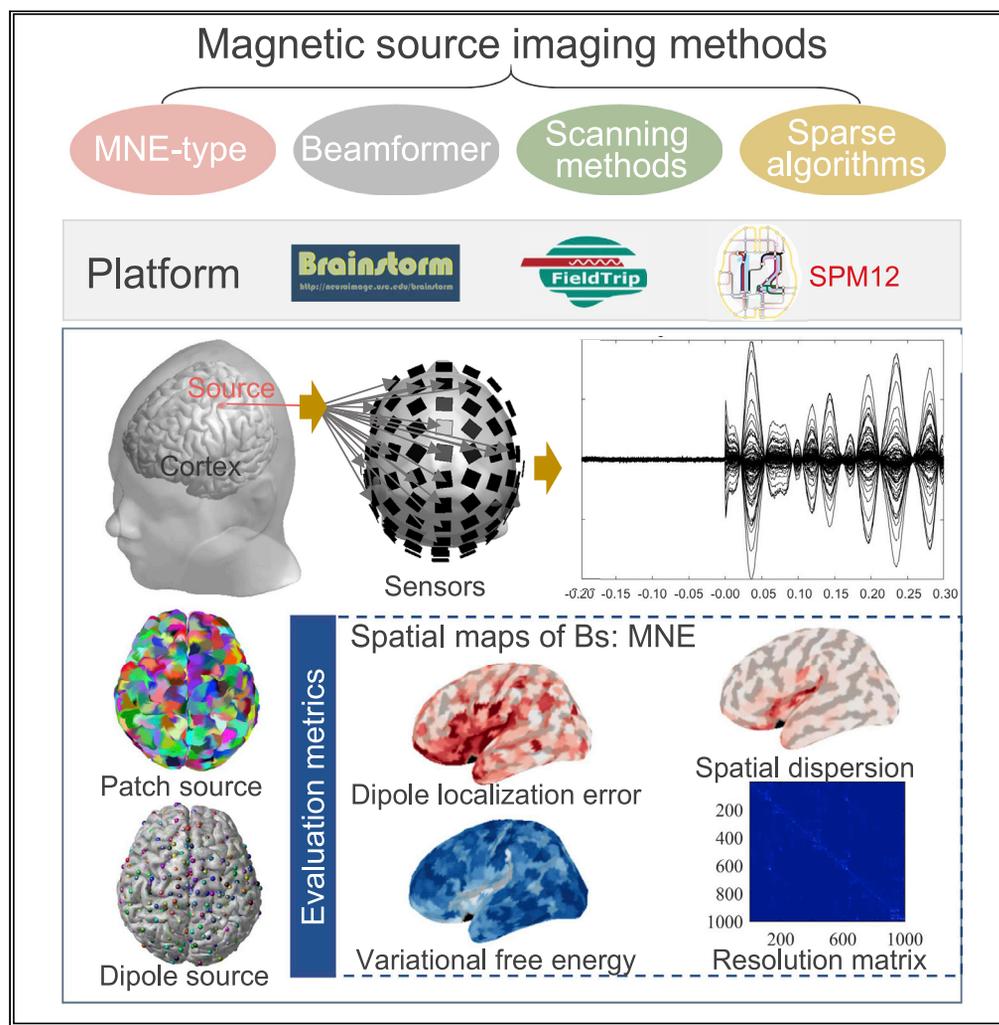


Article

# Spatial accuracy evaluation of magnetic source imaging methods on OPM-based MEG



Nan An, Fuzhi Cao, Wen Li, ..., Dawei Wang, Dexin Yu, Xiaolin Ning

ningxiaolin@buaa.edu.cn

**Highlights**

MSI methods implemented in Brainstorm, FieldTrip, and SPM12 are evaluated

Variational free energy is proposed to compare different source imaging results

Metrics, DLE, SD, and resolution matrix, are included to assess different methods

MSI methods were applied to imaging sources in a median nerve OPM-MEG experiment



## Article

## Spatial accuracy evaluation of magnetic source imaging methods on OPM-based MEG

Nan An,<sup>1,2</sup> Fuzhi Cao,<sup>1,2</sup> Wen Li,<sup>1,2</sup> Wenli Wang,<sup>1,2</sup> Weinan Xu,<sup>1,2</sup> Chunhui Wang,<sup>1,2</sup> Min Xiang,<sup>1,3</sup> Yang Gao,<sup>1,4</sup> Binbin Sui,<sup>5</sup> Dawei Wang,<sup>6,7,8</sup> Dexin Yu,<sup>6,7,8</sup> and Xiaolin Ning<sup>1,3,9,\*</sup>

## SUMMARY

The emergence of the optically pumped magnetometer (OPM)-based magnetoencephalography (MEG) has led to new developments in MEG technology. The source imaging results of different magnetic source imaging (MSI) methods show considerable differences, which makes it difficult for researchers to choose an appropriate method. This study assessed time-domain MSI methods implemented in the Brainstorm, FieldTrip, and SPM12 toolboxes using simulations. We proposed using a metric, variational free energy under the Bayesian framework, as an indicator to evaluate source imaging results because it does not require the ground truth of sources but uses the fitness of the measurement data. Our simulations demonstrated the effectiveness of the variational free energy in indicating the quality of the source reconstruction results. We then applied each MSI method to the real OPM-MEG experimental data. We aimed to highlight the characteristics of each method and provide references for researchers choosing an appropriate MSI method.

## INTRODUCTION

Magnetoencephalography (MEG) is a noninvasive technology used in basic and clinical neuroscience research (Iwasaki and Nakasato, 2019). In recent years, newer emerging sensors, such as high-temperature SQUIDs (Faley et al., 2012) and optically pumped magnetometers (OPMs) (Knappe et al., 2014), have brought vitality to the development of MEG. The new sensors do not require an additional cryogenic cooling system and can be as close to the scalp as possible to obtain higher signals (Chesca et al., 2015; Boto et al., 2016). In particular, OPM-MEG makes new measurement scenarios possible, such as wearable and movable measurements (Iivanainen et al., 2020; Boto et al., 2018), and even in a virtual reality environment (Roberts et al., 2019), which has attracted tremendous enthusiasm among researchers.

MEG measures the magnetic fields directly related to the underlying brain electrical activity with excellent temporal resolution. Unfortunately, its spatial resolution relies heavily on solving an ill-posed problem: inferring the source signals of tens of thousands of sources from a limited number of sensors (hundreds) of measured data. To solve this highly underdetermined problem, multiple magnetic source imaging (MSI) approaches based on different constraints of the solution space have been developed to provide a better estimation of source activities, such as the minimum norm estimate (MNE) (Hämäläinen and Ilmoniemi, 1994) and its variants (Dale et al., 2000; Pascual-Marqui et al., 2002; Pascual-Marqui et al., 2011), beam formers (Van Veen et al., 1997), dipole scanning methods (Mosher et al., 1992), and sparse algorithms (Friston et al., 2008a; Zerouali et al., 2011). These inverse approaches have been implemented and integrated into several open-source toolboxes, such as Brainstorm (Tadel et al., 2011), FieldTrip (Oostenveld et al., 2011), and SPM12 (Henson et al., 2019). These toolboxes offer great convenience to neuroscience investigators for source analysis. However, each approach yields different imaging results (Hincapié et al., 2017; Hedrich et al., 2017; Tait et al., 2021). Even for the same approach, differences in toolbox implementation provide different results (Jaiswal et al., 2020; Westner et al., 2022). A challenge for researchers is to choose an appropriate imaging approach from several available source reconstruction strategies. Therefore, it is necessary to evaluate different source imaging methods, which was the focus of our study. The MSI methods implemented in the three toolboxes were compared, including the MNE, dSPM, sLORETA, LCMV, DipModel, and cMEM methods in Brainstorm, mne, sloreta, eloreta, lcmv, and rv (dipole scanning) methods in Fieldtrip, and IID, COH, EBB, and MSP methods in SPM12.

<sup>1</sup>Zhejiang Provincial Key Laboratory of Ultra-Weak Magnetic-Field Space and Applied Technology, Hangzhou Innovation Institute, Beihang University, Hangzhou 310051, China

<sup>2</sup>Key Laboratory of Ultra-Weak Magnetic Field Measurement Technology, Ministry of Education, School of Instrumentation and Optoelectronic Engineering, Beihang University, Beijing 100191, China

<sup>3</sup>Research Institute of Frontier Science, Beihang University, Beijing 100191, China

<sup>4</sup>School of Physics, Beihang University, Beijing 100191, China

<sup>5</sup>Beijing Tiantan Hospital, Capital Medical University, Beijing 100050, China

<sup>6</sup>Department of Radiology, Qilu Hospital of Shandong University, Jinan 250012, China

<sup>7</sup>Shandong Key Laboratory: Magnetic Field-free Medicine and Functional Imaging, Jinan 250012, China

<sup>8</sup>Research Institute of Shandong University: Magnetic Field-free Medicine and Functional Imaging, Jinan 250012, China

<sup>9</sup>Lead contact

\*Correspondence:

ningxiaolin@buaa.edu.cn

<https://doi.org/10.1016/j.isci.2022.105177>



**Table 1. A brief statistics of the source imaging methods for analyzing the time-domain data provided in three toolboxes: Brainstorm, FieldTrip, and SPM12**

Software	Source imaging methods for analyzing the time-domain data			
	MNE-like solutions	Beamformer	Scanning methods	Sparse algorithms
Brainstorm	MNE/dSPM/sLORETA	LCMV	DipModel	cMEM
FieldTrip	mne/slureta/elureta	lcmv	rv	–
SPM12	IID/COH/EBB	–	–	MSP

See also Table S1.

\*Methods under a similar theoretical framework are listed in the same column and denoted using different names or case sensitivities.

Researchers have evaluated imaging methods by using various metrics. The commonly used metrics are the dipole localization error (DLE) and spatial dispersion (SD) (Hauk et al., 2011), which are used to describe the distance between the reconstructed source peak and the simulated one and how the reconstructed source area is extended, respectively. Grova et al. (2006) proposed using the area under the receiver operating characteristic curve, which considers the imaging threshold, to assess source detection accuracy. Liu et al. (2002) defined crosstalk and point spread metrics to quantify the spatial resolution of the inverse operator, which was also summarized as an analytical resolution matrix (Hedrich et al., 2017). Samuelsson et al. (2021) developed a resolution matrix to make it possible to evaluate linear and nonlinear estimates. The aforementioned metrics can be used to compare different source estimates based on simulations in which the simulated sources are known. However, the “ground truth” is unknown in practical applications, meaning that the metrics mentioned earlier cannot be used to evaluate the imaging results of different methods in this situation. One possible strategy is to perform Monte Carlo simulations using real experimental conditions to compare the imaging results and choose a suitable one. However, this is complicated and inconvenient. Fortunately, under the Bayesian framework, model evidence allows for comparing and evaluating different imaging methods (Mattout et al., 2006; Wipf and Nagarajan, 2009). It does not need to know the “ground truth” of the sources but instead uses an indicator of how well the measured data fit. We proposed using variational free energy (Friston et al., 2007), an approximation of the model evidence, as a metric for evaluating source imaging methods. Our comprehensive simulation experiments demonstrated its validity in indicating the quality of the source reconstruction results.

Our study compared the performances of all time-domain source estimation methods available in the three toolboxes, Brainstorm, FieldTrip, and SPM12. The spatial fidelities of all the methods were evaluated using the proposed metric and general metrics, including the DLE, SD, and resolution matrix. Finally, we applied all the source imaging methods to real median nerve stimulation data measured by OPM-MEG to demonstrate the functional performance of each imaging method.

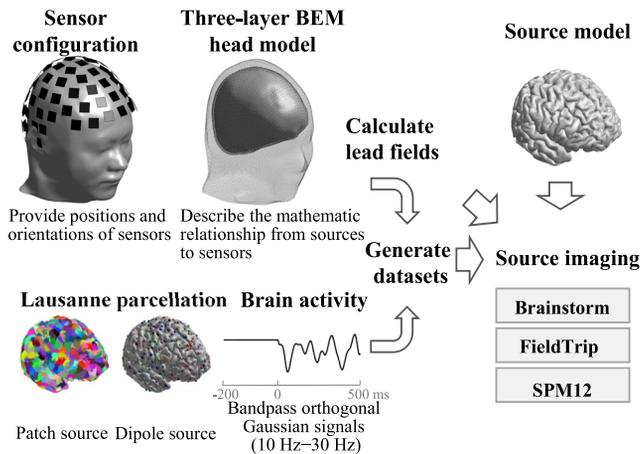
## RESULTS

### Simulations

In this study, we attempted to evaluate all source imaging methods provided in the Brainstorm, FieldTrip, and SPM12 toolboxes to analyze the time-domain data. Table 1 summarizes the MSI methods being compared. A series of simulations were performed to assess different methods. The main parameter settings for each MSI method are summarized in Table S1. The simulation protocol is shown in Figure 1.

#### Effectiveness of variational free energy

Figure 2 illustrates an example of the free energy trends under the estimated source patch with changing spatial sizes and locations relative to the real simulated source. A source patch with an initial extent at the initial position was used to generate datasets. Source patches of different sizes and positions were assumed to simulate the source power distributions estimated using different MSI methods. The presumed source patch was encoded in source covariance  $\hat{Q}_J$  to calculate the corresponding free energy. The calculated free energy was normalized by its maximum value in each group (groups of extent-change and position-change). Figure 2 shows that free energy is at its maximum when the estimated source is at its exact location and has an accurate spatial size.



**Figure 1. Simulation protocol**

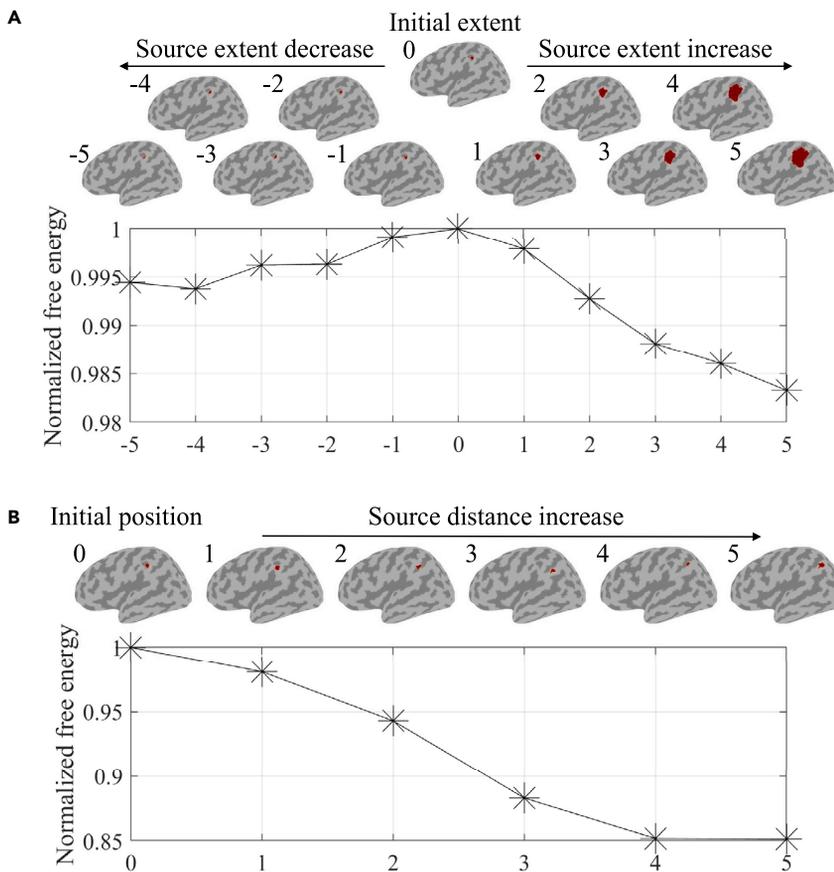
The 85-channel optically pumped magnetometer-based MEG sensor configuration and a three-layer boundary element method (BEM) head model are used to calculate lead fields. Each patch/dipole source is activated to generate datasets. The cortex mesh with 15,002 vertices is used as the distributed source model. The source imaging methods provided in Brainstorm, FieldTrip, and SPM12 are evaluated.

To illustrate the relationship between the source imaging results of each method and the calculated metrics, including the DLE, SD, and variational free energy  $F$ , a typical patch source located in the post central cortex was activated. Its source imaging results are shown in Figure 3. The corresponding DLE, SD, and  $F$  metric values are also displayed. The variational free energy of each method was normalized using the minimum and maximum  $F$  values of all 15 inversion methods on the same dataset. For convenience, the toolboxes referred to are abbreviated as Bs, Ft, and SPM. It can be observed that Bs: cMEM shows the best spatial fidelity and has the highest  $F$  value in this simulation. Overall, when both DLE and SD were relatively small, the  $F$  value was relatively high (the results of the Bs: cMEM, Ft: lcmv, and SPM methods are shown in Figure 3). These results illustrate the effectiveness of the variational free energy in indicating the quality of the source imaging results. Figure 3 provides an intuitive understanding of each imaging method. For example, sLORETA, LCMV, and DipModel in Brainstorm, mne, sloreta, eloreta, and rv in FieldTrip overestimate the source extent.

To further evaluate each method, we activated each patch source provided by the Lausanne parcellation and calculated the metrics for each method. The spatial distributions of DLE, SD, and normalized  $F$  metrics are shown in Figure 4. The spatial maps display the metrics as a function of the source locations. The lower the DLE and SD, and the higher the  $F$  values, the better the performance of each method. It can be seen that the insula and cingulate gyrus showed higher DLE for Bs: MNE, and cMEM, SPM: IID, COH, and EBB methods. Correspondingly, the  $F$  values of the relevant locations were small. Among the methods with similar DLE distributions, such as Bs: LCMV, DipModel, and sLORETA, the method with a higher SD showed smaller  $F$  values. To this extent, the  $F$  values integrate the characteristics of DLE and SD. In addition, we calculated a metric called the Wasserstein distance (Janati et al., 2019; Niklas, 2022), which also shows the ability to incorporate the characteristics of DLE and SD, as shown in Figure S1.

### Evaluation of different methods

**Comparison.** As shown in Figure 4, among all the comparison methods, Bs: sLORETA and Ft: eloreta had the best performance for imaging patch sources over the whole cortex, as indicated by the combination of low DLE and SD and high  $F$  values relative to others. The LCMV method implementations in different toolboxes showed significant discrepancies, which were indicated by the large differences in their  $F$  distributions. This discrepancy was also validated by the DLE and SD distributions. Although Bs: LCMV and Ft: lcmv had low source localization accuracies, Bs: LCMV significantly overestimated the source extent, whereas Ft: lcmv showed smaller SD values, except for some outliers. For the dipole scanning methods (Bs: DipModel and Ft: rv), the low  $F$  values were due to their high SD values over the entire cortex.  $F$  distributions were uneven for the two types of sparse algorithms Bs: cMEM and Ft: MSP. Although the SD values for the two sparse algorithms were the smallest, they had significant source localization errors for some patch sources.

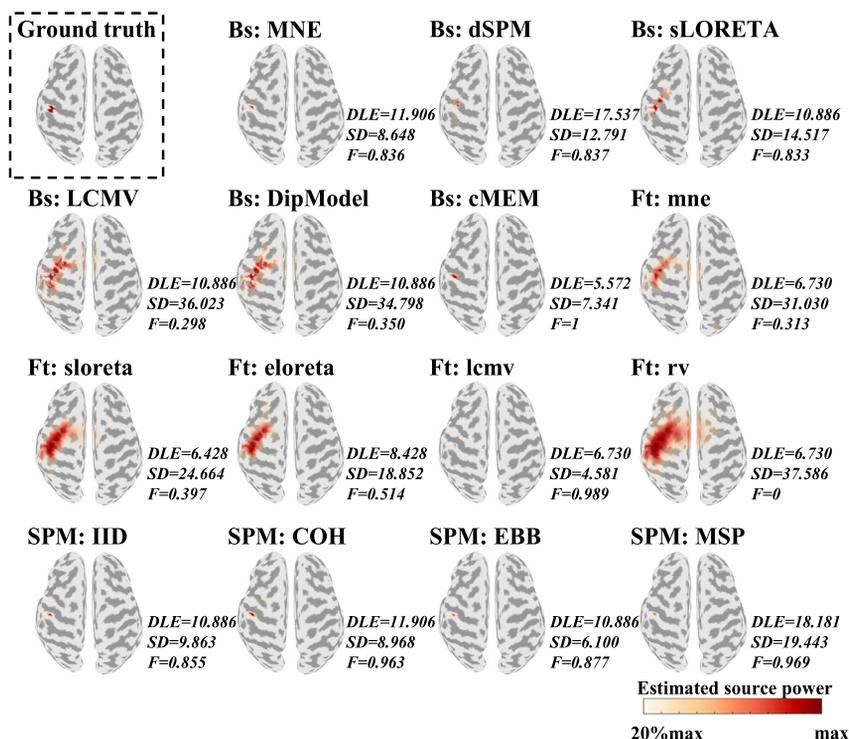


**Figure 2. Example of the normalized free energy trajectories under the change of the source spatial size and location**

(A) The spatial size varies relative to the real simulated source. The source patch (labeled 0) with the initial extent at the initial position was used to generate datasets. The source patches (labeled 0 and  $\sim 0$ ) simulated the source power distribution estimated by MSI methods. The larger the label number, the larger size of the “estimated” source patch. (B) The source location varies relative to the real simulated source. The larger the label number, the far the “estimated” source patch locations are from the simulated source patch location.

There were similarities between the different methods. To further evaluate the similarities, we calculated the average correlation between the spatial distribution of the source power estimated by pairs of algorithms for all 1,000 source patches and plotted a hierarchical clustering tree to show the similarity, as shown in Figure 5. Bs: LCMV and DipModel; Ft: sloreta and eloreta; and SPM: IID and EBB are strongly correlated, with a correlation greater than 0.9. Similarities also exist in the different toolboxes. For example, Bs: MNE is correlated with SPM: IID, COH, and EBB methods with a correlation of at least 0.7.

The empirical resolution matrix provides an intuitive understanding of spatial fidelity. In an ideal situation, the empirical resolution matrix should be diagonal, indicating that the source can be estimated accurately. However, in practical situations, the crosstalk and point spread between sources cannot be avoided because of the uncertainty of the inverse problem. Therefore, the actual resolution matrix was not an ideal diagonal matrix. The closer the resolution matrix is to the diagonal matrix, the better the performance of the corresponding MSI method. Figure 6 shows the ideal and empirical resolution matrices of the representative source imaging methods and the similarity of each resolution matrix to the diagonal matrix. The similarity was defined as the weights of diagonal elements and calculated by  $\gamma_{\hat{\mathbb{R}}} = \text{tr}(\hat{\mathbb{R}}) / \text{sum}(\hat{\mathbb{R}})$ , where  $\text{sum}(\hat{\mathbb{R}})$  was the sum of all elements in  $\hat{\mathbb{R}}$ . The closer  $\gamma_{\hat{\mathbb{R}}}$  is to 1, the closer  $\hat{\mathbb{R}}$  is to the diagonal matrix. It can be seen that the resolution matrices of Bs: LCMV, Ft: mne, and rv showed strong crosstalks and point spreads between sources, indicated when each source patch was activated, a large number of source patches were reconstructed. The situation was better for Bs: sLORETA and Ft: eloreta because the diagonal



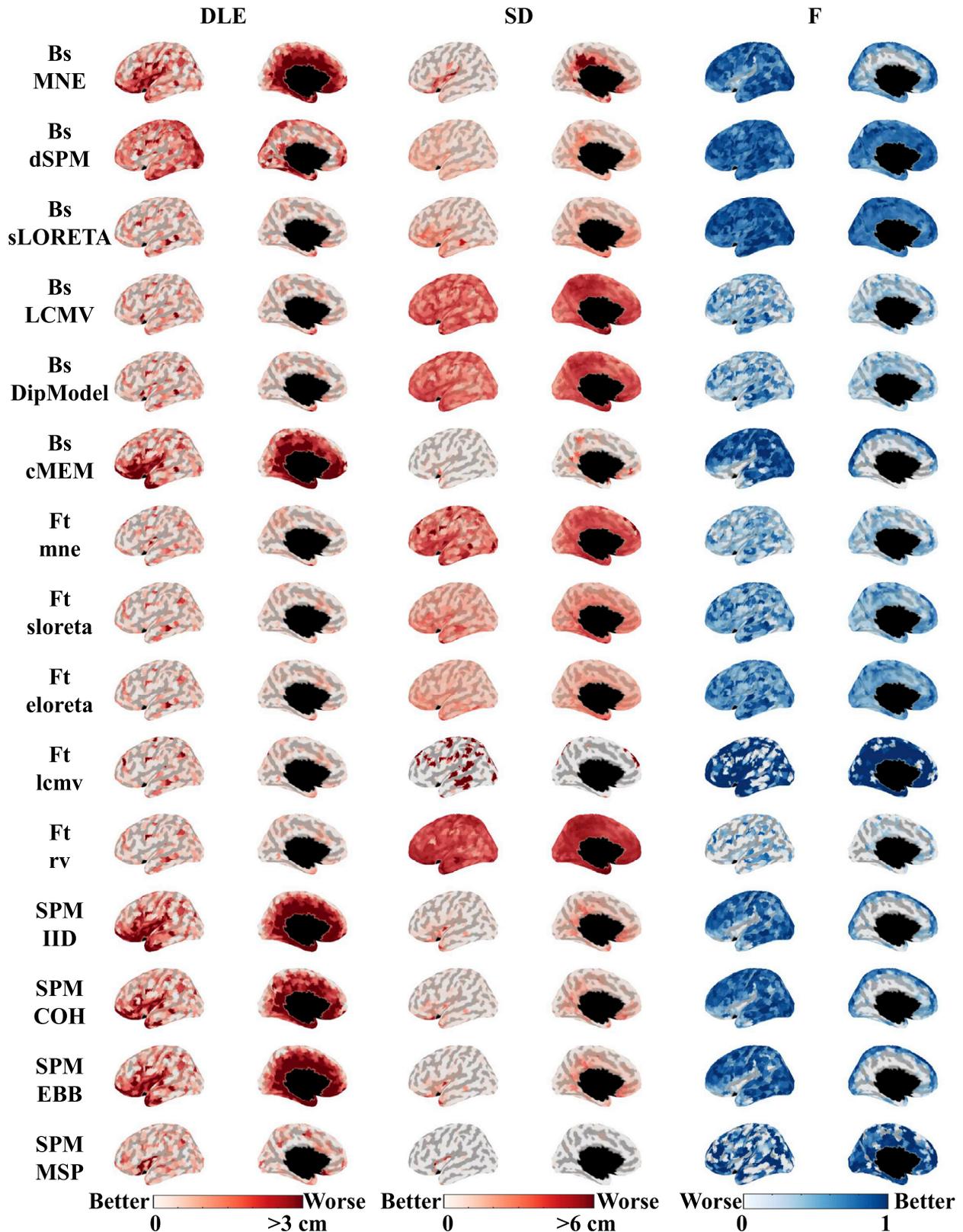
**Figure 3. Typical imaging results of different methods**

The source is located in the postcentral cortex. The imaging results below 20% of the maximum estimated source power are not displayed. The signal-to-noise ratio of the data was 20 dB.

elements could be clearly distinguished from the background. For Bs: MNE and cMEM, and SPM: MSP, although they had few crosstalks between sources, the diagonal elements were not clear enough, indicating that several incorrect source patches were reconstructed.

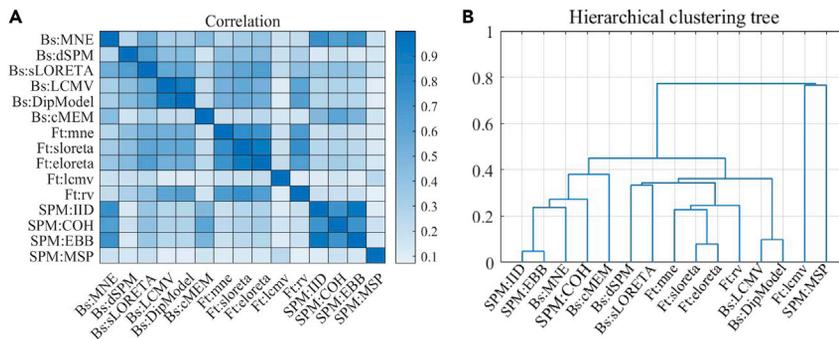
**Effects of source size and SNR.** We tested the performance of all source imaging methods by simulating two types of sources, patch and dipole, at different signal-to-noise ratios (SNRs). Histograms of the three metrics for the patch and dipole sources at an SNR of 20 dB are shown in Figure 7. The metrics distributions were obtained by calculating the metrics across different locations of the patch or dipole sources. The SD histogram shows that each method estimates similar source extents for dipole and focal sources, except for the Ft: lcmv method. The Ft: lcmv method showed the lowest DLE and SD and the highest F values. In other words, Ft: lcmv exhibited the best performance for imaging the focal sources. For Bs: sLORETA, LCMV, and DipModel, and Ft: sloreta, eloreta, lcmv, and rv, their DLE values were almost zero, indicating that they had better accuracies for localizing focal sources than patch sources. To further investigate the effects of source size on the DLE among the MNE-type and dipole scanning methods in Brainstorm and FieldTrip, we regressed the DLE values of the typical algorithm Ft: eloreta against the patch source size, and the results are shown in Figure 8. It can be seen that the DLE and the patch source size had a positive correlation, indicating that the ability of these methods to localize the center of the patch source worsens as the source size increases. Other methods seem to be insensitive to the source size.

We reduced the SNR from 20 to -20 dB and calculated the DLE and SD metrics. The F values were not calculated because F is more suitable for comparing methods under the same measurement scenario. In other words, comparing F values under different SNRs makes no sense. The effect of SNR on each method is shown in Figure 9. It can be seen that with a decrease in the SNR, the performance of each method gradually degrades, as indicated by the growing DLE and SD values. The performance of each method decreased sharply when the SNR decreased from -10 to -20 dB. In particular, the DLE values of the Bs: MNE, sLORETA and cMEM, and Ft: eloreta approaches were even greater than 30 mm at an SNR of -20 dB, far exceeding those of the others.



**Figure 4. Spatial distributions of metrics when simulating patch sources**

Topographic plots of dipole localization error (DLE), spatial dispersion (SD), and normalized variational free energy (F) metrics are shown. Because the results are similar in both hemispheres, only the left hemisphere's cortex results are shown. See also Figure S1.



**Figure 5. Correlation between different source imaging methods**

(A) Pearson correlation coefficient is calculated between each pair of methods. The color bar shows the correspondence between the color and the correlation coefficient. Darker blues indicate a stronger correlation between the two methods.

(B) Hierarchical clustering tree of correlation linkage between source imaging methods. The links between methods are represented as upside-down U-shaped lines. The height of the U indicates the degree of uncorrelation between the methods. The lower the height, the more similar the source imaging results between the two methods.

### Real OPM-MEG data

Cortical responses to median nerve stimulation have been widely studied using MEG, electroencephalography, and functional magnetic resonance imaging, and the approximate source location of the somatosensory cortex is well known (Molins et al., 2008). Therefore, we applied each method to a real 31-channel OPM-MEG dataset that recorded the cortical responses stimulated by the left median nerve. In the OPM-MEG experiment, the OPM sensors (QZFM Gen-2, QuSpin Inc., USA) measured the radial components of the magnetic fields and operated in a single-axis mode. The experiment was reviewed and approved by the Ethics Committee of Beihang University. Data were collected from the same subject as in the simulations, and the subject provided written informed consent. A detailed description of the dataset and preprocessing analysis is available in An et al. (2022).

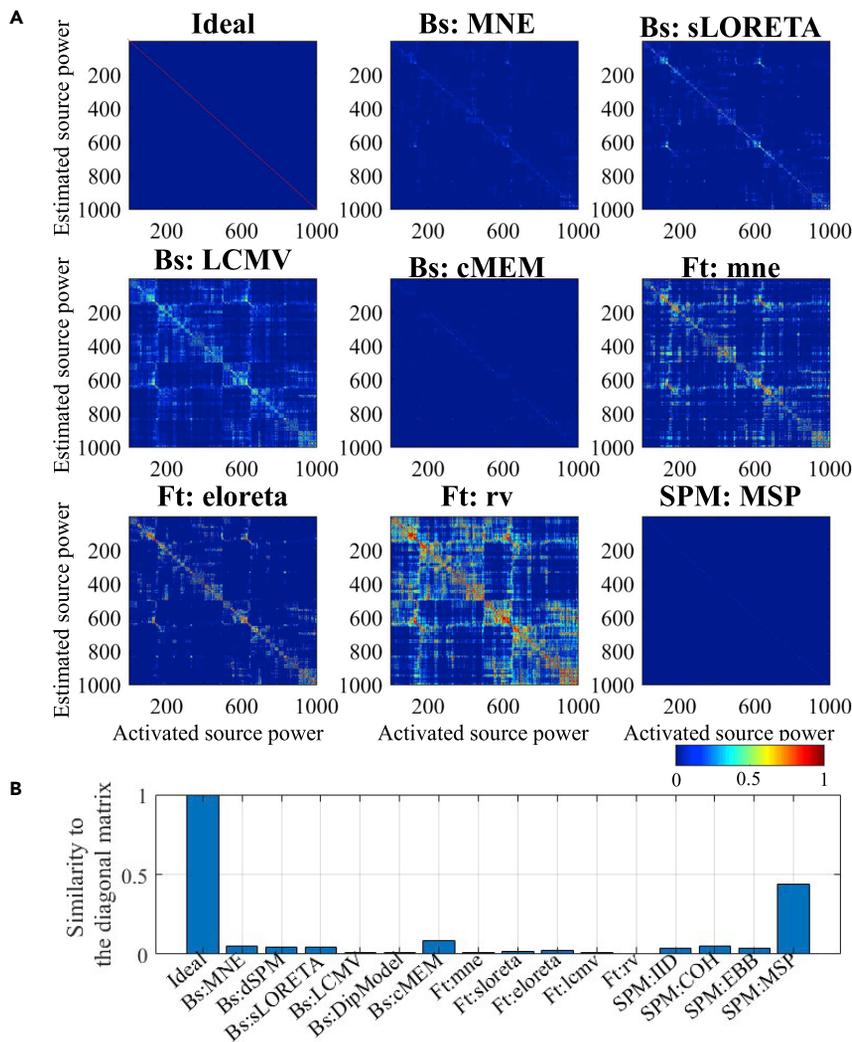
The source imaging results of the measured OPM-MEG data are shown in Figure 10. For each method, the source activation area was mainly in the right hemisphere, contrary to the stimulus side. The center of the reconstructed sources was near the postcentral gyrus, consistent with previous findings (Zetter et al., 2019; Hari and Puce, 2017; Backes et al., 2000). However, the estimated source extents of all methods showed significant differences. It is no surprise that the sparse algorithms Bs: cMEM and SPM: MSP showed more focal reconstruction results, whereas the MNE-type methods showed more blurred source results. The source results of SPM: IID, COH, and EBB were very similar because of the approximate source priors used to solve the inverse problem. The calculated variational free energy provided evidence for choosing the better one from all source imaging results. SPM: MSP had the highest F value in this experiment. It is reasonable to infer that the SPM: MSP imaging results can better explain the measurement data.

### DISCUSSION

The current study has two main objectives. We showed that the variational free energy under the Bayesian framework could be used to evaluate different source imaging results without knowing the ground truth of the brain sources, making it convenient to compare different source reconstruction results in practical applications. We also systematically assessed the spatial fidelity of all time-domain source imaging methods provided in three toolboxes—Brainstorm, FieldTrip, and SPM12—on an 85-channel OPM-MEG system. We attempt to provide users with an intuitive understanding of the characteristics of each method.

### Variational free energy as a metric to evaluate source imaging results

Source reconstruction of MEG is used to make neuroanatomical inferences and has a direct and significant impact on source-level connectivity estimation (Tait et al., 2021; Hincapié et al., 2017). Many source reconstruction methods are available (Becker et al., 2015; Baillet et al., 2001). However, there is no perfect approach to adapt to every situation because methods based on different assumptions about the source prior tend to have different trade-offs to solve the highly underdetermined inverse problem. In practical applications, there is no ground truth for brain sources. Therefore, the assessment of various source



**Figure 6. Illustration of resolution matrices**

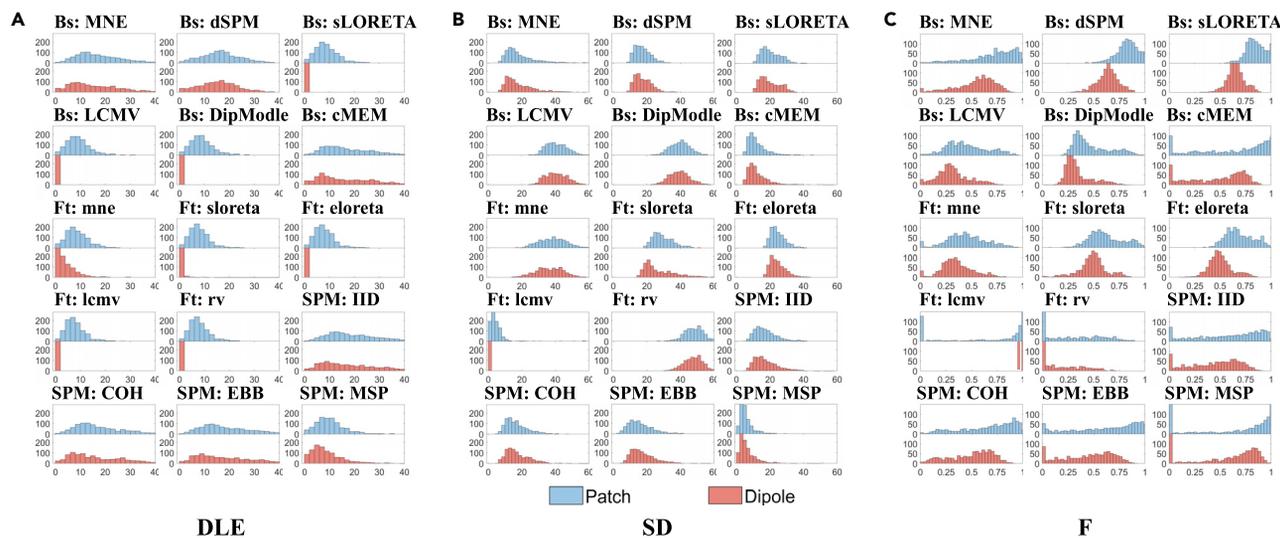
(A) Ideal and empirical resolution matrices  $\mathbb{R}$  of typical source imaging methods. The color bar is shown. For the actual empirical resolution matrix of the MSI method, the  $i$ -th row of represents that the  $i$ -th estimated source is originated from many sources (non-zero elements in the  $i$ -th row). The  $j$ -th column of  $\mathbb{R}$  represents that when the  $j$ -th source is activated, many other sources are estimated to be active.

(B) The similarity of each resolution matrix to the diagonal matrix.

imaging results cannot be performed using general metrics, such as DLE and SD. Our results demonstrated the effectiveness of variational free energy as an indicator for assessing different source imaging results (Figures 2, 3, and 4). The variational free energy incorporates the characteristics of DLE and SD metrics. Under the Bayesian framework and Laplace approximation, the variational free energy is an approximation of model evidence (Friston et al., 2007, 2008b). The MSP method combines variational free energy with the GS or ARD algorithm to search for and compare patch sources with different spatial priors (Friston et al., 2008a). In this study, the source distributions estimated by each MSI method were decoded into source covariance components, and the variational free energy was used to compare different models. The role of the variational free energy was further investigated.

### Characteristics of each source imaging method

Our results show differences in the source estimation results obtained using the same method but implemented in different toolboxes, such as the results of the MNE, sLORETA, and LCMV methods. In our study, we used the same generated datasets, head model, source model, and forward solutions across different

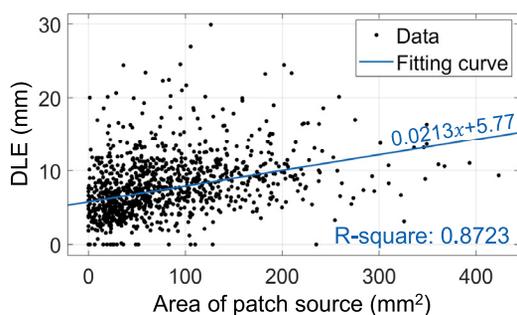


**Figure 7. Metrics histogram for each comparison method**

(A) Histogram of the dipole localization error (DLE). The results of patch sources are shown in blue, and those of dipole sources are shown in red. (B) Histogram of the spatial dispersion (SD). (C) Histogram of the variational free energy (F).

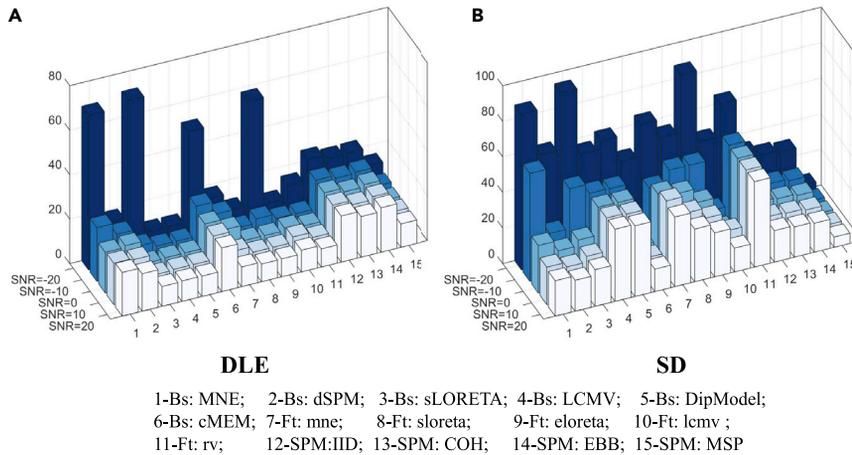
toolboxes to avoid differences in the source estimation results caused by these factors. Therefore, these differences arise mainly when solving the inverse problem. Before calculating the inverse operator, it was necessary to prepare the data. The preparation involves processes such as estimating data and noise covariance, data whitening, data regularization, depth weighting, and spatial or temporal projection. The inconsistency in these process implementations causes differences in the source estimation results. Additionally, discrepancies exist in the core calculation of the inverse operator for the same method. Therefore, a comparison of methods across different toolboxes is required.

In this study, we evaluated MSI methods provided in three commonly used toolboxes: Brainstorm, FieldTrip, and SPM12. The characteristics of all the methods are summarized in Table 2. Similar to the conclusions of other studies, there is no perfect algorithm (Tait et al., 2021). Here, we highlight the characteristics of each type of method and aim to assist researchers in choosing an appropriate source imaging method. For the MNE-like methods, the variant LORETA-type methods showed better performance than the MNE method in each toolbox. For example, there was improved source localization accuracy of Bs: sLORETA in the insula and cingulate gyrus compared with that of Bs: MNE (Figure 4). Bs: LCMV tended to overestimate the source extent, whereas Ft: lcmv was more suitable for reconstructing the focal sources. The performance of Bs: LCMV was closer to the Bs: DipModel method. Dipole scanning methods always exhibit a larger activation area than other methods. The sparse algorithms Bs: cMEM and Ft: MSP exhibited excellent performance in reconstructing the source extent. However, it should be noted that they also had the risk of mispredicting the patch source location compared with the other methods. That is, researchers need to balance two situations: blurred estimation results covering the real source activation area and a more precise source range, but the source location may be biased.



**Figure 8. Regression results between the dipole localization error (DLE) values of Ft: eloreta and the patch source size**

The patch source size is the surface area of all triangle mesh in each patch source. The linear regression is performed using the least absolute residual method. The correlation coefficient R-square is shown. The results are obtained in the datasets with the signal-to-noise (SNR) of 20 dB.

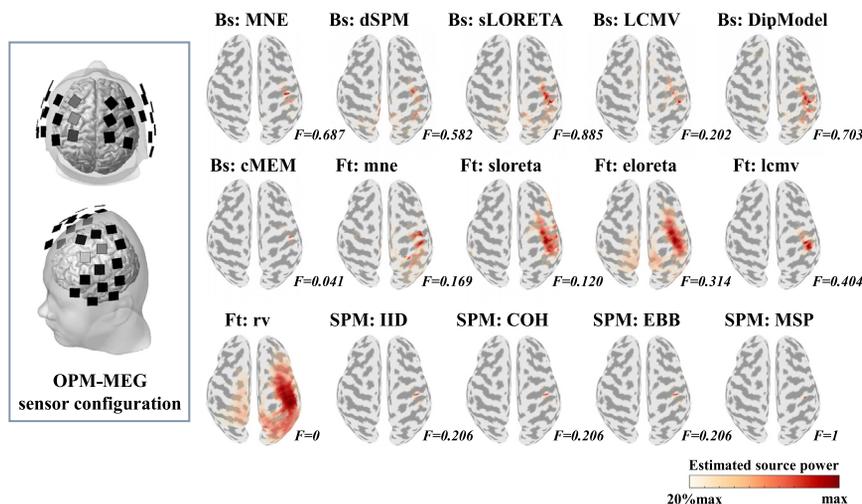


**Figure 9. Effect of the signal-to-noise ratio (SNR) on each source imaging method**

(A) SNR effect on the dipole localization error (DLE) metric.  
(B) SNR effect on the SD metric.

### Limitations of the study

Plenty of MSI methods have been proposed to solve the inverse problem in MEG. In this study, we only focused on evaluating MSI methods implemented in Brainstorm, FieldTrip, and SPM12. For MSI methods, several aspects should be considered in source estimation, such as spatial fidelity, source amplitude reconstruction, and source estimation results, when multiple sources are activated. Our work focused only on the spatial accuracy evaluation of different MSI methods in three commonly used toolboxes for detecting a single source. Comparing the estimated source amplitudes is challenging because the source amplitude units vary among toolboxes. If the estimated source amplitudes are normalized for comparison, they seem to lose their practical value in indicating the intensity of brain sources. A comparison of MSI methods for temporal reconstruction should be conducted within each toolbox. In addition, evaluating situations in which multiple sources are activated is significant. The source reconstruction results may be affected by several factors such as coherency, inconsistent intensity, and distance from each other.



**Figure 10. OPM-MEG sensor configuration and source imaging results of the optically pumped magnetometer (OPM)-based MEG data under left median nerve stimulation**

The variational free energy of each method was normalized using the minimum and maximum F values of all 15 inversion methods on this dataset. The normalized variational free energy is shown.

**Table 2. Characteristics summary of each time-domain source imaging method in Brainstorm, FieldTrip, and SPM12**

Software	Method	Accuracy of source localization	Accuracy of source extent estimation	Better for focus source than extended source?	Robustness to high sensor noise	Additional remark
Brainstorm v3.211107	MNE	***	***	Equal	***	Note 1
	dSPM	****	****	Equal	****	–
	sLORETA	*****	****	Better	***	–
	LCMV	****	**	Better	****	Note 2-1
	DipModel	****	**	Better	****	Note 2-1
	cMEM	***	*****	Equal	***	Note 1
FieldTrip v20201023	mne	****	**	Slightly better	****	–
	sloreta	****	***	Better	****	Note 2-2
	eloreta	****	***	Better	***	Note 2-2
	lcmv	****	***	Better	****	–
	rv	****	**	Better	****	–
SPM12 v7771	IID	***	****	Equal	****	Note 1; note 2-3
	COH	***	****	Equal	****	Note 1
	EBB	***	****	Equal	****	Note 1; note 2-3
	MSP	****	****	Equal	****	–

\*Rate of performance of each method. A higher \* indicates better performance. Note 1: worse in imaging sources in the insula and cingulate gyrus; Note 2-i: the source imaging results of *i*th group of methods are strongly correlated.

Each MSI method has its unique parameters. In our study, the parameters of each method, except for the SNR, used their default or recommended settings in each toolbox. It should be noted that changes in parameters may affect the imaging results. Therefore, when using MSI methods with different parameter settings, it is necessary to consider the differences in the imaging results. Using the variational free energy to compare different imaging results is recommended. In addition, it is possible to use free energy as the objective function to optimize the parameters for each MSI method. When different parameters are set, the estimated source power distribution corresponding to each parameter can be encoded into the source prior, and the free energy can be calculated. The parameters with maximum free energy can be used as the final parameter values.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human subjects
- METHOD DETAILS
  - Inverse problem
  - Source imaging methods to be evaluated
  - Implementations of source imaging methods
  - Evaluation metrics

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2022.105177>.

### ACKNOWLEDGMENTS

We received financial support from the Key Laboratory of Weak Magnetic Detection Technology, Ministry of Education, and the Beijing Municipal Natural Science Foundation (No. 4212012).

## AUTHOR CONTRIBUTIONS

Conceptualization, N.A.; Methodology, N.A. and F.Z.C.; Software, W.L.; Formal analysis, W.L.W., W.N.X., and C.H.W.; Data curation, W.L.W.; Validation, W.N.X.; Supervision, M.X. and Y.G.; Project administration, X.L.N.; Writing – original draft, N.A.; Writing – review and editing, X.L.N.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 9, 2022

Revised: July 24, 2022

Accepted: September 19, 2022

Published: October 21, 2022

## REFERENCES

- Amblard, C., Lapalme, E., and Lina, J.M. (2004). Biomagnetic source detection by maximum entropy and graphical models. *IEEE Trans. Biomed. Eng.* *51*, 427–442. <https://doi.org/10.1109/TBME.2003.820999>.
- An, N., Cao, F., Li, W., Wang, W., Xu, W., Wang, C., Xiang, M., Gao, Y., Sui, B., Liang, A., et al. (2022). Imaging somatosensory cortex responses measured by OPM-MEG: variational free energy-based spatial smoothing estimation approach. *iScience* *25*, 103752. <https://doi.org/10.1016/j.isci.2022.103752>.
- Backes, W.H., Mess, W.H., van Kranen-Mastenbroek, V., and Reulen, J.P.H. (2000). Somatosensory cortex responses to median nerve stimulation: fMRI effects of current amplitude and selective attention. *Clin. Neurophysiol.* *111*, 1738–1744. [https://doi.org/10.1016/S1388-2457\(00\)00420-X](https://doi.org/10.1016/S1388-2457(00)00420-X).
- Baillet, S., Mosher, J.C., and Leahy, R.M. (2001). Electromagnetic brain mapping. *IEEE Signal Proc. Mag.* *18*, 14–30. <https://doi.org/10.1109/79.962275>.
- Becker, H., Albera, L., Comon, P., Gribonval, R., Wendling, F., and Merlet, I. (2015). Brain-source imaging: from sparse to tensor models. *IEEE Signal Proc. Mag.* *32*, 100–112. <https://doi.org/10.1109/MSP.2015.2413711>.
- Belardinelli, P., Ortiz, E., Barnes, G., Noppeney, U., and Preissl, H. (2012). Source reconstruction accuracy of MEG and EEG Bayesian inversion approaches. *PLoS One* *7*, e51985. <https://doi.org/10.1371/journal.pone.0051985>.
- Boto, E., Bowtell, R., Krüger, P., Fromhold, T.M., Morris, P.G., Meyer, S.S., Barnes, G.R., and Brookes, M.J. (2016). On the potential of a new generation of magnetometers for MEG: a beamformer simulation study. *PLoS One* *11*, e0157655. <https://doi.org/10.1371/journal.pone.0157655>.
- Boto, E., Holmes, N., Leggett, J., Roberts, G., Shah, V., Meyer, S.S., Muñoz, L.D., Mullinger, K.J., Tierney, T.M., Bestmann, S., et al. (2018). Moving magnetoencephalography towards real-world applications with a wearable system. *Nature* *555*, 657–661. <https://doi.org/10.1038/nature26147>.
- Cao, F., An, N., Wang, W., Xu, W., Li, W., Wang, C., Xiang, M., Gao, Y., and Ning, X. (2022). Effects of different head models in wearable OPM-MEG. *IEEE Trans. Instrum. Meas.* <https://doi.org/10.1109/TIM.2022.3192287>.
- Chesca, B., John, D., and Mellor, C.J. (2015). Flux-coherent series SQUID array magnetometers operating above 77 K with superior white flux noise than single-SQUIDs at 4.2 K. *Appl. Phys. Lett.* *107*, 162602. <https://doi.org/10.1063/1.4932969>.
- Chowdhury, R.A., Lina, J.M., Kobayashi, E., and Grova, C. (2013). MEG source localization of spatially extended generators of epileptic activity: comparing entropic and hierarchical bayesian approaches. *PLoS One* *8*, e55969. <https://doi.org/10.1371/journal.pone.0055969>.
- Daducci, A., Gerhard, S., Griffa, A., Lemkaddem, A., Cammoun, L., Gigandet, X., Meuli, R., Hagmann, P., and Thiran, J.P. (2012). The connectome mapper: an open-source processing pipeline to map connectomes with MRI. *PLoS One* *7*, e48121. <https://doi.org/10.1371/journal.pone.0048121>.
- Dale, A.M., Liu, A.K., Fischl, B.R., Buckner, R.L., Belliveau, J.W., Lewine, J.D., and Halgren, E. (2000). Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron* *26*, 55–67. [https://doi.org/10.1016/S0896-6273\(00\)81138-1](https://doi.org/10.1016/S0896-6273(00)81138-1).
- Faley, M.I., Poppe, U., Dunin-Borkowski, R.E., Schiek, M., Boers, F., Chocholacs, H., Dammers, J., Eich, E., Shah, N.J., Ermakov, A.B., et al. (2012). High-Tc DC SQUIDs for magnetoencephalography. *IEEE Trans. Appl. Supercond.* *23*, 1600705. <https://doi.org/10.1109/TASC.2012.2229094>.
- Fischl, B. (2012). FreeSurfer. *NeuroImage* *62*, 774–781. <https://doi.org/10.1016/j.neuroimage.2012.01.021>.
- Friston, K., Chu, C., Mourão-Miranda, J., Hulme, O., Rees, G., Penny, W., and Ashburner, J. (2008b). Bayesian decoding of brain images. *NeuroImage* *39*, 181–205. <https://doi.org/10.1016/j.neuroimage.2007.08.013>.
- Friston, K., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., Henson, R., Flandin, G., and Mattout, J. (2008a). Multiple sparse priors for the M/EEG inverse problem. *NeuroImage* *39*, 1104–1120. <https://doi.org/10.1016/j.neuroimage.2007.09.048>.
- Friston, K., Mattout, J., Trujillo-Barreto, N., Ashburner, J., and Penny, W. (2007). Variational free energy and the Laplace approximation. *NeuroImage* *34*, 220–234. <https://doi.org/10.1016/j.neuroimage.2006.08.035>.
- Gramfort, A., Papadopoulos, T., Olivi, E., and Clerc, M. (2010). OpenMEEG: opensource software for quasistatic bioelectromagnetics. *Biomed. Eng. Online* *9*, 1–20. <https://doi.org/10.1186/1475-925X-9-45>.
- Grova, C., Daunizeau, J., Lina, J.M., Bénar, C.G., Benali, H., and Gotman, J. (2006). Evaluation of EEG localization methods using realistic simulations of interictal spikes. *NeuroImage* *29*, 734–753. <https://doi.org/10.1016/j.neuroimage.2005.08.053>.
- Hämäläinen, M.S., and Ilmoniemi, R.J. (1994). Interpreting magnetic fields of the brain: minimum norm estimates. *Med. Biol. Eng. Comput.* *32*, 35–42. <https://doi.org/10.1007/BF02512476>.
- Hari, R., and Puce, A. (2017). *MEG-EEG Primer* (Oxford University Press).
- Hauk, O., Wakeman, D.G., and Henson, R. (2011). Comparison of noise-normalized minimum norm estimates for MEG analysis using multiple resolution metrics. *NeuroImage* *54*, 1966–1974. <https://doi.org/10.1016/j.neuroimage.2010.09.053>.
- Hedrich, T., Pellegrino, G., Kobayashi, E., Lina, J.M., and Grova, C. (2017). Comparison of the spatial resolution of source imaging techniques in high-density EEG and MEG. *NeuroImage* *157*, 531–544. <https://doi.org/10.1016/j.neuroimage.2017.06.022>.
- Henson, R.N., Abdulrahman, H., Flandin, G., and Litvak, V. (2019). Multimodal integration of M/EEG and f/MRI data in SPM12. *Front. Neurosci.* *13*, 300. <https://doi.org/10.3389/fnins.2019.00300>.
- Hincapié, A.S., Kujala, J., Mattout, J., Pascarella, A., Daligault, S., Delpuech, C., Mery, D., Cosmelli, D., and Jerbi, K. (2017). The impact of MEG source reconstruction method on source-space connectivity estimation: a comparison between minimum-norm solution and beamforming. *NeuroImage* *156*, 29–42. <https://doi.org/10.1016/j.neuroimage.2017.04.038>.

- Iivanainen, J., Zetter, R., and Parkkonen, L. (2020). Potential of on-scalp MEG: Robust detection of human visual gamma-band responses. *Hum. Brain Mapp.* 41, 150–161. <https://doi.org/10.1002/hbm.24795>.
- Iwasaki, M., and Nakasato, N. (2019). In MEG in epilepsy and pre-surgical functional mapping, S.S. Magnetoencephalography and C.J. Aine, eds. (Springer-Verlag Berlin Heidelberg), pp. 821–842. <https://doi.org/10.1007/978-3-642-33045-2>.
- Jaiswal, A., Nenonen, J., Stenroos, M., Gramfort, A., Dalal, S.S., Westner, B.U., Litvak, V., Mosher, J.C., Schoffelen, J.M., Witton, C., et al. (2020). Comparison of beamformer implementations for MEG source localization. *Neuroimage* 216, 116797. <https://doi.org/10.1016/j.neuroimage.2020.116797>.
- Janati, H., Bazeille, T., Thirion, B., Cuturi, M., and Gramfort, A. (2019). Group level MEG/EEG source imaging via optimal transport: minimum Wasserstein estimates. In *International Conference on Information Processing in Medical Imaging* (Springer), pp. 743–754. [https://doi.org/10.1007/978-3-030-20351-1\\_58](https://doi.org/10.1007/978-3-030-20351-1_58).
- Knappe, S., Sander, T., and Trahms, L. (2014). In *Optically-pumped magnetometers for MEG*, S.S. Magnetoencephalography and C.J. Aine, eds. (Springer-Verlag Berlin Heidelberg), pp. 993–999. <https://doi.org/10.1007/978-3-642-33045-2>.
- Lin, F.H., Witzel, T., Hämäläinen, M.S., Dale, A.M., Belliveau, J.W., and Stufflebeam, S.M. (2004). Spectral spatiotemporal imaging of cortical oscillations and interactions in the human brain. *Neuroimage* 23, 582–595. <https://doi.org/10.1016/j.neuroimage.2004.04.027>.
- Liu, A.K., Dale, A.M., and Belliveau, J.W. (2002). Monte Carlo simulation studies of EEG and MEG localization accuracy. *Hum. Brain Mapp.* 16, 47–62. <https://doi.org/10.1002/hbm.10024>.
- López, J.D., Litvak, V., Espinosa, J.J., Friston, K., and Barnes, G.R. (2014). Algorithmic procedures for Bayesian MEG/EEG source reconstruction in SPM. *Neuroimage* 84, 476–487. <https://doi.org/10.1016/j.neuroimage.2013.09.002>.
- MacKay, D.J. (1992). Bayesian interpolation. *Neural Comput.* 4, 415–447. <https://doi.org/10.1162/neco.1992.4.3.415>.
- Mattout, J., Phillips, C., Penny, W.D., Rugg, M.D., and Friston, K.J. (2006). MEG source localization under multiple constraints: an extended Bayesian framework. *Neuroimage* 30, 753–767. <https://doi.org/10.1016/j.neuroimage.2005.10.037>.
- Molins, A., Stufflebeam, S.M., Brown, E.N., and Hämäläinen, M.S. (2008). Quantification of the benefit from integrating MEG and EEG data in minimum  $\ell_2$ -norm estimation. *Neuroimage* 42, 1069–1077. <https://doi.org/10.1016/j.neuroimage.2008.05.064>.
- Mosher, J.C., Leahy, R.M., and Lewis, P.S. (1999). EEG and MEG: forward solutions for inverse methods. *IEEE Trans. Biomed. Eng.* 46, 245–259. <https://doi.org/10.1109/10.748978>.
- Mosher, J.C., Lewis, P.S., and Leahy, R.M. (1992). Multiple dipole modeling and localization from spatio-temporal MEG data. *IEEE Trans. Biomed. Eng.* 39, 541–557. <https://doi.org/10.1109/10.141192>.
- Niklas, K. (2022). Wasserstein Distance. <https://github.com/nklb/wasserstein-distance>.
- O’Neill, G.C., Barry, D.N., Tierney, T.M., Mellor, S., Maguire, E.A., and Barnes, G.R. (2021). Testing covariance models for MEG source reconstruction of hippocampal activity. *Sci. Rep.* 11, 1–13. <https://doi.org/10.1038/s41598-021-96933-0>.
- Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* <https://doi.org/10.1155/2011/156869>.
- Pascual-Marqui, R.D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Methods Find. Exp. Clin. Pharmacol.* 24, 5–12.
- Pascual-Marqui, R.D., Lehmann, D., Koukkou, M., Kochi, K., Anderer, P., Saletu, B., Tanaka, H., Hirata, K., John, E.R., Prichep, L., et al. (2011). Assessing interactions in the brain with exact low-resolution electromagnetic tomography. *Philos. Trans. R. Soc., A* 369, 3768–3784. <https://doi.org/10.1098/rsta.2011.0081>.
- Roberts, G., Holmes, N., Alexander, N., Boto, E., Leggett, J., Hill, R.M., Shah, V., Rea, M., Vaughan, R., Maguire, E.A., et al. (2019). Towards OPM-MEG in a virtual reality environment. *Neuroimage* 199, 408–417. <https://doi.org/10.1016/j.neuroimage.2019.06.010>.
- Samuelsson, J.G., Peled, N., Mamashli, F., Ahveninen, J., and Hämäläinen, M.S. (2021). Spatial fidelity of MEG/EEG source estimates: a general evaluation approach. *Neuroimage* 224, 117430. <https://doi.org/10.1016/j.neuroimage.2020.117430>.
- Stenroos, M., Hunold, A., and Hauelsen, J. (2014). Comparison of three-shell and simplified volume conductor models in magnetoencephalography. *Neuroimage* 94, 337–348. <https://doi.org/10.1016/j.neuroimage.2014.01.006>.
- 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.* <https://doi.org/10.1155/2011/879716>.
- Tait, L., Özkan, A., Szul, M.J., and Zhang, J. (2021). A systematic evaluation of source reconstruction of resting MEG of the human brain with a new high-resolution atlas: performance, precision, and parcellation. *Hum. Brain Mapp.* 42, 4685–4707. <https://doi.org/10.1002/hbm.25578>.
- Trujillo-Barreto, N.J., Aubert-Vázquez, E., and Valdés-Sosa, P.A. (2004). Bayesian model averaging in EEG/MEG imaging. *Neuroimage* 21, 1300–1319. <https://doi.org/10.1016/j.neuroimage.2003.11.008>.
- 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. <https://doi.org/10.1109/10.623056>.
- Westner, B.U., Dalal, S.S., Gramfort, A., Litvak, V., Mosher, J.C., Oostenveld, R., and Schoffelen, J.M. (2022). A unified view on beamformers for M/EEG source reconstruction. *Neuroimage* 246, 118789. <https://doi.org/10.1016/j.neuroimage.2021.118789>.
- Wipf, D., and Nagarajan, S. (2009). A unified Bayesian framework for MEG/EEG source imaging. *Neuroimage* 44, 947–966. <https://doi.org/10.1016/j.neuroimage.2008.02.059>.
- Zerouali, Y., Herry, C.L., Jemel, B., and Lina, J.M. (2011). Localization of synchronous cortical neural sources. *IEEE Trans. Biomed. Eng.* 60, 770–780. <https://doi.org/10.1109/TBME.2011.2176938>.
- Zetter, R., Iivanainen, J., and Parkkonen, L. (2019). Optical co-registration of MRI and on-scalp MEG. *Sci. Rep.* 9, 1–9. <https://doi.org/10.1038/s41598-019-41763-4>.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Brainstorm v3.211107	Tadel et al., 2011	<a href="https://neuroimage.usc.edu/bst/download.php">https://neuroimage.usc.edu/bst/download.php</a>
FieldTrip v20201023	Oostenveld et al., 2011	<a href="https://www.fieldtriptoolbox.org/download/">https://www.fieldtriptoolbox.org/download/</a>
SPM12	The Wellcome Centre for Human Neuroimaging	<a href="https://www.fil.ion.ucl.ac.uk/spm/">https://www.fil.ion.ucl.ac.uk/spm/</a>
OpenMEEG	Gramfort et al., 2010	<a href="http://openmeeeg.gforge.inria.fr">http://openmeeeg.gforge.inria.fr</a>
FreeSurfer v7.1.0	Fischl, 2012	<a href="https://surfer.nmr.mgh.harvard.edu/">https://surfer.nmr.mgh.harvard.edu/</a>
MATLAB vR2019b	MathWorks	<a href="https://www.mathworks.com/products/matlab.html">https://www.mathworks.com/products/matlab.html</a>
Wasserstein distance	Niklas, 2022	<a href="https://github.com/nklb/wasserstein-distance">https://github.com/nklb/wasserstein-distance</a>
Algorithms	This paper	<a href="https://doi.org/10.5281/zenodo.7066051">https://doi.org/10.5281/zenodo.7066051</a>
Other		
Human subjects	This paper	N/A

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Xiaolin Ning ([ningxiaolin@buaa.edu.cn](mailto:ningxiaolin@buaa.edu.cn)).

#### Materials availability

This study did not generate new unique reagents.

#### Data and code availability

- All data reported in this article will be shared by the [lead contact](#) on request.
- All original code has been deposited at Zenodo and is publicly available as of the date of publication. DOIs are listed in the [key resources table](#).
- Any additional information required to reanalyze the data reported in this article is available from the [lead contact](#) on request

### EXPERIMENTAL MODEL AND SUBJECT DETAILS

#### Human subjects

One healthy right-handed subject (male, 27-year-old) participated in this study. This study was reviewed and approved by the Ethics Committee of Beihang University and the subject provided written informed consent for the experimental procedure in accordance with the Declaration of Helsinki.

### METHOD DETAILS

#### Inverse problem

For MEG, there is a common assumption that the sources of brain activity can be modeled as a large number of equivalent current dipoles, with their locations fixed and distributed on the cortex and their orientation perpendicular to the local cortex mesh (Liu et al., 2002). Under this assumption, the relationship between the measured MEG data  $\mathbf{Y}$  and the distributed source  $\mathbf{J}$  can be described by a linear model:

$$\mathbf{Y} = \mathbf{L}\mathbf{J} + \boldsymbol{\varepsilon}, \quad (\text{Equation 1})$$

where  $\mathbf{Y}$  is an  $n_c \times n_t$  data matrix measured at  $n_c$  channels and  $n_t$  time samples and  $\mathbf{J}$  is an  $n_d \times n_t$  matrix representing the source amplitudes of the  $n_d$  dipoles.  $\mathbf{L}$  ( $n_c \times n_d$ ) is the forward solution, also known as the lead field matrix, and each column describes the sensitivity of the sensor array to a source with unit

amplitude at a given location and orientation. This lead field matrix can be obtained from the solutions of the forward problem in quasi-static approximation (Mosher et al., 1999).  $\epsilon$  ( $n_c \times n_t$ ) is the additive measurement noise.

The inverse problem for the distributed source model involves estimating the source amplitude  $\mathbf{J}$  based on the measured MEG data and the obtained lead field. The solution to this inverse problem can be derived in various ways, such as the Tikhonov regularization, minimization of expected error, and Bayesian formulation (Liu et al., 2002). These methods yield a similar expression for the solution, as follows:

$$\hat{\mathbf{J}} = \mathbf{Q}_J \mathbf{L}^T (\mathbf{L} \mathbf{Q}_J \mathbf{L}^T + \mathbf{R})^{-1} \mathbf{Y}, \quad (\text{Equation 2})$$

where  $\hat{\mathbf{J}}$  is the estimated source amplitude and  $\mathbf{R}$  is the covariance matrix of the noise  $\epsilon$ .  $\mathbf{Q}_J$  is the source covariance matrix that can be regarded as the prior of the sources. The superscript "T" denotes transpose.  $\mathbf{Q}_J \mathbf{L}^T (\mathbf{L} \mathbf{Q}_J \mathbf{L}^T + \mathbf{R})^{-1}$  is usually denoted as  $\mathbf{W}$ , and is called the inverse operator. The inverse problem of MEG is ill-posed, meaning the solution is not unique. Therefore, additional constraints should be added to the sources, that is, the assumption regarding the source prior  $\mathbf{Q}_J$ .

### Source imaging methods to be evaluated

In this study, we evaluated MNE-like solutions, beamformer, dipole scanning, and two types of sparse algorithms. In this section, we briefly review these source imaging methods and clarify their implementation details.

#### MNE-like solutions

The most widely used solution for the MEG inverse problem is the MNE method. If there is no available source prior, the MNE method usually assumes that the source prior covariance  $\mathbf{Q}_J = \mathbf{I}$  and the noise covariance  $\mathbf{R} = \lambda^2 \mathbf{C}$ , where  $\mathbf{I}_{n_d \times n_d}$  is the identity matrix, and  $\mathbf{C}$  is the estimated noise covariance.  $\lambda^2$  is a regularization parameter estimated from the noise covariance (Lin et al., 2004). The MNE solution is expressed as follows:

$$\hat{\mathbf{J}}_{MNE} = \mathbf{W}_{MNE} \mathbf{Y} = \mathbf{L}^T (\mathbf{L} \mathbf{L}^T + \lambda^2 \mathbf{C})^{-1} \mathbf{Y}. \quad (\text{Equation 3})$$

However, the MNE approach has been shown that it tends to prefer superficial dipoles. Various methods have been developed to eliminate this tendency, such as dSPM (Dale et al., 2000), sLORETA (Pascual-Marqui et al., 2002), and eLORETA (Pascual-Marqui et al., 2011). These variants attempt to standardize the values of the estimated current density using its expected SD and use the normalized results to generate source mapping. This standardization can be defined as follows:

$$P_i = \frac{\hat{\mathbf{J}}_{MNE}(i) \hat{\mathbf{J}}_{MNE}^T(i)}{S_i}, \quad (\text{Equation 4})$$

where  $\hat{\mathbf{J}}_{MNE}(i)$  is the estimated time courses of the  $i$ -th dipole and  $S_i$  is its estimated variance.  $P_i$  is the standardized value of the source power. dSPM regards the variance originated exclusively by the measurement noise, which gives  $S_i = \mathbf{W}_i \mathbf{C} \mathbf{W}_i^T$ , where  $\mathbf{W}_i$  is the  $i$ -th row of  $\mathbf{W}_{MNE}$ . sLORETA takes the variance of the actual sources into account, not only noise, that is,  $S_i = \mathbf{W}_i (\mathbf{L} \mathbf{L}^T + \lambda^2 \mathbf{C}) \mathbf{W}_i^T$ . For eLORETA, it estimates the source variance iteratively. It uses the estimated source variances of the previous iteration to compose the source covariance, that is,  $\mathbf{Q}_J = \text{diag}(1/S_1, 1/S_2, \dots, 1/S_{N_d})$ , and then updates the inverse operator to calculate a new source variance  $S'_i$  where

$$S'_i = \mathbf{W}_i (\mathbf{L} \mathbf{Q}_J \mathbf{L}^T + \lambda^2 \mathbf{C})^{-1} \mathbf{W}_i^T. \quad (\text{Equation 5})$$

The inverse operator continues to update until it converges, and the final source variance is obtained.

Some methods directly assume a prior on the source covariance, such as the SPM-COH and SPM-EBB methods, and then solve the inverse problem using (Equation 2). SPM-COH defines a spatial smoothing before the source covariance, that is  $\mathbf{Q}_J = \exp(\sigma \mathbf{G}_L)$  where  $\sigma$  is the smoothness level and  $\mathbf{G}_L = \mathbf{A} - \text{diag}(\sum_{k=1}^{N_s} \mathbf{A}_{ik})$  is a graph Laplacian matrix (López et al., 2014).  $\mathbf{A}$  is the adjacent matrix and  $\mathbf{A}_i$  is its  $i$ -th row. For SPM-EBB, it assumes the source covariance as a diagonal matrix with its diagonal elements  $q_i = (\mathbf{L}_i^T \mathbf{Q}_Y^{-1} \mathbf{L}_i)^{-1} / (\mathbf{L}_i^T \mathbf{L}_i)^{-1}$ , where  $\mathbf{L}_i$  is the lead field of the  $i$ -th dipole and  $\mathbf{Q}_Y$  is the data covariance (Belardinelli et al., 2012; O'Neill et al., 2021).

### Beamformer

The LCMV method is a widely used beamformer for imaging time-domain data (Van Veen et al., 1997). It defines a spatial filter  $\mathbf{W}$ , similar to the inverse operator. In contrast, it takes  $\mathbf{Q}_J = \mathbf{L}^T \mathbf{Q}_Y^{-1} \mathbf{L}$  in  $\mathbf{W}$  and provides  $\mathbf{W} = (\mathbf{L}^T \mathbf{Q}_Y^{-1} \mathbf{L})^{-1} \mathbf{L}^T \mathbf{Q}_Y^{-1}$ . The spatial filter is used to reconstruct the time course of each dipole. For source localization, the LCMV estimates the variance or strength of each dipole and normalizes it using the estimated noise covariance. The pseudo-statistics of each dipole can be expressed as:

$$P_i = \frac{\text{tr}[\mathbf{L}_i^T \mathbf{Q}_Y^{-1} \mathbf{L}_i]}{\text{tr}[\mathbf{L}_i^T \mathbf{R}^{-1} \mathbf{L}_i]}, \quad (\text{Equation 6})$$

where  $\text{tr}[\cdot]$  is the trace of a matrix and  $\text{tr}[\mathbf{L}_i^T \mathbf{Q}_Y^{-1} \mathbf{L}_i]$  is the estimated source power. This normalization estimate is known as the Neural Activity Index (Van Veen et al., 1997).

### Dipole scanning

The dipole scanning method is the simplest model for solving the inverse problem. It scans the source space with a single dipole and computes the goodness of fit (GOF) (Baillet et al., 2001). The GOF of the  $i$ -th dipole is calculated as follows:

$$\text{GOF}_i = \frac{|\mathbf{Y} - \mathbf{W}_i \mathbf{Y}|_F}{|\mathbf{Y}|_F}, \quad (\text{Equation 7})$$

where  $|\cdot|_F$  is the Frobenius norm of a matrix, and  $\mathbf{W}_i = \mathbf{L}_i \mathbf{L}_i^{-1}$ . The calculated GOFs at all dipole locations were used to produce a dipole scanning map, which could be viewed as the source imaging results.

MEM: The maximum entropy on the mean (MEM) method (Chowdhury et al., 2013) is a sparse algorithm used for source imaging. Similar to the aim of other sparse algorithms, MEM is not only to estimate source locations but also to limit the spatial extent of sources. The inactive brain region is set to zero. It models brain activity using non-overlapping cortical parcels with specific local spatial smoothness; that is, the spatial priors are constrained by  $\mathbf{Q}_J$ . The MEM uses entropy maximization as the criterion to estimate the contrast of the current density within each active parcel. The estimated  $\hat{\mathbf{J}}$  is described as an  $n_J$ -dimensional continuous random variable with a probability density  $dp(\mathbf{J})$ . Kullback's  $\mu$ -entropy is defined as follows (Amblard et al., 2004):

$$S_\mu(dp) = - \int \log \frac{dp}{d\mu} dp, \quad (\text{Equation 8})$$

where  $d\mu$  is prior knowledge about  $\mathbf{J}$ . The solution under the MEM framework can be described as follows:

$$d\hat{p} = \arg \max_{dp \in \mathbf{Q}_J} S_\mu(dp). \quad (\text{Equation 9})$$

The estimated contrast among parcels can be regarded as the source imaging results.

### MSP

The MSP method (Friston et al., 2008a) is similar to MEM but under the Bayesian framework. It also assumes spatial priors on the source prior  $\mathbf{Q}_J$ . The source patches distributed in the source space comprise source components. The estimated source  $\hat{\mathbf{J}}$  can be considered a specific combination of these source components, denoted as  $\mathbf{D}$ . There are several combinations of source components. Bayesian model comparison was introduced to compare different combinations using evidence (Friston et al., 2008a). The model evidence can be approximated by the variational free energy  $F$  (Friston et al., 2007). Thus, the inverse solution in MSP can be expressed as follows:

$$\hat{\mathbf{D}} = \arg \max_{\mathbf{D}_i \in \mathbf{Q}_J} F_{\mathbf{D}_i}, \quad (\text{Equation 10})$$

where  $\hat{\mathbf{D}}$  is the set of patches with maximum free energy and can be regarded as the source imaging results.

### Implementations of source imaging methods

The implementations of the source imaging methods mentioned earlier are based on three widely used toolboxes: Brainstorm, FieldTrip, and SPM12. There are differences in the implementations of the same

method on different toolboxes, for example, differences in the aspects of data whitening, regularization, and estimation of the SNR of data.

### MRI

We acquired T1-weighted MRI scans of one subject using an MPRAGE sequence of a Siemens MAGNETOM Prisma 3T MR system (TR, 2,300 ms; TE, 3.03 ms; TI, 1,100 ms; FA, 8°; field of view, 256 × 256 × 192 mm; voxel size, 1 × 1 × 1 mm<sup>3</sup>). MRI data were preprocessed and segmented using FreeSurfer software (Fischl, 2012). The data obtained in FreeSurfer were then imported into Brainstorm to register the subject's T1 MRI to the subject coordinate system, and the scalp surface and cortex were automatically downsampled and reconstructed. In addition, the registered MRI was segmented by SPM12 to generate the head, innerskull, and outerskull surfaces which were used when solving the forward problem.

### OPM-MEG sensor configuration

Our evaluation was performed using the OPM-MEG sensor configuration. We designed and 3D-printed a helmet for the OPM-MEG system, which had 85 slots for inserting OPM sensors. The simulations were performed with all 85 slots. The sensors were modeled as single-axis sensors with their orientations radial to the head. We co-registered the helmet with the scalp surface extracted from MRI to obtain the relative positions and orientations between the OPM sensors and MRI. The average distance between the sensors and the scalp was  $7.7 \pm 4.1$  mm.

### Head model

We used a three-layer piecewise homogeneous head model and the boundary element method (BEM) to compute the lead field matrix (Mosher et al., 1999; Cao et al., 2022). The BEM computation was performed using OpenMEEG software (Gramfort et al., 2010). Although Brainstorm, FieldTrip, and SPM12 can all call the OpenMEEG program, there are some differences among the computed lead fields owing to the default settings of the parameters. To avoid discrepancies in source imaging results caused by different forward solutions, we used the same segmented head model and the corresponding lead field matrix computed in Brainstorm to solve the inverse problem in our study. Each BEM head model surface consisted of 4,322 vertices. The electrical conductivities of the scalp, skull, and brain were set to 0.33, 0.0042, and 0.33 S/m, respectively (Stenroos et al., 2014).

### Source and source model

We simulated two types of sources, patch source and dipole source, to evaluate the performance of MSI methods. The Lausanne parcellation provided in the Easy Lausanne software (Daducci et al., 2012) was used to subdivide the cortex into 1,000 nonoverlapping parcels covering the brain region, excluding the corpus callosum and the deep brain. Each segmented Lausanne parcel was regarded as the patch source, and the point closest to the center of each parcel was regarded as the dipole source. Therefore, there were 1,000 patch sources and 1,000 dipole sources. We simulated a single source in each dataset, and all sources were activated in turn to generate the datasets. The source signals were bandpass orthogonal Gaussian signals (10–30 Hz). The source was activated within 0–300 ms following 200 ms of baseline data. Note that the activities of the vertices in a patch source were highly coherent because they were generated by simulating identical time series. The cortex segmented from FreeSurfer was down-sampled to a mesh with 15,002 vertices and used as the distributed source model to solve the inverse problem. The orientation of each source was restricted to perpendicular to the local cortical surface.

### SNR settings of the simulated data

For the simulated data, Gaussian noise was added to each group of generated sensor data and the noise was independent across sensors. The intensity of the Gaussian noise was set to scale to obtain the SNR of the data at the desired level. SNR was defined as  $SNR = 20 \log_{10} [rms(Y_{signal}) / rms(Y_{noise})]$ .

### Parameter settings of MSI methods

In our study, most parameters followed the default or recommended settings of each method in each toolbox. The main settings are summarized in Table S1. Among them, the parameters related to regularization were set according to the estimated noise level (ENL) of the data to bring the regularization to a reasonable scale (Samuelsson et al., 2021). The noise levels of the simulated data and the real OPM-MEG data were all estimated by the ratio of the root-mean-square of the baseline data and

post-stimulus data, that is,  $ENL = rms(data_{baseline})/rms(data_{post-stimulus})$ . The source imaging results of all methods were thresholded at 20% of the maximum power value of the estimated sources.

### Evaluation metrics

#### Variational free energy

Under the Bayesian framework, the inverse problem is to maximize the posterior distribution of the source amplitudes, that is,  $p(\mathbf{J}|\mathbf{Y})$ . The source imaging results estimated by an MSI method can be regarded as a specific model. The Bayes rule gives an expression for the posterior distribution condition on the given model  $\mathbf{m}$ ,

$$p(\mathbf{J}|\mathbf{Y}, \mathbf{m}) = \frac{p(\mathbf{Y}|\mathbf{J}, \mathbf{m})P(\mathbf{J}|\mathbf{m})}{p(\mathbf{Y}|\mathbf{m})}, \quad (\text{Equation 11})$$

where  $p(\mathbf{Y}|\mathbf{m})$  represents the model evidence that can be used as a criterion to compare different models for model selection (Trujillo-Barreto et al., 2004; Friston et al., 2008b). The more reliable the model, the greater the value of the model evidence. Generally, it is difficult to give a direct expression of the model evidence. Fortunately, the variational free energy under Laplace approximation provides an approximation to the log evidence of the model (Friston et al., 2007), which can be expressed as follows (Friston et al., 2008a):

$$\ln(p(\mathbf{Y}|\mathbf{m})) \cong F = -\frac{n_t}{2} \text{tr}(\mathbf{Q}_Y \boldsymbol{\Sigma}^{-1}) - \frac{n_t}{2} \ln|\boldsymbol{\Sigma}| - \frac{n_t n_c}{2} \ln 2\pi + \frac{1}{2} \ln|\boldsymbol{\Sigma}_\lambda \boldsymbol{\Pi}^{-1}| - \frac{1}{2} (\boldsymbol{\mu}_\lambda - \boldsymbol{\eta})^T \boldsymbol{\Pi}^{-1} (\boldsymbol{\mu}_\lambda - \boldsymbol{\eta}) \quad (\text{Equation 12})$$

where  $\boldsymbol{\Sigma} = \mathbf{e}^{\lambda_m} \mathbf{L} \mathbf{Q}_J \mathbf{L}^T + e^{\lambda_\varepsilon} \mathbf{R}$  is the sensor-level covariance, and the hyperparameters  $\{\lambda|\lambda_m, \lambda_\varepsilon\}$  with the Gaussian hyperpriors  $N(\boldsymbol{\eta}, \boldsymbol{\Pi})$  and Gaussian posteriors  $N(\boldsymbol{\mu}_\lambda, \boldsymbol{\Sigma}_\lambda)$  can be estimated by maximizing free energy using the expectation maximization algorithm (Friston et al., 2008b; López et al., 2014).

Free energy reflects a trade-off between the simplicity of the model and its capability for data fitting (MacKay, 1992). In addition, the computation of model evidence does not require the ground truth of  $\mathbf{J}$ . Therefore, evaluating the source imaging results in practical applications where the real source distribution is unknown is more helpful. In this study, we introduced it to evaluate different source imaging results.

Each MSI method estimates a spatial distribution of the source power  $\hat{\mathbf{P}}$ , which can be directly obtained by calculating  $\hat{P}_i = \hat{\mathbf{J}}(i) \hat{\mathbf{J}}^T(i)$  or using pseudo-statistics such as in Equations (6) and (7). The source amplitude units vary among toolboxes. Therefore, to compare source estimation results in different toolboxes, the estimated source power was then normalized by its maximum value,  $P_{\max}$ . The normalized source power can be regarded as model  $\mathbf{m}$  to be compared. A simpler way to model the estimated source power as a source imaging model is to give more weight to the sources with larger power. The estimated source power is used as weight and encoded into the model:

$$\hat{\mathbf{Q}}_J = \text{diag}(P_1 / P_{\max}, P_2 / P_{\max}, \dots, P_{N_d} / P_{\max}). \quad (\text{Equation 13})$$

where the sources are assumed to be independent. Among the comparison of different MSI methods, (Equation 13) is substituted into (Equation 12) to calculate the corresponding free energy. Under the Bayesian framework, the estimated spatial distribution with maximum free energy was regarded as the best.

#### Validation metrics

**DLE.** DLE is defined as the Euclidean distance between the location of the maximum amplitude of the estimated source power map and the center of the true simulated source. A near-zero DLE value indicates a better localization accuracy.

**SD.** SD evaluates the degree of the spatial extent of the estimated source. SD is defined as the sum of the distances from the dipole with the maximum amplitude of the estimated source power to other dipoles in the source space. The distances are weighted by the estimated source power (Hedrich et al., 2017). SD can be expressed as follows:

$$SD = \frac{\sum_{i=1}^{n_d} \bar{p}_i d_i}{\sum_{i=1}^{n_d} \bar{p}_i}, \quad (\text{Equation 14})$$

where  $d_i$  is the distance between the location of the peak reconstructed power and the location of the  $i$ -th dipole in the source space.  $\bar{p}$  is the source power estimated by each method, normalized to a scale of [0,1], and  $\bar{p}_i$  is the power value corresponding to the  $i$ -th dipole.

**Resolution matrix.** The conventional resolution matrix  $\mathbb{R}$  is defined as  $\mathbb{R} = \mathbf{W}\mathbf{L}$ . However, this expression is only suitable for the linear inverse solver. Among the MSI methods to be evaluated, there are several nonlinear estimates. Therefore, we introduced an empirical resolution matrix  $\hat{\mathbb{R}}$  proposed in the study by Samuelsson et al. (2021) to quantify the performance of each source imaging method. The element of the  $i$ -th row and  $j$ -th column of the empirical resolution matrix can be expressed as follows:

$$\hat{\mathbb{R}}_{ij} = \frac{\sum_{k=1}^{n_{D_i}} \bar{p}_{(D_i),k}}{n_{D_i}}, \quad (\text{Equation 15})$$

where  $\bar{p}_{(D_i)}$  is the estimated source power normalized to 0–1 when the  $j$ -th source patch is activated, and  $\bar{p}_{(D_i),k}$  is its  $k$ -th power value and  $k \in D_i$ .  $n_{D_i}$  is the number of dipoles in patch  $D_i$ . In other words, the  $i$ -th row of  $\hat{\mathbb{R}}$  reflects the  $i$ -th source patch response to each patch or dipole activated. The  $j$ -th column of  $\hat{\mathbb{R}}$  reflects the  $j$ -th activated source patch or dipole motivation for each patch or dipole.