporating $l1$-norm penalty regularizers that favor sparse collections of focal dipolar sources. Other sparse approaches include hierarchical reconstructions of cortical and subcortical sources (Gramfort et al., 2012; Krishnaswamy et al., 2017; Rezaei et al., 2021). While these methods have shown some success, they tend to be computationally demanding and have limited accuracy when dealing with complex multi-dipole configurations.

On the other hand, *discrete multiple dipole localization* methods avoid the ill-posedness associated with distributed methods by finding a small set of equivalent current dipoles (ECDs) whose field best matches the M/EEG measurements in a least-squares sense (Mosher et al., 1992). Dipole localization methods offer a more classical approach to brain source localization and provide more intuitive interpretations of brain activity by estimating the location, orientation and amplitude of neural sources. The most well-known methods are beamformers (Van Veen et al., 1997; Vrba and Robinson, 2001) and MUltiple SIgnal Classification (MUSIC) (Mosher et al., 1992), and their recursive variants Recursively Applied and Projected MUSIC (RAP-MUSIC) (Mosher and Leahy, 1999), Truncated RAP-MUSIC (Mäkelä et al., 2018), and RAP Beamformer (Ilmoniemi and Sarvas, 2019). While recursive variants generally perform better than their non-recursive counterparts, they still face limitations such as reduced effectiveness, reliance on high signal-to-noise ratio (SNR), and potential cancellation of correlated sources. Recent advancements in this field have addressed some of these concerns, including Alternating Projections (AP) (Adler et al., 2022), double-scanning (DS-MUSIC) (Mäkelä et al., 2017; Ilmoniemi and Sarvas, 2019), hemispherical harmonics MUSIC (HSH-MUSIC) (Giri et al., 2018), head harmonics MUSIC (H^2 MUSIC) (Giri et al., 2019), and Flex-MUSIC (Hecker et al., 2023). Estimating the number of independent signal components is a prerequisite for dipole localization methods to accurately estimate dipole

sources. However, determining the correct number of active sources contributing to the recorded signals remains a fundamental challenge in MEG data analysis (Wendel et al., 2009), significantly impacting the success of brain source localization. We focus on addressing this specific problem here.

Estimating the number of active sources in MEG data poses significant challenges due to multiple factors. First, MEG signals generally exhibit a low SNR, which makes it difficult to differentiate between simultaneously active sources. Second, the presence of correlated sources adds complexity by potentially causing multiple sources to be mistakenly identified as a single source. Last, errors in translational head modeling and variations in individual anatomy introduce additional noise and variability, hampering accurate estimation of the location and strength of underlying sources.

Early attempts to estimate the number of dipoles relied on subjective thresholds (Bartlett, 1954; Lawley, 1956; Chen et al., 1991). These approaches involved setting a threshold that separated the eigenvalues of the data covariance matrix from the complete set of eigenvalues. Chen et al. (1991) proposed a method that detected the number of sources by imposing an upper bound on the eigenvalue magnitudes of the correlation matrix derived from the array output. In addition to conventional eigenvalue-based techniques, a few methods have also employed eigenvectors for estimating the number of sources (Di and Tian, 1984; Jiang and Ingram, 2004).

To overcome the limitations of subjective thresholds, two main classes of methods have been developed for estimating the number of signal sources using the distribution of the eigenvalues of the data covariance matrix. The first class involves techniques based on principal component analysis (PCA) (Green et al., 1988; Yao et al., 2018), independent component analysis (ICA) (Ikeda and Toyama, 2000), and factor analysis (Malinowski, 1977a,b). The second class consists of information theoretic approaches (Wax and Kailath, 1985; Knösche et al., 1998), such as the Akaike information criterion (AIC) (Akaike, 1974) and the minimum description length (MDL) (Rissanen, 1978; Schwarz, 1978). The work from Wax and Kailath (1985) derived the eigenvalue forms of AIC and MDL methods, which can be directly applied to array signal processing problems. These methods aim to strike a balance between model fit and complexity using principles from information theory. However, these approaches assume source independence, which is not always valid in the brain. As a result, they tend to perform poorly (Zhang et al., 1989; Chen et al., 1991; Salman et al., 2015; Yao et al., 2018), especially in the presence of correlated sources, noise, low SNR, and limited time samples. Hence, there is a need for more robust and accurate methods to estimate the number of active sources in MEG signals.

To address these limitations, we propose a robust method for accurately estimating the number of active sources in the brain using the F-ratio statistical approach. Our method introduces formal decision criteria that sequentially increase the number of sources in the multiple dipole localization problem until all sources are found. Our method is based on the F-ratio test, which is commonly used in statistics to compare the variances of two samples. It is sensitive to differences in the variances of the samples and can be used to determine whether adding a source to the model significantly improves the fit of the model to the data. The F-ratio

statistical approach allows for a comparison between a full model with a higher number of sources and a reduced model with fewer sources.

We validated the F-ratio-based method on simulated, real phantom, and human MEG data, and compared its performance to that of other state-of-the-art statistical approaches, such as AIC and MDL. We found that the F-ratio-based method outperformed competing methods in terms of accuracy and reliability. One crucial aspect we investigated was the selection of appropriate thresholds for the F-ratio values. We found that this selection played a critical role in determining the overall performance of our method. Through systematic analyses, we identified optimal thresholds that needed to be adjusted according to the number of sources and SNR levels. Importantly, these thresholds exhibited remarkable consistency across different inter-source correlations, translational modeling inaccuracies, and cortical anatomies. When fine-tuned with the optimal selection of thresholds, our F-ratio-based method emerged as a precise and robust tool for estimating the true number of active sources in MEG data.

2. Materials and methods

In this section, we provide a concise overview of the notations used to describe the measurement data, forward matrix, and sources. We also present the problem formulation for estimating multiple ECDs in the brain. Subsequently, we describe the F-ratio statistical procedure, which serves as the foundation to estimate the number of active sources in the brain, and outline the experimental procedures we use to assess performance.

2.1. Measurement model and notations

Consider an array of M MEG sensors detecting signals from Q ECD sources located at positions $\{\mathbf{p}_q\}_{q=1}^{Q}$. At time t, the MEG measurement vector $\mathbf{y}(t)$ can be described as a superposition of the contributions from Q source signals $\{s_q(t)\}_{q=1}^{Q}$ and additive noise:

$$\mathbf{y}(t) = \sum_{q=1}^{Q} \mathbf{l}(\mathbf{p}_q) s_q(t) + \mathbf{n}(t), \quad \text{where} \quad Q < M \quad (1)$$

The topography $\mathbf{l}(\mathbf{p}_q)$ of the qth dipole at location $\mathbf{p}_q$ is defined as $\mathbf{l}(\mathbf{p}_q) = \mathbf{L}(\mathbf{p}_q)\mathbf{o}$, where $\mathbf{L}(\mathbf{p}_q) \in \mathbb{R}^{M \times 3}$ is the lead field matrix and $\mathbf{o} \in \mathbb{R}^{3 \times 1}$ is the orientation vector. Depending on the problem, the orientation vector $\mathbf{o}$ may either be known, referred to as a *fixed-oriented* dipole, or it may be unknown, referred to as *freely-oriented*. Additionally, the measurements are subject to the presence of additive white Gaussian noise, which is represented by $\mathbf{n}(t) \in \mathbb{R}^{M \times 1}$.

Several source localization methods exist for estimating the Q ECD sources, with each dipole source characterized by its location, orientation, and amplitude. Recently, we introduced a method called Alternating Projections (AP) (Adler et al., 2022), which offers several advantages. AP source localization method is robust to forward model errors, can handle high inter-source correlation values, and is effective even in low SNR scenarios. It

is important to note that estimating the true number of active sources Q is a fundamental requirement for all the aforementioned dipole localization methods to accurately estimate the dipole source parameters.

2.2. F-ratio based method

The F-test is a widely used statistical technique that leverages the F-ratio to assess the presence of a significant difference between the variances of two data sets. In the context of determining the true number of sources, this technique holds particular value. It enables us to test the hypothesis that incorporating an additional source results in a substantial improvement in the variance accounted for by the model. By employing the F-test, we can make informed decisions regarding the optimal number of sources to include in order to provide the most accurate explanation for the observed data.

In probability theory and statistics, the F-statistic, also known as the F-ratio, is defined as the ratio between two independent chi-square distributions, denoted as $\mathbf{X_1} \sim \chi^2_{\text{DOF}_1}$ and $\mathbf{X_2} \sim \chi^2_{\text{DOF}_2}$, where DOF_1 and DOF_2 represent the respective degrees of freedom. Mathematically, the F-ratio is expressed as:

$$\mathbf{F} = \frac{\mathbf{X_1}/\text{DOF}_1}{\mathbf{X_2}/\text{DOF}_2} \quad (2)$$

This formula provides a means to calculate the F-ratio by dividing the observed value of $\mathbf{X_1}$ normalized by its degrees of freedom, DOF_1, by the observed value of $\mathbf{X_2}$ normalized by its degrees of freedom, DOF_2. In this study, we use this idea to compare two hypothesized models based on the variance they explain. The first model, referred to as the "reduced model", explains the data with K number of sources:
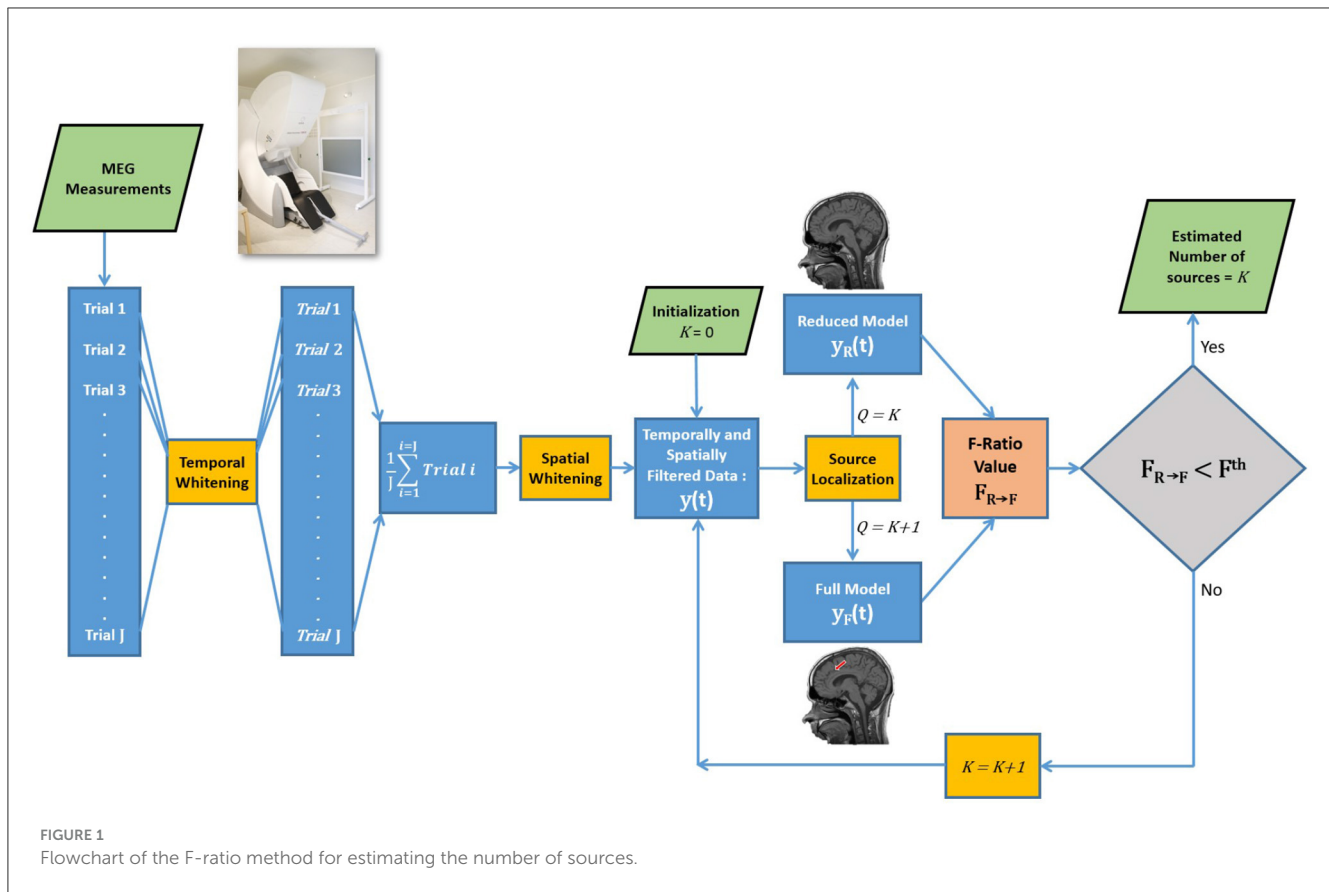
$$\mathbf{y}(t) = \mathbf{y}_{\text{R}}(t) + \mathbf{n}_{\text{R}}(t), \quad (3)$$

$$\mathbf{y}_{\text{R}}(t) = \sum_{q=1}^{K} \mathbf{l}(\mathbf{p}_q) s_{q_{\text{R}}}(t) \quad (4)$$

where $\mathbf{y}_{\text{R}}(t)$ represent the estimated signals and $\mathbf{n}_{\text{R}}(t)$ represents noise of a reduced model. On the other hand, the second model, called the "full model," includes one more source compared to the reduced model, with $K + 1$ sources:

$$\mathbf{y}(t) = \mathbf{y}_{\text{F}}(t) + \mathbf{n}_{\text{F}}(t), \quad (5)$$

$$\mathbf{y}_{\text{F}}(t) = \sum_{q=1}^{K+1} \mathbf{l}(\mathbf{p}_q) s_{q_{\text{F}}}(t) \quad (6)$$

where $\mathbf{y}_{\text{F}}(t)$ represent the estimated signals and $\mathbf{n}_{\text{F}}(t)$ represents noise of a full model. The estimation of $\mathbf{y}_{\text{R}}(t)$ and $\mathbf{y}_{\text{F}}(t)$ signals is achieved by solving an inverse problem using a dipole fitting method. By comparing the residual variance of these two models, we can assess their performance and determine the most appropriate model for the given data. Since we assume that the noise added to the measured signal is white Gaussian noise, it can be deduced that the sum of square errors between the measured signal $\mathbf{y}(t)$ and the estimated signals $\mathbf{y}_{\text{R}}(t)$ and $\mathbf{y}_{\text{F}}(t)$, denoted as

FIGURE 1
Flowchart of the F-ratio method for estimating the number of sources.

$\|\mathbf{y}(t) - \mathbf{y}_{R(t)}\|_2^2$ and $\|\mathbf{y}(t) - \mathbf{y}_F(t)\|_2^2$ respectively, follows chi-square distributions. Therefore, the F-ratio test can be written as Supek and Aine (1993):

$$\mathbf{F}_{R \to F} = \frac{\sum_{t=1}^{N} \|\|\mathbf{y}(t) - \mathbf{y}_R(t)\|\|_2^2}{\text{DOF}_R} \bigg/ \frac{\sum_{t=1}^{N} \|\|\mathbf{y}(t) - \mathbf{y}_F(t)\|\|_2^2}{\text{DOF}_F}$$

(7)

where N is the number of time samples. For the fixed-oriented case, the degrees of freedom (DOF) for the reduced and full models are given by $\text{DOF}_R = MN - (3 + N)K$ and $\text{DOF}_F = MN - (3 + N)(K + 1)$, respectively. For the freely-oriented case, the DOF for the reduced and full models are $\text{DOF}_R = MN - (4 + N)K$ and $\text{DOF}_F = MN - (4+N)(K+1)$, respectively. These formulas account for a total of MN degrees of freedom, with three degrees of freedom deducted for position, one for orientation (in the freely-oriented case), and N for amplitude, for each dipole.

Note that the formulas for estimating the DOF assume independence among all data points, which is not the case in experimental data. To address this issue, in experimental data we implemented a two-step whitening process involving both temporal and spatial filtering. The first step involved temporal whitening, as depicted in the flowchart of the F-ratio method (Figure 1). The MEG data of each trial was whitened temporally using a six-order linear predictive coding (LPC) technique. This method helped alleviate temporal correlations within the data, reducing their impact on the results. Subsequently, the LPC-filtered trials were averaged. The second step was spatial whitening, which was achieved by applying a whitening filter derived from inverting the noise covariance matrix. This step further mitigated inter-dependencies among the data points, enhancing the reliability of the analysis. A more detailed discussion of the two-step whitening process is presented in Section 2.4.

It is important to note that the calculated F-ratio values are influenced by the residuals obtained after dipole fitting, and therefore, are dependent on the chosen source localization method used for solving the ECD localization problem in MEG. In our study, we specifically examined the behavior of the F-ratio statistical procedure when employing the AP localization method. The AP method solves the inverse problem iteratively and sequentially by minimizing the least-squares (LS) criterion. For a more detailed and comprehensive discussion of the AP method, we refer readers to Adler et al. (2022).

To estimate the number of sources, a systematic approach involving the comparison of a reduced model and a full model is illustrated in Figure 1. The formal comparison begins by initializing the reduced model with zero sources and the full model with one source to compute the F-ratio value. The decision regarding the number of sources involves comparing the resulting F-ratio value, denoted as $\mathbf{F}_{0 \to 1}$, with a threshold value, $\mathbf{F}^{th}$. If the reduced model is rejected ($\mathbf{F}_{0 \to 1} > \mathbf{F}^{th}$), indicating evidence of at least one source, the analysis proceeds to compare a reduced model with one source against a full model with two sources, represented as $\mathbf{F}_{1 \to 2}$. This sequential process continues, increasing the number of sources, until reaching a step where the reduced model cannot be rejected, providing an estimation of the true number of sources.

2.3. Performance evaluation with simulations

We evaluated the performance of the F-ratio method in diverse simulated scenarios, considering variations in the number of sources, SNR levels, inter-source correlations, translational modeling errors, and cortical anatomies.

The SNR was defined as the ratio between the Frobenius norm of the signal-magnetic-field spatiotemporal matrix and that of the noise matrix, following the approach described in Sekihara et al. (2001). To quantify inter-source correlation, we employed the Pearson's correlation coefficient. To establish the desired correlation among the sources, we utilized the Cholesky decomposition method. Initially, we generated fundamental cosine signals for each simulated source, with randomized phase and frequencies ranging from 10Hz to 30Hz. Next, we applied the Cholesky decomposition to factorize the symmetric positive definite target correlation matrix into the product of a lower triangular matrix and its conjugate transpose. By multiplying lower triangular matrix with the fundamental cosine signals, we generated a set of correlated dipole waveforms. This procedure ensured that the resulting source time courses closely matched the correlation coefficients specified by the target correlation matrix, thus incorporating the desired inter-source correlations in our simulations.

The sensor array geometry was based on the Megin Triux MEG system, which consists of a 306-channel probe unit with 204 planar gradiometer sensors and 102 magnetometer sensors. The MEG source space geometries were modeled using the cortical manifold extracted from MR data of adult humans, employing Freesurfer (Fischl et al., 2004). In our analysis, we used cortical anatomies from four different adult humans. Simulated sources were restricted to approximately 15,000 grid points distributed over the cortex. The reconstructed sources were estimated on a distinct grid of 50,000 points covering the cortex. To avoid the "inverse crime" problem, where identical parameters are used for data synthesis and inversion in an inverse problem, the simulation and reconstruction grids were non-overlapping with an average distance of 0.7 mm between neighboring points (Colton and Kress, 1998). The forward matrix for both grids was computed using the boundary element method implemented in OpenMEEG (Gramfort et al., 2010) within the BrainStorm software (Tadel et al., 2011). Simulated MEG data was generated by randomly selecting sources from the simulation grids. Gaussian white noise was then added to the MEG sensors to model instrumentation noise at specified SNR levels. In order to evaluate the effect of head model errors, we introduced translations to the reconstruction grid before computing the forward matrix. Last, we employed the AP method to solve the inverse problem in the *fixed oriented* case. All experiments were conducted with 100 Monte Carlo simulations to ensure statistical robustness.

2.4. Performance evaluation with a real phantom

We assessed the performance of the F-ratio method using the *freely-oriented* dipoles model with phantom data provided in the phantom tutorial (Taylor et al., 2016) of the Brainstorm software (Tadel et al., 2011). The phantom experiment was conducted using the Megin Neuromag system, which consists of a 306-channel probe unit with 204 planar gradiometer sensors and 102 magnetometer sensors.

The data comprised MEG recordings obtained from the sequential activation of 32 artificial dipoles. To activate the phantom dipoles, an internal signal generator was used along with an external multiplexer box that connected the signal to each individual dipole. Each dipole was activated 20 times with an amplitude of 200 nAm, resulting in a total of 20 trials for each experimental condition. It is important to note that the chosen amplitude of 200 nAm falls within the range typically observed in inter-ictal spikes associated with epilepsy, as observed in raw data (Oishi et al., 2002).

For each dipole and each trial, the MEG data was whitened temporally using a six-order LPC technique. In particular, baseline data of a 200 ms pre-stimulus interval was used to compute the LPC coefficients of sixth order. These coefficients were then averaged across sensors and subsequently applied to the post-stimulus data modeled as an moving average (MA) filter. The purpose of this step was to eliminate temporal dependencies in the post-stimulus data, as observed in the auto-regressive (AR) model of baseline data. Following this, the LPC-filtered post-stimulus measurements were averaged across the 20 trials. In addition to temporal prewhitening, spatial prewhitening was also performed on the average data using a regularized noise covariance matrix in the Brainstorm software (Tadel et al., 2011). The regularization process included adding an identity matrix scaled to 10% of the largest eigenvalue of the noise covariance matrix.

To simulate the concurrent activation of multiple sources, we combined averaged data from different dipoles since only one dipole could be activated at a time. To introduce variability and avoid perfect coherence, we added a random delay ranging from 0 to 50 ms for each dipole.

The reconstruction source space was defined as a sphere centered within the MEG sensor array, with a radius of 64.5 mm. It was sampled using a regular volumetric grid of points with a resolution of 2.5 mm, resulting in a total of 56,762 grid points. The forward matrix was estimated based on a single sphere head model using the BrainStorm software (Tadel et al., 2011). The performance of the F-ratio method was evaluated using the AP method for localizing dipoles in the *freely-oriented* case.

2.5. Performance evaluation with human MEG data

The effectiveness of the F-ratio method in practical scenarios was assessed using human MEG data recorded from a single human participant during an auditory task. Prior to participation, the subject provided written informed consent, and the study was approved by the local ethics committee (Institutional Review Board of the Massachusetts Institute of Technology), following the principles of the Declaration of Helsinki. During the experiment, binaural sounds (beeps) were delivered to the subject using tubal-insert earphones. These auditory stimuli are known to elicit specific

brain responses that are localized in the bilateral primary auditory cortex. A total of 166 trials were recorded, with an interstimulus interval of 2150 ms between each auditory stimulus. The MEG data were acquired using a MEGIN Triux MEG system, which includes a 306-channel probe unit consisting of 102 magnetometers and 204 planar gradiometers.

The forward matrix was estimated using BrainStorm based on an overlapping spheres head model. The reconstruction source space samples the brain volume in an adaptive way, with a higher density near the surface where we expect a higher spatial resolution due to the proximity to the sensors. The density decreased gradually toward the center of the brain, resulting in a total of 33,073 grid points. This grid is constructed using a specific adaptive algorithm in Brainstorm: It begins with a brain envelope containing 10,000 vertices as the initial number. Then, each previous layer is successively shrunk and downsampled by a factor of 3. This operation is repeated for 17 layers or until no more vertices are available.

The performance of the F-ratio method was evaluated using the AP method for localizing dipoles in the *freely-oriented* case. Before localization, the raw data underwent prewhitening in both the temporal and spatial domains following the same procedure as in the real phantom data.

3. Results

3.1. Optimal modeling of MEG data requires the adjustment of F-ratio nominal thresholds

To investigate the behavior of the F-ratio method under different experimental conditions, we conducted a thorough simulation analysis. Our objective was to assess the accuracy of estimating the true number of sources by varying the threshold values across various experimental scenarios. These scenarios included different numbers of active sources, varying SNR levels, inter-source correlation values, translational modeling errors, and cortical anatomies.

We observed that the accuracy of estimating the true number of sources using a specific F-ratio threshold was strongly influenced by the actual number of sources Q. This relationship is depicted in Figures 2A–C, where the estimation accuracy varied significantly for different values of Q. Notably, as the number of true sources increased, a lower F-ratio threshold was required to achieve higher performance. Similarly, we discovered a strong correlation between the accuracy of estimating the true number of sources and the SNR level. Figures 2D–F demonstrates this dependency for various SNR values. As the SNR level increased, a lower F-ratio threshold became necessary to achieve higher accuracy in estimating the true number of sources.

In contrast, we made the important observation that the optimal F-ratio threshold remained independent of the inter-source correlation level. This finding is illustrated in Figures 2G–I, where we tested different inter-source correlation values (0.1, 0.5, and 0.9). The accuracy in estimating the number of sources peaked at the same threshold value for all correlation levels. This robustness indicates that the optimal F-ratio thresholds

were not influenced by the inter-source correlation values of the active sources. Consequently, researchers may rely on a consistent threshold value regardless of the degree of correlation among the sources, enhancing the practical applicability and reliability of the F-ratio method. We obtained similarly robust results when testing the accuracy of estimating the true number of sources across different model errors (Figures 2J–L) and cortical anatomies (Figures 2M–O). In the case of model errors, we introduced registration errors by translating the reconstruction source space relative to the source simulation space. Specifically, we applied translations of 1 mm posterior (X-axis), right (Y-axis), and upward (Z-axis). Importantly, despite the presence of these registration errors, the F-ratio method remained highly robust. Similarly, when evaluating the F-ratio thresholds across the cortical anatomies of four different adult humans, we observed consistent and robust results.

In summary, our observations indicated that the performance of F-ratio thresholds varied significantly depending on the true number of sources and the SNR levels of the data. However, we found that F-ratio thresholds remained robust across different inter-source correlation values, translational model errors, and cortical anatomies. These findings highlight the need to adapt and optimize threshold procedures for the F-ratio test based on the specific number of sources and SNR levels in the data. In the next section, we determined these optimal thresholds.

3.2. Computation and evaluation of adjusted F-ratio thresholds in simulated data

In this section, we aimed to determine adjusted F-ratio thresholds for accurately estimating the number of active sources in MEG data. To accomplish this, we used a specific cortical anatomy as a reference (referred to as Anatomy 1). Optimal threshold values for Anatomy 1 were computed by identifying the F-ratio value that yielded the highest average accuracy in estimating the number of sources, considering the $\rho \in \{0.1, 0.5, 0.9\}$ inter-source correlations and 1mm translational modeling errors in x, y, z. It may be noted that our study specifically examined the robustness of the method against translational modeling errors as a representative example. The resulting adjusted F-ratio thresholds were obtained for various SNR levels and numbers of sources, as depicted in Figure 3. Our findings indicate that higher threshold values are necessary in scenarios characterized by a high SNR and a low number of sources.

To assess the effectiveness of the adjusted threshold values obtained for the reference anatomy, we conducted tests on three additional cortical anatomies (Figure 4). The performance of the adjusted F-ratio thresholds in estimating the number of sources was evaluated at 0 dB SNR and correlation levels $\rho \in \{0.1, 0.5, 0.9\}$. Remarkably, the results showed that the performance of the adjusted thresholds was comparable to that of the reference cortical anatomy (Anatomy 1). These findings demonstrate the reliability and robustness of the calculated optimal F-ratio thresholds across a wide range of simulation scenarios, including variations in the number of sources, SNR levels, inter-source correlation values, translational modeling errors, and cortical anatomies.
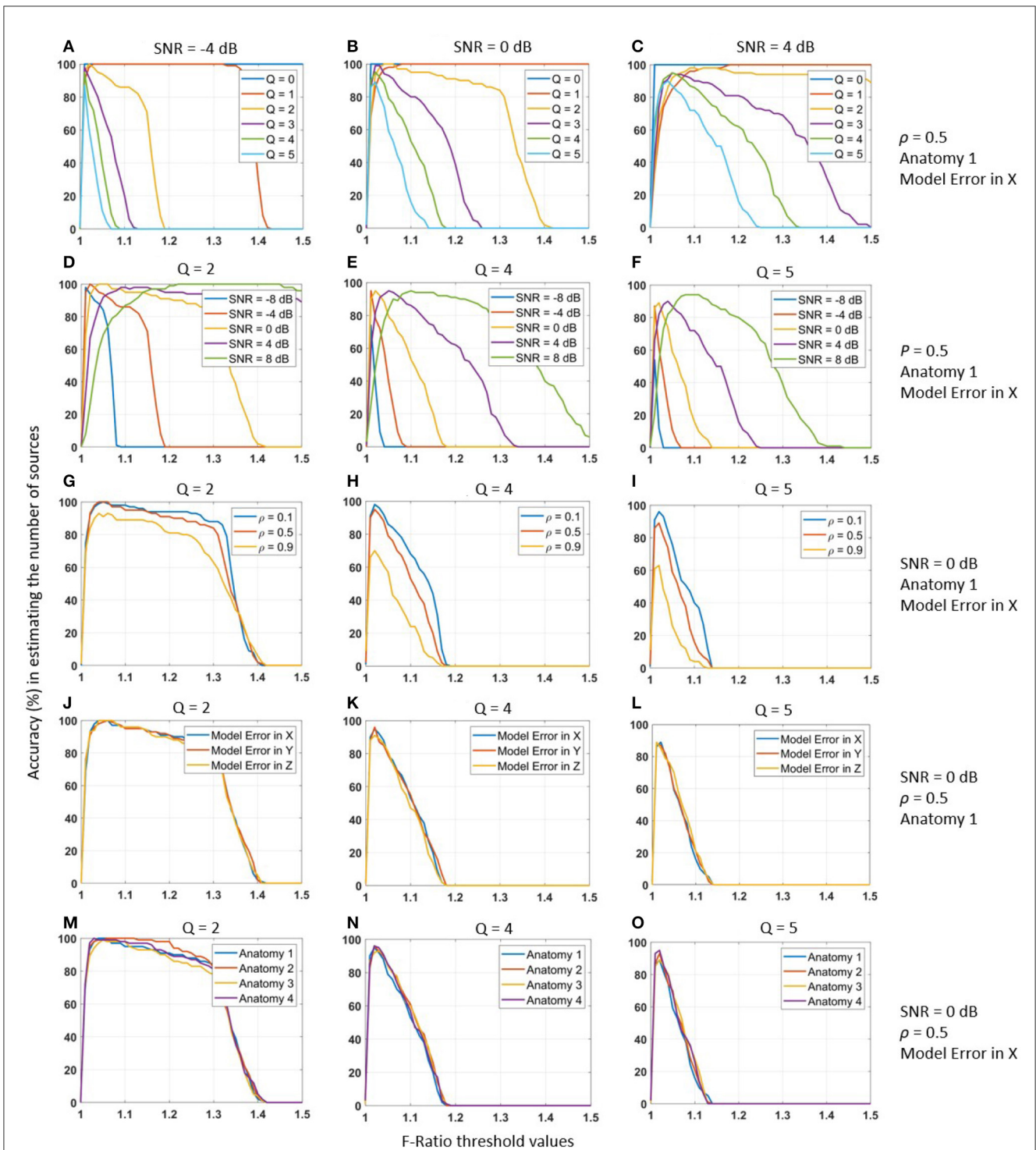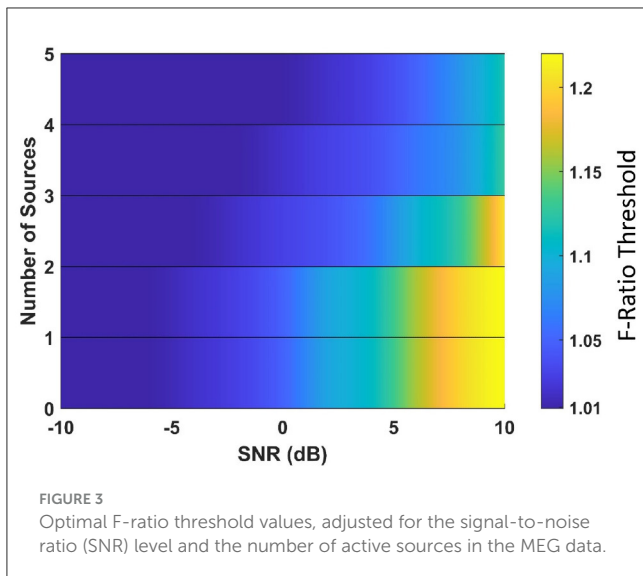
FIGURE 2

Accuracy (%) of the F-ratio method for estimating the number of active MEG sources (on vertical axis) under different threshold values (on horizontal axis). Performance evaluation was conducted across various experimental conditions: **(A–C)** varying number of true sources Q, **(D–F)** SNR levels from -8 to 8 dB, **(G–I)** inter-source correlation values ρ from 0.1 to 0.9, **(J–L)** different model errors, and **(M–O)** different cortical anatomies. Each experimental condition was tested using 100 Monte-Carlo repetitions to ensure statistical robustness. To account for head registration errors, we incorporated inaccuracies into the lead field matrix by applying a translation of 1mm posterior (X-axis), rightward (Y-axis), and upward (Z-axis).

We proceeded by conducting a comparative analysis between the proposed F-ratio method with adjusted thresholds and two commonly used methods, namely the information criterion AIC and MDL method, for estimating the number of sources. These methods rely on likelihood functions derived from information theory to assess and choose the optimal model. The comparison results of the F-ratio, AIC, and MDL methods across different SNR conditions and correlation levels are depicted in Figure 5. The

FIGURE 3
Optimal F-ratio threshold values, adjusted for the signal-to-noise ratio (SNR) level and the number of active sources in the MEG data.

results correspond to cortical anatomy 4 with 1 mm translational modeling error in X. Remarkably, the proposed F-ratio method with adjusted thresholds outperformed both the AIC and MDL methods in terms of accuracy and reliability.

3.3. Performance of the F-ratio method in estimating the number of active dipoles in phantom data

We assessed the performance of the F-ratio method in estimating the number of active dipoles in phantom data (Figure 6A). The locations of the 32 artificial dipoles of the MEGIN phantom are shown in Figure 6B. To simulate the activation of multiple MEG sources simultaneously, we combined the data obtained from individually activated dipoles. To avoid perfect coherence, a random delay ranging from 0 to 50 ms was introduced between the dipole time courses. Figure 6C illustrates an example of MEG sensor data from two active dipoles with a randomly selected temporal delay of 29 ms. The time courses are displayed following temporal and spatial prewhitening, and had an estimated SNR of 5.5 dB.

We conducted 100 Monte Carlo repetitions of phantom data simulations for each scenario involving 0 to 5 active sources. For each repetition, we applied the optimal F-ratio thresholds (presented in Figure 3) based on the estimated SNR of the phantom data and the corresponding number of tested sources. The rationale behind using the same optimal F-ratio value from simulations to both the phantom and experimental data lies in the observation that, while the performance of F-ratio thresholds varied significantly based on the number of sources and SNR levels, these thresholds remained largely invariant to variations in inter-source correlations, translational modeling inaccuracies, and different cortical anatomies.

The performance of the F-ratio method in accurately estimating the true number of active dipoles is shown in Figure 6D. The method successfully identified the correct number of sources up

to 2, surpassing both the AIC and MDL methods. In contrast, the latter methods failed entirely, with no correct estimations among the 100 simulated scenarios (results not depicted). The AIC method yielded mean estimates of 41 for no true sources and 42 for 1 to 5 true number of sources, respectively, across 100 Monte Carlo repetitions. Similarly, the MDL method estimated the number of sources as 41 in all cases of 0 to 5 true numbers of sources across the same 100 Monte Carlo repetitions. It was found that AIC and MDL consistently overestimated the number of sources.
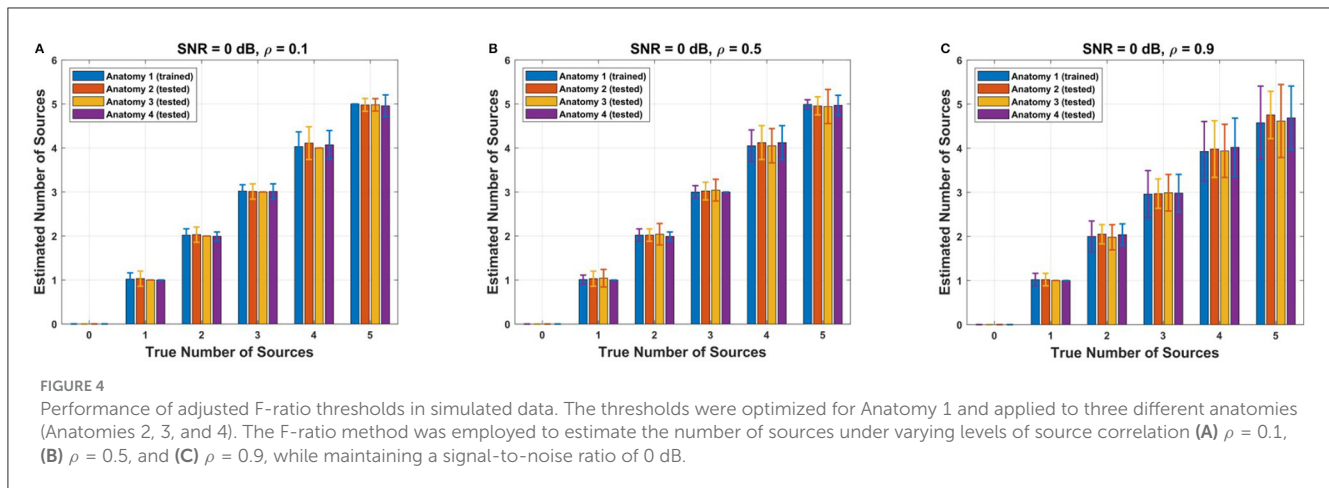
It is important to note that although the performance of the F-ratio method in experimental data was not as remarkable as in the simulated data, we attribute this to two factors. First, the specific configuration of the phantom dipoles played a critical role. The 32 phantom dipoles were closely spaced and shared similar orientations. In the 100 Monte Carlo repetitions, sources were randomly chosen with no constraints in their simultaneous activation. Consequently, there was a substantial probability of selecting adjacent sources, and this likelihood increased as the number of sources grew (3, 4, and 5). Additionally, the time courses of the phantom dipoles had random delays ranging from 0 to 50 ms, resulting in instances of minimal delay and strong correlation. The combined effect of proximate source selection and small time course delay exacerbated the challenge of accurately estimating the number of sources, particularly when dealing with a larger number of sources.

3.4. Performance of the F-ratio method in estimating the number of sources in human auditory data

To evaluate the effectiveness of the F-ratio method in human data, we utilized it to analyze brain responses captured during an auditory task. We employed the AP source localization method to fit dipoles within the time interval of 100–130 ms relative to the onset of the auditory stimuli, which approximately corresponded to the peak MEG response. The SNR within this interval was determined to be 5.78 dB.

Figure 7A illustrates the adjusted F-ratio thresholds at 5.78 dB SNR, along with the corresponding values estimated from the human data. By employing the F-ratio sequential procedure, we observed the rejection of the reduced models with zero or one source, while finding no evidence to reject (and thus accepting) a model with two distinct sources. To further validate these findings, we plotted the sum of squares of residuals for the competing models, as shown in Figure 7B. This plot reinforces our conclusion of the presence of two sources, as there was no significant reduction in the sum of squares of residuals beyond the model order of 2. Additionally, we visualized the dipoles detected using the AP localization method for the cases of 1, 2, and 3 sources (Figures 7C–E). In the case of the two-source model, the dipoles were localized bilaterally and coincided with the well-established regions in the primary auditory cortices.

It is worth mentioning that alternative methods such as the AIC and MDL yielded different estimates for the number of sources. Specifically, the AIC method suggested 33 sources, while the MDL method indicated only 31 sources. These estimates far exceed the

FIGURE 4
Performance of adjusted F-ratio thresholds in simulated data. The thresholds were optimized for Anatomy 1 and applied to three different anatomies (Anatomies 2, 3, and 4). The F-ratio method was employed to estimate the number of sources under varying levels of source correlation **(A)** ρ = 0.1, **(B)** ρ = 0.5, and **(C)** ρ = 0.9, while maintaining a signal-to-noise ratio of 0 dB.

expected number of active sources in human auditory responses and are not in line with the existing knowledge in the field.

Lastly, we conducted a further analysis on the human data with a reduced SNR of approximately 0 dB, achieved by averaging a reduced number of 44 trials (originally 166). As depicted in Supplementary Figure 1, this analysis highlights the robustness of the F-ratio test in accurately detecting 2 sources, consistent with the findings from the higher SNR case. It was found that AIC and MDL consistently overestimated the number of sources even in the presence of lower SNR. This suggests that the overestimation nature of AIC and MDL (in phantom and real data) might be related to the complex nature of human experimental data.
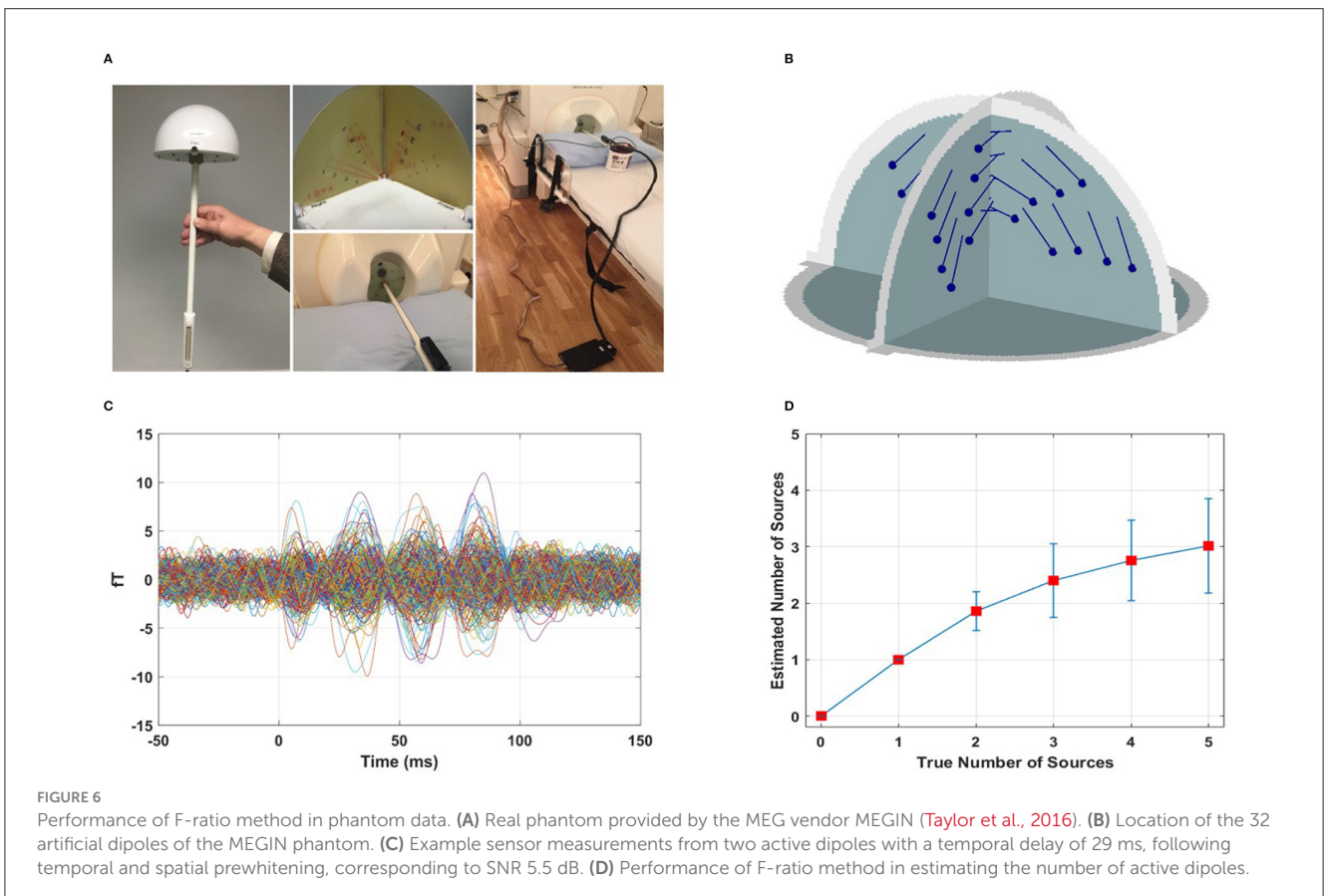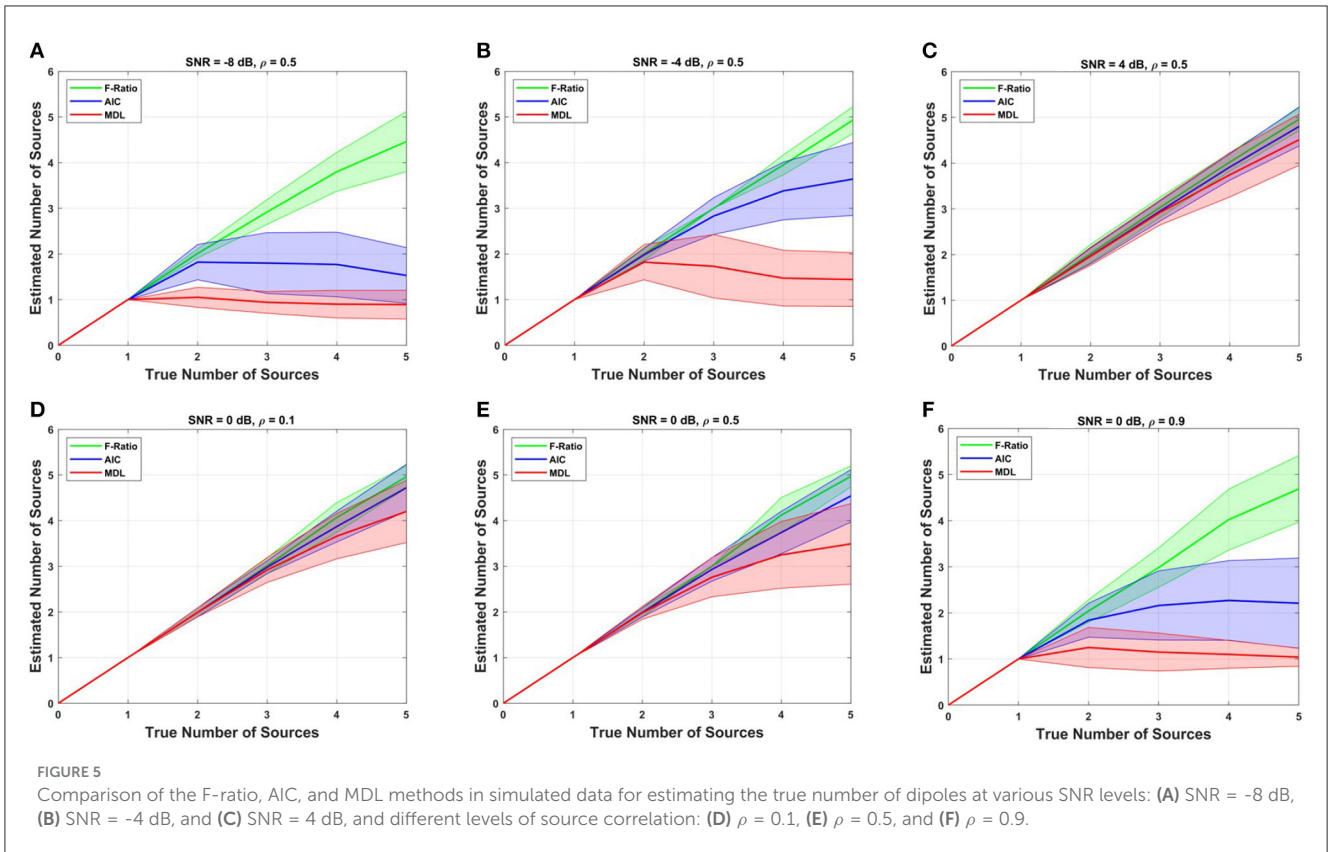
4. Discussion

We have proposed and validated an F-ratio-based method to detect the number of active sources in MEG data. We initially aimed to tune the method with a universal F-ratio threshold value that, once selected, could be applied across diverse simulation scenarios. However, we found that the performance of the F-ratio method was inherently dependent on the SNR and number of sources. Thus, we have concluded that it is not feasible to determine a single threshold value that guarantees optimal performance across all cases. Instead, we proposed a methodology that adjusts F-ratio threshold values based on the estimated SNR and the corresponding number of tested sources. Our results demonstrated the reliability and robustness of the calculated optimal F-ratio thresholds across a wide range of simulation scenarios, including variations in the number of sources, SNR levels, inter-source correlation values, modeling errors, and cortical anatomies.
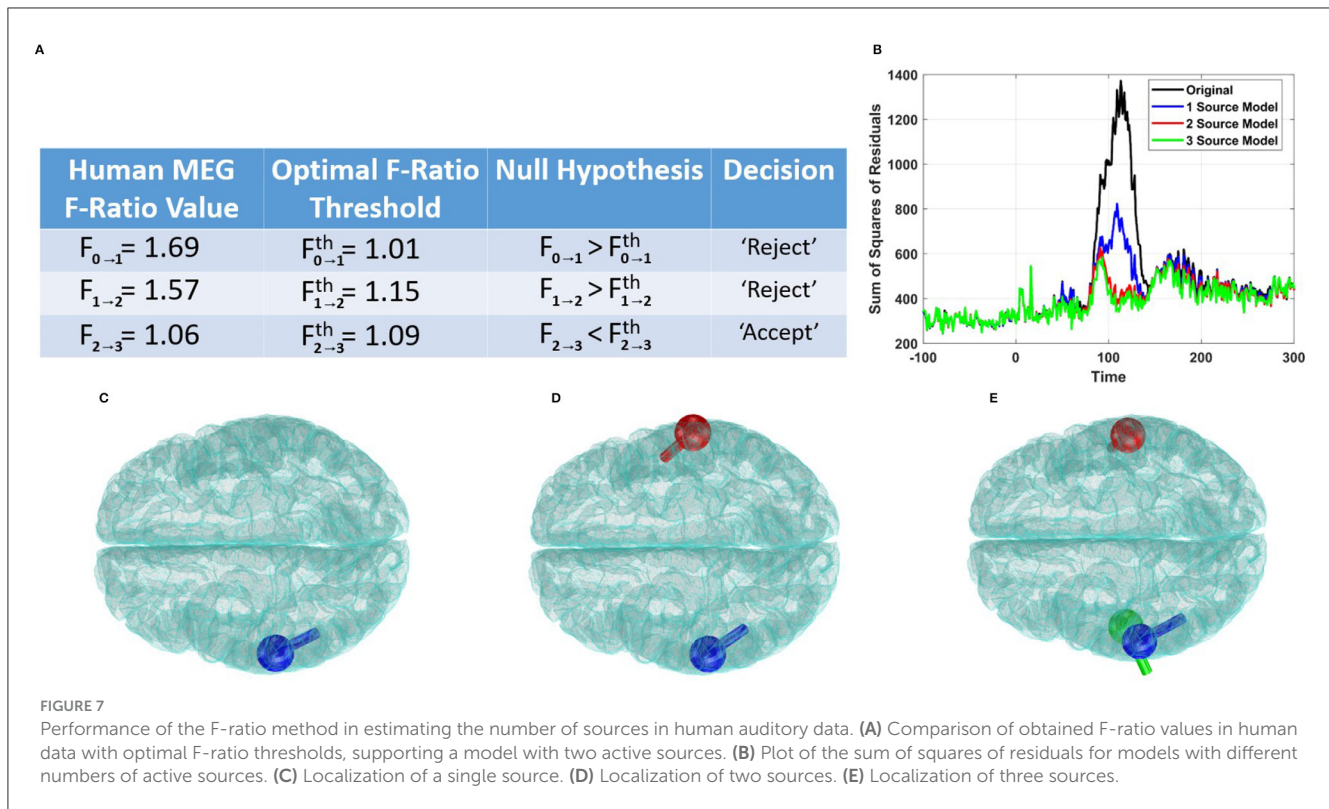
However, it is crucial to acknowledge that the adjusted threshold values obtained in this study are specific to the MEG system analyzed and may need to be adjusted for other devices or modalities, such as EEG. When applying the F-ratio method in different devices, it would be necessary to determine appropriate threshold values that are specific to each case. Similarly, the effectiveness of the proposed method is inherently linked to the choice of the source localization technique. In this study, we employed the AP method to compute and validate the F-ratio

thresholds. However, different source localization methods may yield varying results and require different threshold adjustments. Therefore, it is crucial to determine the optimal threshold values for a particular set of experimental settings and source localization method to ensure accurate estimation of the true number of active brain sources. Despite these considerations, our proposed method provides researchers with a precise tool to estimate the true number of active brain sources and effectively model brain function. By calculating threshold values that are tailored to the specific modality and source localization method, researchers can enhance the accuracy and reliability of their source estimation process.

A study by Supek and Aine (1993) aimed to evaluate the efficacy of three statistical measures, namely percent of variance, reduced chi-square, and F-ratio, in determining the correct number of sources (model order). The authors advocated for the reduced chi-square method as a reliable measure of goodness-of-fit, whereas they were less favorable to the percent of variance and F-ratio because they ignored noise contributions. Although we agree that the percent of variance has limited utility, we contend that the efficacy of the F-ratio method was underestimated. Indeed, the simulation results presented in Supek and Aine (1993) demonstrated that the F-ratio remained stable across different noise levels and successfully identified the true number of sources. However, it is important to note that the study was confined to simulations on a simple spherical head model, lacked assessments using real data, and did not provide clear threshold decision criteria for determining the correct number of sources. In this work, we proposed a methodology to estimate the number of sources by employing F-ratio threshold values based on the estimated SNR and the corresponding number of tested sources.

The Bayesian multi-dipole estimation method, Sequential Semi-Analytic Monte-Carlo Estimation (SESAME) (Sommariva and Sorrentino, 2014; Sorrentino et al., 2014; Luria et al., 2020) is an iterative Monte Carlo algorithm that approximates the posterior distribution for an a-priori unknown number of dipoles. The output of SESAME is a posterior distribution for a variable number of dipoles and their parameters. From this distribution, a cortical probability map is computed, quantifying for each voxel the posterior probability of containing a dipolar source. Additionally, the method provides a point estimate of the dipole location,

FIGURE 5
Comparison of the F-ratio, AIC, and MDL methods in simulated data for estimating the true number of dipoles at various SNR levels: **(A)** SNR = -8 dB, **(B)** SNR = -4 dB, and **(C)** SNR = 4 dB, and different levels of source correlation: **(D)** ρ = 0.1, **(E)** ρ = 0.5, and **(F)** ρ = 0.9.

FIGURE 6
Performance of F-ratio method in phantom data. **(A)** Real phantom provided by the MEG vendor MEGIN (Taylor et al., 2016). **(B)** Location of the 32 artificial dipoles of the MEGIN phantom. **(C)** Example sensor measurements from two active dipoles with a temporal delay of 29 ms, following temporal and spatial prewhitening, corresponding to SNR 5.5 dB. **(D)** Performance of F-ratio method in estimating the number of active dipoles.

FIGURE 7
Performance of the F-ratio method in estimating the number of sources in human auditory data. **(A)** Comparison of obtained F-ratio values in human data with optimal F-ratio thresholds, supporting a model with two active sources. **(B)** Plot of the sum of squares of residuals for models with different numbers of active sources. **(C)** Localization of a single source. **(D)** Localization of two sources. **(E)** Localization of three sources.

determined as the peak of the cortical probability map. In future research, it would be valuable to compare the performance of SESAME against our proposed method.

It is essential to acknowledge the fact that the performance of the F-ratio method in phantom data was not as remarkable as in the simulated data, we attribute this to two factors. First, the specific configuration of the phantom dipoles played a critical role. The 32 phantom dipoles were closely spaced and shared similar orientations. Additionally, the time courses of the phantom dipoles had random delays ranging from 0 to 50 ms, resulting in instances of minimal delay and strong correlation. The combined effect of proximate source selection and small time course delay exacerbated the challenge of accurately estimating the number of sources, particularly when dealing with a larger number of sources.

In the literature, it is well known that the localization errors tend to increase when dealing with sources that have small spatial separation. This phenomenon has also observed in the AP method, as we reported in our previous work (Adler et al., 2019). Second, estimating the number of sources beyond two is generally an exceptionally challenging problem in MEG. Consequently, a significant body of work has focused on solving the problem of localizing up to two or three sources within simulation or controlled environments (Mosher et al., 1992; Mosher and Leahy, 1999; Mäkelä et al., 2018; Adler et al., 2019; Ilmoniemi and Sarvas, 2019; Giri et al., 2020).

Finally, it is important to mention that the MEG human experiment revealed the presence of two active sources. Notably, the F-ratio method consistently exhibited excellent performance in simulations and phantom experiments with this specific number of sources. However, it is crucial to acknowledge that as the

number of sources increases, the suitability of the F-Ratio method might diminish, necessitating further investigation in future studies.

5. Conclusion

We have validated our F-ratio-based method on simulated, real phantom, and human MEG data. In comparison to other state-of-the-art statistical approaches like AIC and MDL, which rely on certain assumptions that often do not hold in real-world situations, our method demonstrated superior performance in terms of accuracy and reliability. One crucial aspect we emphasized is the selection of appropriate thresholds for the F-ratio values, which significantly impacts the overall performance of the method. We identified optimal thresholds and showed that these thresholds needed to be adjusted for the number of sources and SNR levels. Notably, these thresholds exhibited remarkable consistency across different inter-source correlations, head translation modeling errors, and cortical anatomies. Overall, by fine-tuning the selection of thresholds, our F-ratio-based method provides researchers with a precise and robust tool for accurately estimating the true number of active sources in MEG data. Further research is needed to explore and validate the proposed method in different modalities and with various source localization techniques. By refining the threshold determination process and investigating its applicability across different experimental conditions, we can extend the utility of this method to a wider range of neuroimaging studies and enhance our understanding of the underlying mechanisms of brain function.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://github.com/Amita-Giri/FRatio-based-Method-for-Source-Enumeration.

Ethics statement

The studies involving humans were approved by Institutional Review Board of the Massachusetts Institute of Technology. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

AG and DP developed the methodology. AG performed the coding, data analysis, and wrote the first manuscript draft. DP made major contributions to the writing of the final manuscript. AA and DP provided funding, resources, and supervision. All authors contributed to conception of the method, interpreted and discussed the results, critically reviewed the manuscript, and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnhum.2023.1235192/full#supplementary-material

References

Adler, A., Wax, M., and Pantazis, D. (2019). Localization of MEG and EEG Brain Signals by Alternating Projection. *arXiv. [Preprint]*. arXiv:1908.11416.

Adler, A., Wax, M., and Pantazis, D. (2022). "Brain source localization by alternating projection," in *2022 IEEE 19th International Symposium on Biomedical Imaging* (ISBI), (IEEE), p. 1–5.

Ahveninen, J., Jääskeläinen, I. P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., et al. (2006). Task-modulated "what" and "where" pathways in human auditory cortex. *Proc. Nat. Acad. Sci.* 103, 14608–14613. doi: 10.1073/pnas.0510480103

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* 19, 716–723. doi: 10.1109/TAC.1974.1100705

Baillet, S. (2017). Magnetoencephalography for brain electrophysiology and imaging. *Nat. Neurosci.* 20, 327–339. doi: 10.1038/nn.4504

Bartlett, M. S. (1954). A note on the multiplying factors for various χ 2 approximations. *J. Royal Stat. Soc. Series B (Methodological)*. 1954, 296–298. doi: 10.1111/j.2517-6161.1954.tb00174.x

Chen, W., Wong, K. M., and Reilly, J. P. (1991). Detection of the number of signals: a predicted eigen-threshold approach. *IEEE Trans. Signal Proc.* 39, 1088–1098. doi: 10.1109/78.80959

Colton, D. L., and Kress, R. (1998). *Inverse Acoustic and Electromagnetic Scattering Theory*. Cham: Springer.

Dale, A.M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., et al. (2020). Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron* 26, 55–67. doi: 10.1016/S0896-6273(00)81138-1

Darvas, F., Pantazis, D., Kucukaltun-Yildirim, E., and Leahy, R. (2004). Mapping human brain function with meg and eeg: methods and validation. *Neuroimage* 23, S289–S299. doi: 10.1016/j.neuroimage.2004.07.014

Di, A., and Tian, L. (1984). "Matrix decomposition and multiple source location," in ICASSP'84. IEEE International Conference on Acoustics, Speech, and Signal Processing. San Diego, CA: IEEE, 722–725.

Fischl, B., Salat, D. H., Van Der Kouwe, A. J., Makris, N., Ségonne, F., Quinn, B. T., et al. (2004). Sequence-independent segmentation of magnetic resonance images. *Neuroimage* 23, S69–S84. doi: 10.1016/j.neuroimage.2004.07.016

Giorgetta, C., Grecucci, A., Bonini, N., Coricelli, G., Demarchi, G., Braun, C., et al. (2013). Waves of regret: A meg study of emotion and decision-making. *Neuropsychologia* 51, 38–51. doi: 10.1016/j.neuropsychologia.2012.10.015

Giri, A. (2022). *Spatial and anatomical harmonics domain based brain source localization*. PhD thesis, IIT Delhi. doi: 10.21203/rs.3.rs-953226/v1

Giri, A., Kumar, L., and Gandhi, T. (2018). "EEG dipole source localization in hemispherical harmonics domain," in *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. Honolulu, HI: IEEE, 679–684.

Giri, A., Kumar, L., and Gandhi, T. (2019). "Head harmonics based EEG dipole source localization," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*. Pacific Grove, CA: IEEE, 2149–2153.

Giri, A., Kumar, L., and Gandhi, T. K. (2020). Brain source localization in head harmonics domain. *IEEE Trans. Instrum. Meas.* 70, 1–10. doi: 10.1109/TIM.2020.3026511

Giri, A., Kumar, L., Kurwale, N., and Gandhi, T. K. (2022). Anatomical harmonics basis based brain source localization with application to epilepsy. *Sci. Rep.* 12, 11240. doi: 10.1038/s41598-022-14500-7

Gramfort, A., Kowalski, M., and Hämäläinen, M. (2012). Mixed-norm estimates for the m/eeg inverse problem using accelerated gradient methods. *Phys. Med. Biol.* 57, 1937. doi: 10.1088/0031-9155/57/7/1937

Gramfort, A., Papadopoulo, T., Olivi, E., and Clerc, M. (2010). OpenMEEG: opensource software for quasistatic bioelectromagnetics. *Biomed. Eng. Online* 9, 1–20. doi: 10.1186/1475-925X-9-45

Gramfort, A., Strohmeier, D., Haueisen, J., Hämäläinen, M. S., and Kowalski, M. (2013). Time-frequency mixed-norm estimates: Sparse m/eeg imaging with non-stationary source activations. *Neuroimage* 70, 410–422. doi: 10.1016/j.neuroimage.2012.12.051

Green, A. A., Berman, M., Switzer, P., and Craig, M. D. (1988). A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Trans. Geosci. Remote Sens.* 26:65–74. doi: 10.1109/36.3001

Hämäläinen and Ilmoniemi,Hämäläinen, M. S., and Ilmoniemi, R. J. (1994). Interpreting magnetic fields of the brain: minimum norm estimates. *Med. Biol. Eng. Comput.*32, 35–42. doi: 10.1007/BF02512476

Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., and Lounasmaa, O. (1993). Magnetoencephalography theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev. Mod. Phys.* 65, 413–497. doi: 10.1103/RevModPhys.65.413

Hecker, L., Tebartz van Elst, L., and Kornmeier, J. (2023). Source localization using recursively applied and projected MUSIC with flexible extent estimation. *Front. Neurosci.* 17:1170862. doi: 10.3389/fnins.2023.1170862

Ikeda, S., and Toyama, K. (2000). Independent component analysis for noisy data—MEG data analysis. *Neural Networks* 13, 1063–1074. doi: 10.1016/S0893-6080(00)00071-X

Ilmoniemi, R., and Sarvas, J. (2019). *Brain Signals: Physics and Mathematics of MEG and EEG*. Cambridge, MA: MIT Press.

Jiang, J.-S., and Ingram, M. A. (2004). "Robust detection of number of sources using the transformed rotational matrix," in *2004 IEEE Wireless Communications and Networking Conference (IEEE Cat. No. 04TH8733)*. Atlanta, GA: IEEE, 501–506.

Klepp, A., Niccolai, V., Buccino, G., Schnitzler, A., and Biermann-Ruben, K. (2015). Language-motor interference reflected in MEG beta oscillations. *Neuroimage* 109, 438–448. doi: 10.1016/j.neuroimage.2014.12.077

Knösche, T. R., Berends, E., Jagers, H., and Peters, M. J. (1998). Determining the number of independent sources of the EEG: a simulation study on information criteria. *Brain Topogr.* 11, 111–124. doi: 10.1023/A:1022202521439

Krishnaswamy, P., Obregon-Henao, G., Ahveninen, J., Khan, S., Babadi, B., Iglesias, J. E., et al. (2017). Sparsity enables estimation of both subcortical and cortical activity from meg and eeg. *Proc. Nat. Acad. Sci.* 114, E10465–E10474. doi: 10.1073/pnas.1705414114

Lawley, D. (1956). Tests of significance for the latent roots of covariance and correlation matrices. *Biometrika*. 43, 128–136. doi: 10.1093/biomet/43.1-2.128

Luria, G., Duran, D., Visani, E., Rossi Sebastiano, D., Sorrentino, A., Tassi, L., et al. (2020). Towards the automatic localization of the irritative zone through magnetic source imaging. *Brain Topogr.* 33, 651–663. doi: 10.1007/s10548-020-00789-y

Mäkelä, N., Stenroos, M., Sarvas, J., and Ilmoniemi, R. J. (2017). Locating highly correlated sources from MEG with (recursive)(R) DS-MUSIC. *bioRxiv [Preprint]*. doi: 10.1101/230672

Mäkelä, N., Stenroos, M., Sarvas, J., and Ilmoniemi, R. J. (2018). Truncated RAP-MUSIC (TRAP-MUSIC) for MEG and EEG source localization. *Neuroimage* 167, 73–83. doi: 10.1016/j.neuroimage.2017.11.013

Malinowski, E. R. (1977a). Determination of the number of factors and the experimental error in a data matrix. *Anal. Chem.* 49, 612–617. doi: 10.1021/ac50012a027

Malinowski, E. R. (1977b). Theory of error in factor analysis. *Anal. Chem.* 49, 606–612. doi: 10.1021/ac50012a026

Mosher, J., Lewis, P., and Leahy, R. (1992). Multiple dipole modeling and localization from spatio-temporal MEG data. *IEEE Trans. Biomed. Eng.* 39, 541–557. doi: 10.1109/10.141192

Mosher, J. C., and Leahy, R. M. (1999). Source localization using recursively applied and projected (RAP) MUSIC. *IEEE Trans. Signal Proc.* 47, 332–340. doi: 10.1109/78.740118

Oishi, M., Otsubo, H., Kameyama, S., Morota, N., Masuda, H., Kitayama, M., et al. (2002). Epileptic spikes: magnetoencephalography versus simultaneous electrocorticography. *Epilepsia* 43, 1390–1395. doi: 10.1046/j.1528-1157.2002.10702.x

Pancholi, S., Giri, A., Jain, A., Kumar, L., and Roy, S. (2023). "Source aware deep learning framework for hand kinematic reconstruction using

EEG signal," in *IEEE Transactions on Cybernetics, Vol. 53* (IEEE), 4094–4106. doi: 10.1109/TCYB.2022.3166604

Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sloreta): technical details. *Methods Find. Exp. Clin. Pharmacol.* 24, 5–12.

Rezaei, A., Lahtinen, J., Neugebauer, F., Antonakakis, M., Piastra, M. C., Koulouri, A., et al. (2021). Reconstructing subcortical and cortical somatosensory activity via the RAMUS inverse source analysis technique using median nerve SEP data. *Neuroimage* 245, 118726. doi: 10.1016/j.neuroimage.2021.118726

Rissanen, J. (1978). Modeling by shortest data description. *Automatica*. 14, 465–471. doi: 10.1016/0005-1098(78)90005-5

Salman, T., Badawy, A., Elfouly, T. M., Mohamed, A., and Khattab, T. (2015). "Estimating the number of sources: An efficient maximization approach," in *2015 International Wireless Communications and Mobile Computing Conference (IWCMC)*. Dubrovnik: IEEE, 199–204.

Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Statist.* 461–464. doi: 10.1214/aos/1176344136

Sekihara, K., Nagarajan, S. S., Poeppel, D., Marantz, A., and Miyashita, Y. (2001). Reconstructing spatio-temporal activities of neural sources using an MEG vector beamformer technique. *IEEE Trans. Biomed. Eng.* 48, 760–771. doi: 10.1109/10.930901

Sommariva, S., and Sorrentino, A. (2014). Sequential Monte Carlo samplers for semi-linear inverse problems and application to magnetoencephalography. *Inverse Probl.* 30, 114020. doi: 10.1088/0266-5611/30/11/114020

Sorrentino, A., Luria, G., and Aramini, R. (2014). Bayesian multi-dipole modelling of a single topography in MEG by adaptive sequential Monte Carlo samplers. *Inverse Probl.* 30, 045010. doi: 10.1088/0266-5611/30/4/045010

Strohmeier, D., Bekhti, Y., Haueisen, J., and Gramfort, A. (2016). The iterative reweighted mixed-norm estimate for spatio-temporal MEG/EEG source reconstruction. *IEEE Trans. Med. Imaging* 35, 2218–2228. doi: 10.1109/TMI.2016.2553445

Supek, S., and Aine, C. J. (1993). Simulation studies of multiple dipole neuromagnetic source localization: model order and limits of source resolution. *IEEE Trans. Biomed.l Eng.* 40, 529–540. doi: 10.1109/10.237672

Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., and Leahy, R. M. (2011). Brainstorm: a user-friendly application for MEG/EEG analysis. *Comput. Intell. Neurosci.* 2011, 1–13. doi: 10.1155/2011/879716

Taylor, K., Tadel, F., and Mosher, J. C. (2016). *MEG Current Phantom (Elekta-Neuromag)*. Available online at: https://neuroimage.usc.edu/brainstorm/Tutorials/PhantomElekta

Van Veen, B. D., Van Drongelen, W., Yuchtman, M., and Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans. Biomed. Eng.* 44, 867–880. doi: 10.1109/10.623056

Vrba, J., and Robinson, S. E. (2001). Signal processing in magnetoencephalography. *Methods* 25, 249–271. doi: 10.1006/meth.2001.1238

Wax, M., and Kailath, T. (1985). Detection of signals by information theoretic criteria. *IEEE Trans. Acoust.* 33, 387–392. doi: 10.1109/TASSP.1985.1164557

Wendel, K., Väisänen, O., Malmivuo, J., Gencer, N. G., Vanrumste, B., Durka, P., et al. (2009). EEG/MEG source imaging: methods, challenges, and open issues. *Comput. Intell. Neurosci.* 2009, 656092. doi: 10.1155/2009/656092

Westlake, K. P., Hinkley, L. B., Bucci, M., Guggisberg, A. G., Findlay, A. M., Henry, R. G., et al. (2012). Resting state alpha-band functional connectivity and recovery after stroke. *Exp. Neurol.* 237, 160–169. doi: 10.1016/j.expneurol.2012.06.020

Wilkinson, C. M., Burrell, J. I., Kuziek, J. W., Thirunavukkarasu, S., Buck, B. H., and Mathewson, K. E. (2020). Predicting stroke severity with a 3-min recording from the muse portable EEG system for rapid diagnosis of stroke. *Sci. Rep.* 10, 1–11. doi: 10.1038/s41598-020-75379-w

Xu, M., Sanz, D. L., Garces, P., Maestu, F., Li, Q., and Pantazis, D. (2021). A graph gaussian embedding method for predicting alzheimer's disease progression with MEG brain networks. *IEEE Trans. Biomed. Eng.* 68, 1579–1588. doi: 10.1109/TBME.2021.3049199

Yao, Z., Zhang, Y., Bai, Z., and Eddy, W. F. (2018). Estimating the number of sources in magnetoencephalography using spiked population eigenvalues. *J. Am. Stat. Assoc.* 113, 505–518. doi: 10.1080/01621459.2017.1341411

Zhang, Q.-T., Wong, K. M., Yip, P. C., and Reilly, J. P. (1989). Statistical analysis of the performance of information theoretic criteria in the detection of the number of signals in array processing. *IEEE Trans. Acoust.* 37, 1557–1567. doi: 10.1109/29.35394