VSSI-GGD: A Variation Sparse EEG Source Imaging Approach Based on Generalized Gaussian Distribution

Ke Liu, Shu Peng, Chengzhi Liang, Zhuliang Yu, *Member, IEEE*, Bin Xiao, Guoyin Wang, *Senior Member, IEEE*, and Wei Wu, *Senior Member, IEEE*

Abstract—Electroencephalographic (EEG) source imaging (ESI) is a powerful method for studying brain functions and surgical resection of epileptic foci. However, accurately estimating the location and extent of brain sources remains challenging due to noise and background interference in EEG signals. To reconstruct extended brain sources, we propose a new ESI method called Variation Sparse Source Imaging based on Generalized Gaussian Distribution (VSSI-GGD). VSSI-GGD uses the generalized Gaussian prior as a sparse constraint on the spatial variation domain and embeds it into the Bayesian framework for source estimation. Using a variational technique, we approximate the intractable true posterior with a Gaussian density. Through convex analysis, the Bayesian inference problem is transformed entirely into a series of regularized L_{2p}-norm ($0 < p < 1$) optimization problems, which are efficiently solved with the ADMM algorithm. Imaging results of numerical simulations and human experimental dataset analysis reveal the superior performance of VSSI-GGD, which provides higher spatial resolution with clear boundaries compared to benchmark algorithms. VSSI-GGD can potentially serve as an effective and robust spatiotemporal EEG source imaging method. The source code of VSSI-GGD is available at https://github.com/Mashirops/VSSI-GGD.git.

Index Terms—EEG source imaging, generalized Gaussian distribution, variation sparsity, ADMM.

Ke Liu is with Chongqing Key Laboratory of Computational Intelligence and the School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China (e-mail: liuke@cqupt.edu.cn).

Shu Peng, Chengzhi Liang, and Bin Xiao are with the School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China (e-mail: S210201074@stu.cqupt.edu.cn; 2020212196@stu.cqupt.edu.cn; xiaobin@cqupt.edu.cn).

Zhuliang Yu is with the College of Automation Science and Engineering, South China University of Technology, Guangzhou 510641, China (e-mail: zlyu@scut.edu.cn).

Guoyin Wang is with Chongqing Key Laboratory of Computational Intelligence and the Key Laboratory of Cyberspace Big Data Intelligent Security of Ministry of Education, Chongqing University of Posts and Telecommunications, Chongqing 400065, China (e-mail: wanggy@cqupt.edu.cn).

Wei Wu is with the Shanghai Key Laboratory of Emotions and Affective Disorders, Shanghai Jiao Tong University School of Medicine, Shanghai 201600, China (e-mail: weiwuneuro@gmail.com).

I. INTRODUCTION

ELECTROENCEPHALOGRAPHY (EEG) is widely employed for brain imaging due to the millisecond-level time resolution [1], [2]. However, EEG has limited spatial resolution because of the volume conduct effect [3]. EEG source imaging (ESI) is widely recognized as an effective technique for improving the spatial resolution of scalp EEG signals. It is a technique that uses EEG signals and mathematical models of human head tissues to derive the image of intracranial neural activities, precisely reconstructing their locations and extents [3], [4]. Previous studies on electrocorticography (ECoG) and functional magnetic resonance imaging (fMRI) have shown that brain activity exhibits spatial correlations, indicating that EEG signals may originate from widespread active regions of the cortex [5]. In other words, it requires an activated brain area (usually no less than 6 cm^2) to generate a spike on the scalp [6]. Reconstructing the spatial extent of cortical brain activity, along with the location, is not only possible but also holds significant scientific and medical importance [7], such as stroke recovery [8], [9], as well as detection and surgical resection of epileptic foci [10], [11]. Additionally, ESI provides signals with a higher spatiotemporal resolution for brain-computer interfaces (BCI) and related applications [12], [13].

For ESI problems, the distributed current density (DCD) model is a widely employed source model [3], [14]. The DCD model divides the cortex into several triangular grids, where each triangle represents a source with a fixed location. Typically, the number of sources is around several thousand and much larger than the number of scalp electrodes (usually 32 to 256 electrodes) [14], [15]. Hence, ESI is a severely ill-posed inverse problem, where countless potential source configurations can fit the scalp measurements. It is similar to compressed sensing, as it relies on utilizing a limited quantity of observed data to reconstruct intricate signals. To obtain a

unique and accurate solution, appropriate constraints need to be applied to narrow the solution space.

Conventional ESI methods employ the L_2-norm regularize to constrain the potential sources. For example, the minimum norm estimation (MNE) applies L_2-norm regularization to select the source configuration with the minimum energy that matches the scalp measurement [14]. However, MNE is not sensitive to deep sources because the electric fields generated by superficial sources have less energy [3]. To address this biased estimation, weighted MNE (wMNE) proposes to weight the regularization term with the norms of each column of the lead-field matrix [16]. To obtain spatial coherence and smooth solutions, low-resolution electromagnetic tomography (LORETA) proposes to minimize the L_2-norm of the second-order spatial derivative of the sources [17]. Nonetheless, the imaging results of these L_2-norm based methods are often blurry and cover most of the cortical regions [3], [18]. To improve spatial resolution, many studies have employed sparse constraints such as the L_1-norm further to narrow the solution space [19], [20]. Other studies incorporate constraints as prior information into the Bayesian framework. One such method is sparse Bayesian learning (SBL) [21], [22]. Under the SBL framework, the study in [23] combines the hyperparameters of the source variance along with the full-structure multivariate Gaussian noise for joint estimation, leading to promising outcomes. Cai et al. [24] utilize SBL to estimate the model data covariance and reconstruct potential sources with adaptive beamformers, effectively mitigating the effects of correlated brain sources and yielding good results.

The sparse constraint methods can accurately localize the activated brain sources with high spatial resolution. However, due to the sparsity enforcement on the source space, these methods usually generate only one or a few points within the active areas, thereby missing the extent information of source activities [14], [25]. To remedy this issue, several studies applied sparse constraints in the transform domains, such as the variation domain [26], [27], [28], or a combination of multiple domains [29]. These constraints can provide more accurate estimates of source location and extent. Nonetheless, most of these methods apply the L_1-norm constraint in the spatial transform domains, which tends to overestimate the extent of the sources, especially for small ones [30]. To reduce this bias, the study in [31] has pointed out that non-convex constraints like L_p-norm ($0 < p < 1$) can acquire more accurate results in contrast to L_1-norm-based methods, especially when measurements are noisy and restricted isometry property (RIP) is violated, which is always the case for ESI. Studies have indicated that the L_p-norm is sparser than the L_1-norm, as the inclusion of p allows for more flexible balancing between measurement data and sparsity [32]. In addition, L_p-norm-based methods require fewer measurements to yield reliable results [33].

In the Bayesian probabilistic framework, the solution of the typical L_p-norm regularization methods is equivalent to the maximum a posterior (MAP) estimation with the generalized Gaussian prior. However, according to [34] and [35], this non-convex MAP reconstruction often encounters many shallow local optima. The posterior, which incorporates the generalized Gaussian prior (L_p-norm constraint with $0 < p < 1$), is asymmetric and exhibits multiple local modes. The MAP estimation is unable to account for the overall mass of the posterior and tends to become stuck in shallow optima. To address this issue, Bayesian integration can be employed to average out these shallow local optima and obtain the mean of the posterior distribution [35], [36].

Similar to the mixed L_{21}-norm [25], [37], this work used the L_{2p}-norm regularization to reconstruct the cortical extended sources. As a matrix form extension of L_p-norm, we define this regularization operator to impose the L_2-norm regularization in the temporal domain and the L_p-norm regularization in the spatial variation domain. By formulating the ESI problem within the Bayesian framework, we introduced a new EEG source imaging algorithm called variation sparse source imaging based on generalized Gaussian distribution (VSSI-GGD). Using a variational technique, VSSI-GGD approximates the intractable posterior density with a Gaussian distribution. Through convex analysis, the approximated Gaussian posterior distribution is obtained by solving a series of regularized L_{2p}-norm constraint problems. The alternating direction method of multipliers (ADMM) algorithm [38] is employed to solve each efficiently regularized L_{2p}-norm constraint problem. By combining the generalized Gaussian prior and variational Bayesian inference technique, we anticipate that the proposed method will yield more accurate reconstructions of cortical sources compared to existing techniques. This is validated by simulations using synthetic and experimental EEG data.

The rest of this paper is summarized as follows. Section II provides a detailed description of VSSI-GGD. Section III presents the simulation design, benchmark algorithms, and performance metrics used to validate the performance of our proposed method. Section IV presents the results obtained using both simulated and real EEG data, followed by a discussion in Section V and a brief conclusion in Section VI.

Matrices and vectors in this paper are denoted using uppercase and lowercase boldface letters, respectively. $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ represents a Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. f^* and f^{-1} are the conjugate and inverse function of f, respectively. $\boldsymbol{x}_i$ and $\boldsymbol{x}_{j,:}$ denote the ith column and the jth row of matrix $\boldsymbol{X}$, respectively.

II. METHODS

A. Background

According to the quasi-static approximation of Maxwell's equations, the relationship of EEG signals and current sources is described as [3], [14],

$$\boldsymbol{B} = \boldsymbol{L}\boldsymbol{S} + \boldsymbol{\varepsilon} \tag{1}$$

where $\boldsymbol{B} \in \mathbb{R}^{N_b \times N_t}$ is the EEG data collected from N_b electrodes at N_t time sampling points, and $\boldsymbol{S} \in \mathbb{R}^{N_s \times N_t}$ is the cortical source of N_s potential brain sources at N_t time points. $\boldsymbol{L} \in \mathbb{R}^{N_b \times N_s}$ is the lead-field matrix, which models how potential brain activities, $\boldsymbol{S}$, are related to scalp EEG recordings. $\boldsymbol{\varepsilon}$ denotes the measurement noise which is assumed to be drawn from $\mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})$ independently, and $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$ is the

covariance obtained from pre-stimulus data. For generality, spatial whitening is applied to both the lead-field matrix and EEG measurements based on the estimated Σ_{ε}, resulting in $\varepsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ [25].

Based on (1), ESI aims to estimate the appropriate potential source configuration S from the EEG measurement data B with a given lead-field matrix L. However, the number of suitable source configurations tends to approach infinity because the number of EEG electrodes is much less than the number of potential sources, i.e., $N_b \ll N_s$. Thus, prior spatiotemporal constraints are needed to narrow the solution space for the ill-posed problem. For the temporal constraint, we assume that the source matrix S is a linear combination of K temporal basis functions (TBFs) [15], [25], $\Phi = [\phi_1, \cdots, \phi_K]^\top \in \mathbb{R}^{K \times N_t}$, where $\phi_k \in \mathbb{R}^{N_t \times 1}$ is the kth TBF. By projecting EEG measurements B and sources S onto the subspace spanned by the TBFs, we can capture the temporal properties of brain sources through the projection coefficients $\tilde{B} = B\Phi^\top \in \mathbb{R}^{N_b \times K}$, and $\tilde{S} = S\Phi^\top \in \mathbb{R}^{N_s \times K}$. Hence, the data model in (1) can be written as

$$\tilde{B} = L\tilde{S} + \tilde{\varepsilon} \tag{2}$$

where $\tilde{\varepsilon} = \varepsilon\Phi^\top \in \mathbb{R}^{N_b \times K}$. Here, we construct TBFs from singular value decomposition (SVD) of EEG signals and use the Kaiser criterion to determine the value of K [25]. For orthogonal TBFs, we obtain $\tilde{\varepsilon}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ [15]. After obtaining $\tilde{S}$, we recover the potential source S as $S = \tilde{S}\Phi$. The use of TBFs not only reduces the dimension of the ESI problem but also helps stabilize the source reconstruction. For notation simplicity, we use S and B instead of $\tilde{S}$ and $\tilde{B}$ throughout the remainder of this paper.

For spatial constraint, to obtain locally smooth and globally clustered cortical sources [5], [15], we enforce source sparsity on the spatial variation domain and penalize the amplitude differences between adjacent dipoles. Specifically, we apply a linear transformation to the source through the variation operator V to obtain the variation source U. V is defined as [26], [39]

$$V = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1N_s} \\ v_{21} & v_{22} & \cdots & v_{2N_s} \\ \vdots & \vdots & \ddots & \vdots \\ v_{M1} & v_{M2} & \cdots & v_{MN_s} \end{bmatrix}$$
$$\begin{cases} v_{mi} = 1, v_{mj} = -1, i < j; & \text{if source } i, j \text{ share edge } m \\ v_{mi} = 0; & \text{otherwise} \end{cases}$$
$$\tag{3}$$

where M is the number of triangular edges in the source model. Each row of V corresponds to two adjacent dipoles that share the same edge in the triangular grid on the cerebral cortex, indicated by the 1 and -1 values. Hence, each row of variation sources $U = VS \in \mathbb{R}^{M \times K}$ represents the amplitude difference between two adjacent sources. The use of the variation transformation enforces sparsity on the variation source, preventing the loss of spatial information caused by enforcing sparsity in the original domain [7].

B. Inverse Solution

To reconstruct extended sources, we assume that the variation source U is sparse, and each row of U is independent of each other. We achieve this by incorporating the generalized Gaussian prior for each row of U [40]

$$p(\boldsymbol{u}_{i,:}) = \frac{p}{2\lambda\Gamma(1/p)}\exp\left[-\left(\frac{\|\boldsymbol{u}_{i,:}\|_2}{\lambda}\right)^p\right] \tag{4}$$

where λ and p are the scale and shape parameters of the generalized Gaussian distribution, and $\Gamma(\cdot)$ is the Gamma function. Then the prior of U can be expressed as

$$p(U) = \prod_{i=1}^{M} p(\boldsymbol{u}_{i,:}) \tag{5}$$

Based on the data model (1) and prior (5), the posterior distribution of S is derived as

$$p(S|B) = Z^{-1} p(B|S) p(U), \quad U = VS$$
$$Z = p(B) = \int p(B|S) p(U) dS \tag{6}$$

where Z is the model evidence. However, due the the non-Gaussian prior, the integration of Z is intractable. To efficiently leverage information from the posterior distribution, we utilize a variational technique to find a Gaussian density $Q(S|B)$ to fit the true posterior $p(S|B)$.

According to the convex analysis, the prior $p(\boldsymbol{u}_{i,:})$ can be approximated as [41]

$$p(\boldsymbol{u}_{i,:}) = \max_{\gamma_i} \mathcal{N}(\mathbf{0}, \gamma_i^{-1}\mathbf{I}) f(\gamma_i) \tag{7}$$

where $f(\gamma_i) = \sqrt{\frac{2\pi}{\gamma_i}}\exp\left(g^*\left(\frac{\gamma_i}{2}\right)\right)$, g^* is the concave conjugate of g, $g(x) = \log\lambda + \lambda^{-p}x^{p/2} + \mathcal{C}$, and $\mathcal{C}$ is a constant. Therefore, we have

$$p(U) = \prod_{i=1}^{M} p(\boldsymbol{u}_{i,:}) = \max_{\boldsymbol{\gamma}} \prod_{t=1}^{K} \mathcal{N}(\boldsymbol{u}_t|0, \boldsymbol{\Gamma}^{-1}) \prod_{i=1}^{M} f(\gamma_i) \tag{8}$$

in which $\boldsymbol{u}_t$ denotes the tth column of U and $\boldsymbol{\Gamma} = \text{diag}(\boldsymbol{\gamma})$, where $\boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_M]^\top$ is a vector comprising M unknown nonnegative hyperparameters that govern the prior covariance of U. If $\boldsymbol{\gamma}$ is known, $p(U)$ is lower bounded by $\mathcal{C}_1 \prod_{t=1}^{K} \mathcal{N}(\boldsymbol{u}_t|0, \boldsymbol{\Gamma}^{-1})$, where $\mathcal{C}_1 = \prod_{i=1}^{M} f(\gamma_i)$ is a constant. With this Gaussian-form lower bound, the approximated Gaussian posterior are obtained as

$$Q(S|B; \boldsymbol{\gamma}) \propto p(B|S) \prod_{t=1}^{K} \mathcal{N}(\boldsymbol{u}_t|0, \boldsymbol{\Gamma}^{-1}) \tag{9}$$

Given a value of $\boldsymbol{\gamma}$, the mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ of $Q(S|B; \boldsymbol{\gamma})$ are represented as follows

$$\boldsymbol{\mu} = \boldsymbol{\Sigma}L^\top B$$
$$\boldsymbol{\Sigma} = A^{-1}, \quad A = L^\top L + V^\top \boldsymbol{\Gamma} V \tag{10}$$

As a result, we aim to determine the optimal value of $\boldsymbol{\gamma}$ that makes Gaussian density $Q(S|B; \boldsymbol{\gamma})$ the tightest to $p(S|B)$. A natural way to fit $p(S|B)$ by $Q(S|B; \boldsymbol{\gamma})$ is to maximize the

model evidence Z in (6). Since $Q(S|B; \gamma) = \mathcal{N}(S|\mu, A^{-1})$, we have $\max_S Q(S|B) = |2\pi A^{-1}|^{-\frac{K}{2}} \int Q(S|B)dS$. Hence,

$$\max_S p(B|S)p(U) = |2\pi A^{-1}|^{-\frac{K}{2}} \int p(B|S)p(U)dS$$

$$\Rightarrow Z = \int p(B|S)p(U)dS = |2\pi A^{-1}|^{\frac{K}{2}} \max_S p(B|S)p(U)$$

(11)

The inference problem falls to maximize the right hand of Eq. (11) w.r.t. S and γ. Combining (8) and taking the negative logarithm of (11), we have

$$(\hat{\gamma}, \hat{S}) = \arg \min_{\gamma, S} \psi(\gamma, S)$$

$$= \arg \min_{\gamma, S, \lambda} \left[K \log |A| - 2 \sum_{i=1}^{M} g^*(\frac{\gamma_i}{2}) + R(\gamma, S) \right],$$

$$R(\gamma, S) = \|B - LS\|_{\mathcal{F}}^2 + \sum_{t=1}^{K} u_t^\top \Gamma u_t$$

(12)

For computationally convenience, we construct auxiliary functions using convex analysis to form an alternative objective function for $\log |A|$. Since it is concave in γ, by Fenchel duality, it can be expressed as a minimum over upper-bounding hyperplanes [21], [39]

$$\log |A| = \min_{\xi} \xi^\top \gamma - f^*(\xi)$$

(13)

where $f^*(\xi)$ is the concave conjugate of $\log |A|$ with $\xi = [\xi_1, \cdots, \xi_M]^\top$ as auxiliary variables [21], [25], [39], which is obtained from simple geometric considerations as the slope at the current γ of $\log |A|$

$$\xi = \nabla_\gamma \log |A| = \text{diag}(V A^{-1} V^\top)$$

(14)

In this work, we utilize the Lanczos algorithm to provide a scalable low-rank approximation of the covariance A^{-1} [42].

Combining (12) with (13) and removing the minimizations, the modified cost function is derived as

$$\psi(\gamma, S, \xi) = K \xi^\top \gamma - 2 \sum_{i=1}^{M} g^*(\frac{\gamma_i}{2}) + R(\gamma, S) - K f^*(\xi)$$

(15)

The hyperparameter γ is obtained by minimizing the cost function $\psi(\gamma, S, \xi)$ with other variables fixed

$$\gamma_i = \arg \min_{\gamma_i} K \xi_i \gamma_i + \gamma_i \|u_{i,:}\|_2^2 - 2g^*(\frac{\gamma_i}{2})$$

$$= \arg \min_{\gamma_i} 2 \left[\frac{\gamma_i}{2}(K \xi_i + \|u_{i,:}\|_2^2) - g^*(\frac{\gamma_i}{2}) \right]$$

(16)

Taking the derivation of γ_i, we have

$$K \xi_i + \|u_{i,:}\|_2^2 - \frac{g^*(\frac{\gamma_i}{2})}{\partial \gamma_i} = 0$$

$$\Rightarrow g^{*\prime}(\frac{\gamma_i}{2}) = K \xi_i + \|u_{i,:}\|_2^2$$

(17)

Note that $g(x)$ is concave, so the inverse function of its concave conjugate is itself [41]. Thus

$$\gamma_i = 2g^{*\prime-1}(K \xi_i + \|u_{i,:}\|_2^2) = 2g'(K \xi_i + \|u_{i,:}\|_2^2)$$

$$= p \lambda^{-p} x_i^{\frac{p-2}{2}} |_{x_i = K \xi_i + \|u_{i,:}\|_2^2}$$

(18)

Considering the terms related to S in $\psi(\gamma, S, \xi)$, the source S is updated as

$$S = \arg \min_S \|B - LS\|_{\mathcal{F}}^2 + \sum_{i=1}^{M} \left\{ \gamma_i [K \xi_i + \|u_{i,:}\|_2^2] - 2g^*(\frac{\gamma_i}{2}) \right\}$$

(19)

Since

$$g(x) = \min_\gamma \gamma^\top x - g^*(\gamma)$$

(20)

we have,

$$\gamma_i \left[K \xi_i + \|u_{i,:}\|_2^2 \right] - 2g^*(\frac{\gamma_i}{2}) = 2g(K \xi_i + \|u_{i,:}\|_2^2)$$

(21)

Hence, S is obtained by solving the following regularized optimization problem

$$S = \arg \min_S \|B - LS\|_{\mathcal{F}}^2 + 2\lambda^{-p} \sum_{i=1}^{M} \left(\sqrt{K \xi_i + \|u_{i,:}\|_2^2} \right)^p$$

$$= \arg \min_S \frac{1}{2} \|B - LS\|_{\mathcal{F}}^2 + \lambda^{-p} \sum_{i=1}^{M} \left(\sqrt{K \xi_i + \|u_{i,:}\|_2^2} \right)^p$$

(22)

Eq. (22) can be fully solved by ADMM algorithm w.r.t. the augmented Lagrangian function of (22)

$$\mathcal{L}(S, U, Z) = \frac{1}{2} \|B - LS\|_{\mathcal{F}}^2 + \lambda^{-p} \sum_{i=1}^{M} (K \xi_i + \|u_{i,:}\|_2^2)^{p/2}$$

$$+ \text{tr}[Z^\top (VS - U)] + \frac{\rho}{2} \|VS - U\|_{\mathcal{F}}^2$$

(23)

where $Z \in \mathbb{R}^{M \times K}$ is the Lagrangian multiplier and $\rho > 0$ is penalty parameter. S, U, and Z are updated alternately as

$$S^{k+1} = (L^\top L + \rho V^\top V)^{-1} \left[L^\top B + \rho V^\top \left(U^k - \frac{1}{\rho} Z^k \right) \right]$$

$$U^{k+1} = \arg \min_U \frac{\rho}{2} \|V S^{k+1} + \frac{1}{\rho} Z^k - U\|_{\mathcal{F}}^2 + f(U)$$

$$Z^{k+1} = Z^k + \rho(V S^{k+1} - U^{k+1})$$

(24)

where S^k, U^k, and Z^k are the corresponding output of the kth ADMM iteration, and $f(U) = \lambda^{-p} \sum_{i=1}^{M} (K \xi_i + \|u_{i,:}\|_2^2)^{p/2}$. Here, the optimization problem of U^{k+1} is solved using iteratively reweighted least squares (IRLS) algorithm [43]

$$U^{k+1} = \arg \min_U \frac{\rho}{2} \|Y - U\|_{\mathcal{F}}^2 + \text{tr}(U^\top W U)$$

$$\text{s.t.} \quad Y = V S^{k+1} + \frac{1}{\rho} Z^k$$

(25)

where $W = \text{diag}(w_1, w_2, \cdots, w_M)$, and $w_i = \lambda^{-p} \left(K \xi_i + \|u_{i,:}\|_2^2 \right)^{\frac{p-2}{2}}$.

Here, the scale parameter λ is chosen using cross-validation. The ADMM penalty parameter ρ is selected using adaptive residual balancing strategy [44], and the update for ρ is performed after every 10 ADMM iterations. To determine the value of p, according to prior studies [31], [32], [33]

TABLE I
LIST OF SYMBOLS IN THIS PAPER

Variable	Meaning	Variable	Meaning
B	EEG signal	N_b	Number of Electrodes
S	Source	N_s	Number of Sources
L	Lead-field Matrix	N_t	Number of sampling points
ϵ	Noise	Φ	Temporal Basis Functions
K	Number of TBFs	V	Variation Operator
U	Variation Source	M	Number of Triangular Edges
p	Shape Parameter	λ	Scale Parameter
μ	Mean	γ	Hyperparameters of U
Σ	Covariance	A	Inverse of Σ
ξ	Auxiliary Variables	ρ	Penalty Parameter
Z	Lagrangian multiplier	W	IRLS' Coefficient

and our numerical results provided in the supplementary materials, we experimentally set $p = 0.8$. The proposed algorithm was conducted on a standard PC (Corei9-10980XE CPU 3 GHz and 128GB RAM). During our numerical simulation, VSSI-GGD converges after ten cycles, where each cycle consists of 400 ADMM iterations which takes about 30 seconds. A schema of VSSI-GGD above is presented in Algorithm 1, and the code of it is offered at https://github.com/Mashirops/VSSI-GGD.git.

Algorithm 1 VSSI-GGD Algorithm

Input: EEG data $B \in \mathbb{R}^{N_b \times N_t}$, lead-field matrix $L \in \mathbb{R}^{N_b \times N_s}$, TBF matrix $\Phi \in \mathbb{R}^{K \times N_t}$, variation operator $V \in \mathbb{R}^{M \times N_s}$, hyperparameters λ, p.

Output: Potential source $S \in \mathbb{R}^{N_s \times N_t}$.

1: Project EEG data onto TBFs, i.e. $B = B\Phi^{\mathsf{T}}$.
2: Initialize $S = 0$, $U = 0$, $Z = 0$, $\xi = 1$, $\rho = 1$.
3: **repeat**
4:     **repeat**
5:         update S^k, U^k, Z^k with (24).
6:     **until** $\frac{\|S^k - S^{k-1}\|_{\mathcal{F}}}{\|S^k\|_{\mathcal{F}}}$ is smaller than a specific value, e.g. 1e-4.
7:     update ξ and γ using (14) and (17), respectively.
8: **until** relative change of the cost function (12) is smaller than a tolerance, e.g. 1e-4.
9: $S = S\Phi^{\mathsf{T}}$

III. SIMULATION PROTOCOL AND PERFORMANCE METRICS

This section provides a detailed description of the benchmark algorithms, Monte Carlo numerical simulations, and performance metrics used to validate the performance of the proposed method.

A. Benchmark ESI Algorithms

The performance of the proposed method, VSSI-GGD, was compared with the following algorithms: (1) wMNE [16], (2) LORETA [17], which are two widely-employed L_2-norm based methods with the regularization parameter learned by Bayesian Minimum-Norm [4], (3) SBL [21], (4) VSSI-MAP [26], [39], the temporal extension of VB-SCCD, which

reconstructs the sources by solving

$$S = \arg \min_S \frac{1}{2} \|B - LS\|_{\mathcal{F}}^2 + \sum_{i=1}^{M} \lambda \|u_{i,:}\|_2 \qquad (26)$$

(5) VSSI-CM [39], and (6) VSSI-L_{2p}, which solves the following formula

$$S = \arg \min_S \frac{1}{2} \|B - LS\|_{\mathcal{F}}^2 + \lambda \|U\|_{2,p}^p \qquad (27)$$

where we define

$$\|U\|_{2,p}^p = \sum_{i=1}^{M} \left(\sqrt{\sum_{j=1}^{N} u_{ij}^2} \right)^p = \sum_{i=1}^{M} (\|u_{i,:}\|_2)^p \qquad (28)$$

The regularization parameter λ for VSSI-MAP and VSSI-L_{2p} was derived by cross-validation on the EEG sensors [25].

B. Simulation Design

We conducted a series of Monte Carlo numerical simulations to verify the performance of these ESI algorithms due to the lack of ground truth. In this work, we utilized Brainstorm with the default ICBM 152 magnetic resonance imaging (MRI) to generate the head model [45]. The cortex surface was divided into 6006 triangular grids, with each triangle representing a potential brain source. The lead-field matrix was calculated based on the 64-channel Neuroscan Quik-cap sensor system and the OpenMEEG [46] (with two non-EEG channels abandoned). To construct a simulated cortical source, a seed point was randomly selected on the cortical triangular grids, and adjacent triangles were added to the cluster until the source region reached a specified value. The time series of the simulated ground truth, denoted as S_{real}, was divided into pre- and post-stimulus parts. The post-stimulus parts were simulated by applying several Gaussian-damped sinusoidal signals to the activated source region. Interference noise, labeled as S_{noise}, was added to mimic background source activities. Specifically, five clusters outside the real sources were randomly chosen as interference sources with an extent of 3-5 cm^2. Pink noise with $1/f$ power was applied to these interference sources. The final simulated source signals S were obtained by the weighted sum of normalized S_{real} and S_{noise} under a specific signal-to-noise/interference ratio (SNIR),

$$S = \frac{S_{real}}{\|S_{real}\|_{\mathcal{F}}} + \varphi \frac{S_{noise}}{\|S_{noise}\|_{\mathcal{F}}} \qquad (29)$$

Here, φ is the ratio of real brain source signals to interference noise, and SNIR is defined as $10 \lg \frac{1}{\varphi}$ [15]. Gaussian white noise was added as the measurement noise under a specified SNR defined as SNR $= 10 \lg \frac{\|LS\|_{\mathcal{F}}}{\|\varepsilon\|_{\mathcal{F}}}$. Consequently, the simulated observed EEG data, denoted as B, was obtained through the forward model.

The performance of ESI methods was compared under various scenarios

1) Various extents - we tested the influence of sources with different extents on ESI. Specifically, we considered source extents of 2 cm^2, 5 cm^2, 10 cm^2, 18 cm^2, and 32 cm^2 for a single source;

2) *Various SNRs* - similarly, we evaluated the performance of the proposed methods under different measurement noise levels (0, 3, 5, and 10 dB) with a single patch source of 8 cm^2;

3) *Various numbers of channels* - we investigated the effect of varying amounts of data on the algorithm results by considering different numbers of channels with SNR and SNIR set to 5 dB. Specifically, we down-sampled the EEG channels and analyzed cases with 100% (62 channels), 75% (46 channels), 50% (32 channels), and 25% (16 channels) of the data. In each simulation, channels were manually selected for non-complete cases (refer to the supplement document for detailed information).

In the aforementioned scenarios, unless specified otherwise, the SNR and SNIR were set to 5 dB with a single source of 8 cm^2. In the supplement document, we also present the results of different ESI methods with various SNIR levels. Each case involved 50 Monte Carlo simulations.

C. Performance Metrics

Four performance metrics are employed to evaluate the performance of ESI algorithms.

1) The area under the receiver operating characteristic (ROC) curve (AUC) [4], [47] evaluates the detection accuracy and sensitivity of the reconstructed sources.

2) Spatial dispersion (SD) [29], [48] measures the spatial blurring of the estimated sources compared with the ground truth.

3) The distance of localization error (DLE) [29], [48] measures the localization error of the reconstructed sources.

4) The shape error (SE) [25], [39] quantifies the relative squared error between the normalized reconstructed and simulated sources.

For the details of each performance metric, please refer to [4] and [39]. In general, higher AUC values and lower SD, DLE, and SE values indicate better performance of the ESI methods. To assess the statistical significance, we utilized the Kruskal-Wallis test. If the statistic from this test is significant, we then conducted Wilcoxon rank sum tests, which are adjusted by the false discovery rate (FDR), within VSSI-GGD against each benchmark algorithm. Moreover, we employed Otsu's threshold to visualize the imaging results as in [25].

IV. RESULTS

A. Results of Monte Carlo Simulations

1) Influence of Various Extents: We first evaluated the influence of different source extents (2, 5, 10, 18, and 32 cm^2) for ESI methods with a single patch. The performance metrics of the results are displayed in Fig. 1. For very small size sources (i.e., 2 cm^2), all methods except SBL show poor performance with lower AUC and higher SD, DLE, and SE values. SBL demonstrates exceptional performance, indicated by nearly one AUC and zero SD, DLE, and SE values. This is attributed to SBL over-enforcing sparsity, which always produces point estimations. As the source extents increase, the SD and DLE values of SBL consistently remain low, but the AUC values significantly decrease. On the contrary, LORETA
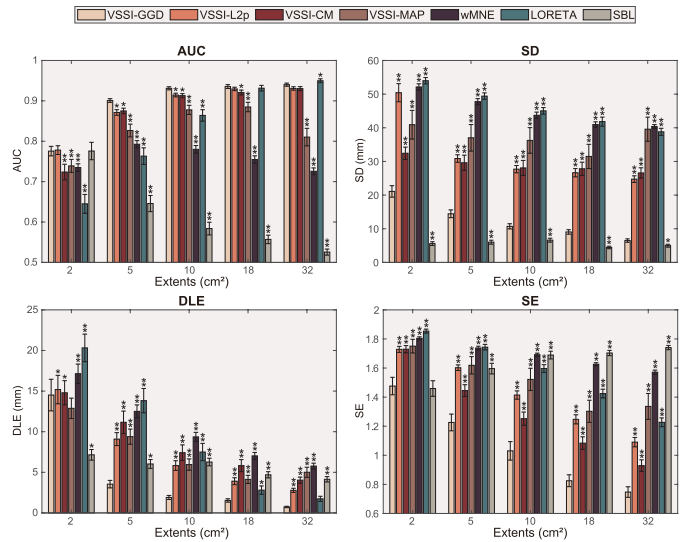

Fig. 1. Performance metrics of various extents. This figure presents the Mean $\pm$ SEM of the results for 50 Monte-Carlo simulations. * and ** indicate that the corresponding performance metric of VSSI-GGD is significantly different from those of the compared methods with $p_{\text{correct}} < 0.05$ and $p_{\text{correct}} < 0.01$ respectively.

often generates diffuse estimations, as evidenced by the high SD values. For source extents greater than 18 cm^2, LORETA shows the highest AUC values, indicating it is highly proficient at detecting large-scale sources, which aligns with the findings of [39] and [47]. In most cases, VSSI-GGD achieves larger AUC values and smaller SD (except that of SBL), DLE, and SE values than the compared methods. The results demonstrate the superior performance of VSSI-GGD in reconstructing the location and extent of extended sources.

One example of source imaging of different extents is presented in Fig. 2, in which all imaging results are thresholded by Otsu's method. The results of both wMNE and LORETA are diffused across all extents. While the results of SBL always shrink to several points, in the case of extended cortical activities, SBL loses most of the extent information. Four VSSI-based methods provide appropriate extents estimation, while VSSI-L_{2p}, VSSI-CM, and VSSI-MAP generate some spurious sources around the ground truth. Overall, the estimated results from VSSI-GGD are closer to the simulated sources in terms of localization and extent among seven ESI algorithms, which is also supported by the performance metrics provided in the supplementary document.

2) Influence of Various SNRs: We next assessed the performance of each ESI method against different SNR levels (0 dB, 3 dB, 5 dB, 10 dB), and the performance metrics are shown in Fig. 3. As SNR increases, all algorithms show improved performance, as evidenced by increased AUC and decreased SD, DLE, and SE values. Since SBL always produces focal estimations, the SD and SE values of SBL are not sensitive to noise. However, as found in the previous scenarios, SBL provides little information on source extents. VSSI-GGD outperforms the other benchmark methods under various measurement noise levels, indicated by higher AUC and lower SD and SE values.
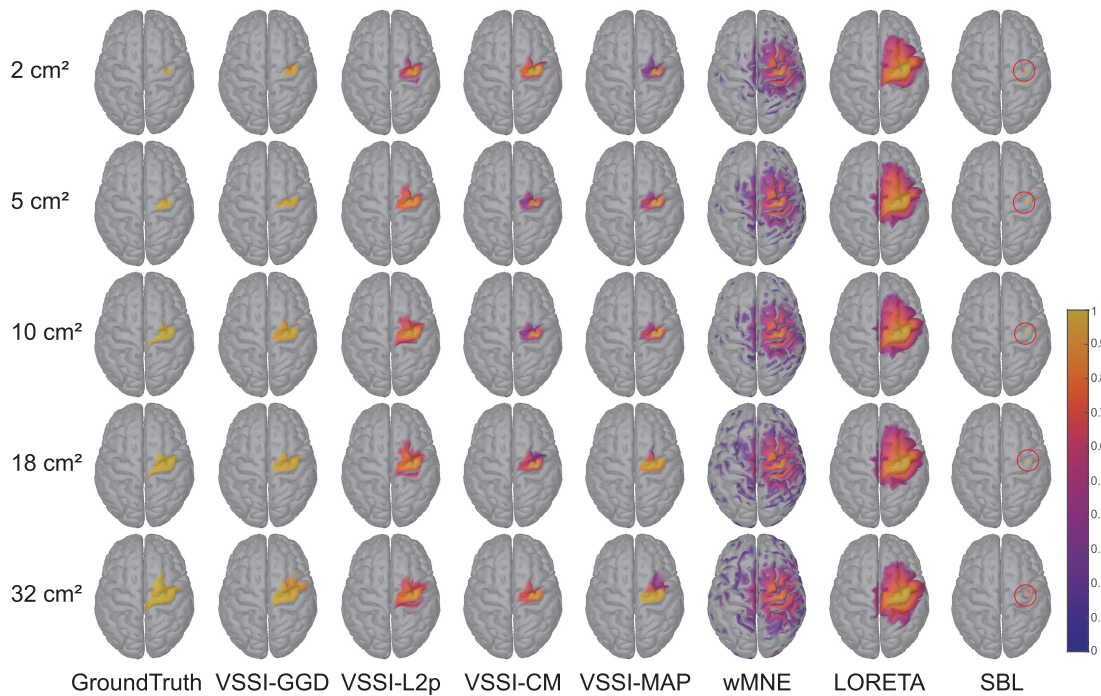
Fig. 2. One imaging example of the simulated and estimated sources by ESI algorithms under various extents, and the performance metrics of the results corresponding to this instance are listed below. The thresholds of each estimated source are determined by Otsu's method.
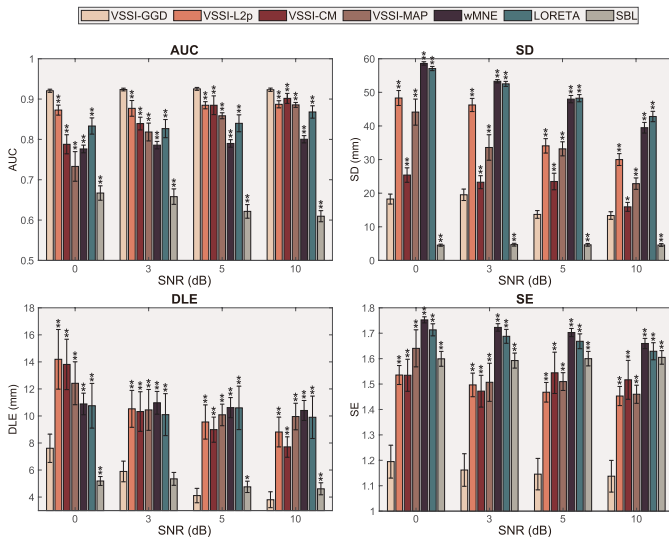

Fig. 3. Performance metrics of various SNRs. This figure shows the Mean $\pm$ SEM of the results for 50 Monte-Carlo simulations. * and ** indicate that the corresponding performance metric of VSSI-GGD is significantly different from those of the compared methods with $p_{correct} < 0.05$ and $p_{correct} < 0.01$ respectively.

3) Influence of Various Number of Channels: To evaluate the influence of data volume, we down-sampled the 62 EEG channels to 46, 32, and 16 channels. In the supplement document, we provide the electrode distribution map in 4 cases. Most methods show a declined performance as the number of channels decreases, indicated by the decreased AUC and increased SD, DLE, and SE values. However, the performance metrics of SBL remain almost unchanged. This stability can be attributed to its focal estimations. Except for the SD values of SBL,

VSSI-GGD achieves larger AUC, and lower SD, DLE, and SE values than the compared methods in all conditions, suggesting the superior performance and robustness of VSSI-GGD in the presence of missing data.

B. Results of Experimental EEG Data Analysis

We further apply these ESI algorithms on two real human EEG datasets to assess the practical efficacy of each method.

1) Analysis of Localize-MI Dataset: The Localize-MI dataset is an open dataset that includes high-density EEG data recorded from 61 sessions of 7 subjects with drug-resistant epilepsy to evaluate source location methods [49], which is available at. This dataset includes the spatial locations of stimulating contacts within the brain. These contacts were used to precisely deliver electrical currents during the presurgical evaluation of patients with drug-resistant epilepsy. Hence, Localize-MI provides the stimulation site as the ground truth. The EEG data were recorded at a sampling rate of 8000 Hz across 256 channels. We re-referenced the data to the average of all good channels and clipped the time window from -2 to 2 ms. The dataset also provides the surface model with 8196 sources and a corresponding lead-field matrix. We applied those ESI methods to the averaged EEG data of each session to reconstruct the potential sources and compare them with the ground truth. The MNE is used to visualize the results [50].

The source reconstruction results for the average time course of session 1 of subject 01 are depicted in Fig. 5. Imaging results of other subjects are provided in the supplementary material due to space limitations. In each imaging result, the green point indicates the ground truth. As shown in Fig. 5, the reconstructed source of VSSI-MAP misses the

| Subjects	VSSI-GGD	VSSI-L_{2p}	VSSI-CM	VSSI-MAP	SBL	wMNE	LORETA
sub-01	**2.26 ± 0.91**	4.21 ± 0.24**	6.12 ± 1.44**	5.47 ± 1.09**	3.30 ± 0.82**	4.10 ± 0.59**	3.94 ± 0.37**
sub-02	**4.74 ± 0.84**	5.86 ± 0.49	6.58 ± 0.96*	7.58 ± 1.02**	5.44 ± 1.39	5.67 ± 0.83	5.66 ± 0.61
sub-03	**3.42 ± 1.16**	4.96 ± 1.02	5.79 ± 0.33**	6.62 ± 1.94**	3.77 ± 1.19	4.91 ± 1.15	4.81 ± 1.24
sub-04	**2.80 ± 0.89**	5.59 ± 1.49**	6.19 ± 0.68**	6.54 ± 0.87**	5.68 ± 2.16**	5.83 ± 1.64**	5.37 ± 1.22**
sub-05	**3.13 ± 1.25**	4.97 ± 0.93**	7.79 ± 0.91**	8.02 ± 1.39**	3.41 ± 1.15	4.76 ± 0.68**	4.85 ± 0.67**
sub-06	**3.62 ± 1.15**	5.39 ± 0.87*	7.97 ± 0.60**	7.51 ± 1.21**	4.35 ± 0.94	5.31 ± 0.89*	5.42 ± 0.49*
sub-07	**4.95 ± 0.29**	6.25 ± 0.59**	8.55 ± 1.24**	8.54 ± 1.74**	4.96 ± 0.32	5.89 ± 0.43**	6.03 ± 0.70*
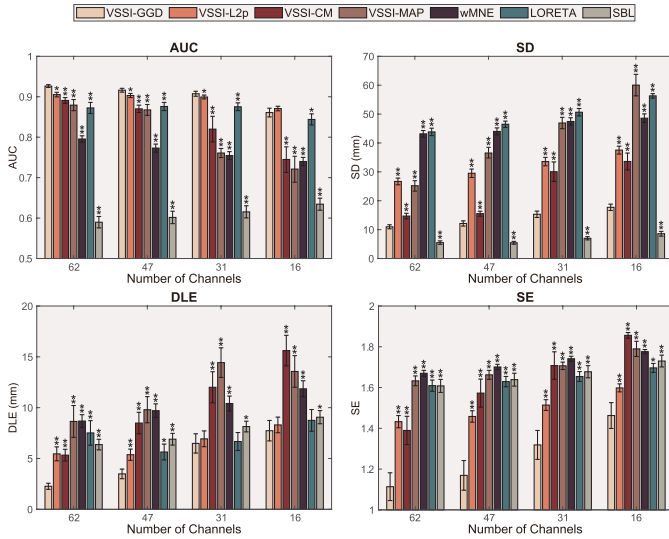
Fig. 4. Performance metrics of various number of channels. This figure shows the Mean ± SEM of the results for 50 Monte-Carlo simulations. * and ** indicate that the corresponding performance metric of VSSI-GGD is significantly different from those of the compared methods with $p_{correct} < 0.05$ and $p_{correct} < 0.01$ respectively.

ground truth point, whereas SBL estimates some point sources around the ground truth. Among the other six methods, VSSI-GGD provides more compact estimated sources with clear boundaries. To quantitatively assess the performance of each algorithm, we computed the spatial dispersion defined in [49] for all 61 sessions, and results are presented in Table II. VSSI-GGD provides the lowest SD values across all subjects in line with the imaging results.

2) Analysis of Public Epilepsy Dataset: We further applied the proposed method to the public epilepsy EEG dataset that belongs to the Brainstorm tutorial. The dataset is available at https://neuroimage.usc.edu/brainstorm/DatasetEpilepsy. The EEG data was recorded from a patient suffering from focal epilepsy and various types of seizures. Besides, this dataset was fully analyzed and reported in [27] and [51]. Here, we follow the Brainstorm tutorial to derive the head model, lead-field matrix, and averaged EEG data for ESI analysis. Source estimation results at the peak of averaged EEG data (0 ms) are presented in Fig. 6. All methods locate the source on the left frontal lobe, while wMNE and LORETA are widely spread, covering a large area. In contrast, SBL provides a focal estimation that consists of only one point source. Results of four VSSI-based methods are highly similar and consistent

with the clinical findings in [51] and the results in [27]. Among all, VSSI-GGD provides the clearest boundary between active and non-active sources.

V. DISCUSSIONS

We proposed an EEG extended source imaging algorithm, VSSI-GGD, within the Bayesian framework. VSSI-GGD employs the generalized Gaussian prior in the variation domain to enforce the sparsity of brain sources. Using variational inference, we obtained the approximated posterior by solving a series of regularized L_{2p}-norm regularization problems, which can be efficiently solved using the ADMM method. Both Monte Carlo numerical simulations and real EEG data analysis have revealed the superior performance of VSSI-GGD in estimating the location and extent of brain sources.

ESI plays a crucial role in stroke rehabilitation and resection of epileptic foci, as previous studies have confirmed the capability of EEG to capture potential neural activities [1], [7]. However, it is limited by its low spatial resolution due to the volume conduction effect. Furthermore, the number of potential sources is much larger than the number of EEG electrodes, making this problem inherently ill-posed. To improve the spatial resolution, appropriate constraints are necessary to narrow the source space.

The L_2-norm constraint is a widely used method in source estimation techniques such as wMNE and LORETA due to its computational convenience [3], [14], [30]. However, one drawback of this method is its tendency to produce diffuse estimations, as shown in Fig. 2, where the results of wMNE and LORETA cover a large portion of the cortical areas with low spatial resolution. Ideally, an activated potential brain source is expected to be a patch of moderate size with clear boundaries. Unfortunately, the information on source location and extent provided by wMNE and LORETA is often too blurred. To improve spatial resolution, some studies have employed sparse constraint methods such as L_1-norm and sparse Bayesian learning [19], [22], [24]. However, as depicted in Fig. 2, the results of these methods often collapse to a single point within the ground truth due to the direct imposition of sparse constraints in the original domain [7], [25], [39]. Several studies have shown that potential brain sources are sparse in certain transform domains, such as the variation domain, rather than the original domain [19], [25], [27], [39]. Therefore, enforcing sparsity in the original domain can lead to information loss.
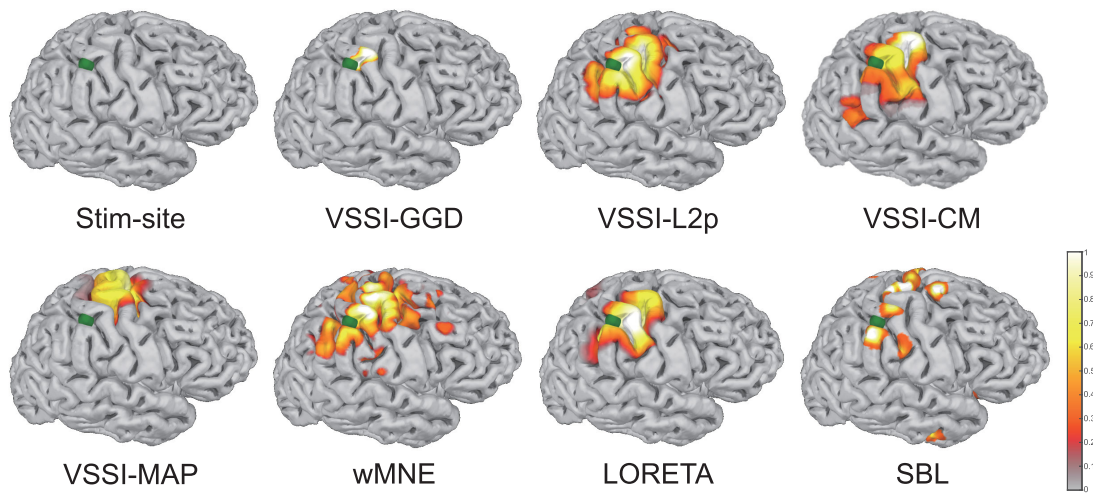
Fig. 5.   Source reconstruction results for the first session of subject 01 from the Localize-MI dataset. The green point is the sEEG simulation site.
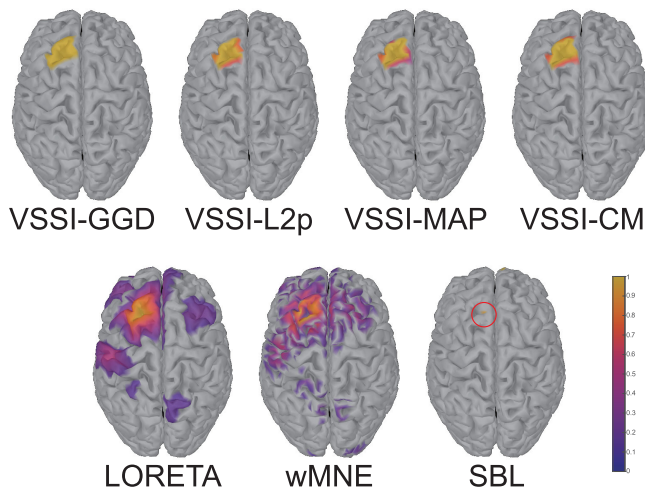

Fig. 6.   Reconstructed sources of the Brainstorm epilepsy tutorial data.

Keeping these factors in mind, in this work, we project the sources onto the variation domain to obtain sparse variation sources, following the strategy as prior studies [25], [26], [39]. Previous studies have demonstrated that variation sparsity-based methods are effective in providing clear information about source location and extent [7], [25], [26], [27], [39]. For VSSI-GGD, we assume that each row of the variation source follows the generalized Gaussian distribution. This statistical model is more flexible and suitable for non-Gaussian distribution variables like EEG signals, compared to the common Gaussian model [40]. The source sparsity is ensured by limiting the shape parameter p range from 0 to 1. The generalized Gaussian prior is then embedded into the Bayesian framework for statistical inference. It's important to note that our objective is not to directly maximize the posterior distribution, but rather to obtain the posterior's mean by maximizing the model evidence's lower bound. This is because the MAP estimation only gives the mode of $p(S|B)$ and is blind to the mass, especially for the multimodal asymmetric posterior distribution [39], [42].

The Bayesian framework provides a flexible framework where many ESI algorithms can be derived and explained.

There is a direct connection between many classical distributions and norms within this framework. For instance, when the prior is assumed to be a Gaussian distribution, the L_2-norm based method MNE corresponds to the MAP estimation of this prior [21]. Compared to regularization methods, the Bayesian framework is powerful because it allows for explicit modeling of prior assumptions using probability distributions and hyper-parameters in a fully data-driven manner. In the case of VSSI-GGD, the generalized Gaussian prior we use corresponds to the L_p-norm ($0 < p < 1$). The L_p-norm is also a sparse constraint similar to the L_1-norm. Both norms are approximations of the NP-hard L_0-norm. However, as suggested by [30], methods based on the L_1-norm on the variation domain tend to overestimate the extent of sources, particularly for small-sized sources. On the other hand, the more sparse mathematical nature of the L_p-norm allows it to better fit the measurement and source signals. This is also confirmed in the study of [31], where they compared these two sparse constraint-based methods and revealed that L_p-norm can mitigate the bias caused by L_1-norm.

VSSI-GGD is novel as we transform the generalized Gaussian prior source estimation problem into a regularized L_{2p}-norm optimization problem through convex analysis. The L_{2p}-norm combines the L_2-norm in the temporal domain and the L_p-norm in the spatial domain, similar to the L_{21}-norm. This allows for effective localization of active sources and improved estimation of their spatial extent compared to traditional single-norm methods. However, the L_{21}-norm also suffers from the mathematical properties of the L_1-norm we mentioned above. To overcome this limitation, the L_{2p}-norm employs the L_p-norm to constrain the spatial domain of the potential sources for better spatial sparsity. VSSI-GGD outperforms VSSI-CM in reconstructing extended sources, as suggested by the higher AUC values and lower SD, DLE, and SE values achieved by VSSI-GGD shown in Fig. 1. Similarly, this conclusion holds for VSSI-L_{2p} and VSSI-MAP.

In general, the proposed VSSI-GGD narrows the solution space through temporal and spatial constraints in order to obtain a unique solution like most state-of-the-art methods. Specifically, VSSI-GGD utilizes a better sparse generalized

Gaussian prior and performs variational inference within the Bayesian framework, transforming the intractable posterior estimation problem into a regularized optimization problem. Besides, efficient solutions can be achieved using existing methods such as ADMM and IRLS. The results of both simulation experiments and real data analysis show that it is precisely because of the sparser constraint and Bayesian variation inference that, for extended source reconstruction, VSSI-GGD outperforms other ESI methods in terms of accuracy in localization and precision in extent estimation. Nevertheless, there are certain areas where improvements can be made to this method. Firstly, similar to other optimization methods, the computation time of VSSI-GGD is not short. It takes several minutes to complete the source estimation using our method, which does not meet the requirement of real-time systems, where millisecond-level response is expected. Secondly, determining an appropriate value of the shape parameter (referred to as the value of p) of the generalized Gaussian prior is crucial for accurately fitting the measurements in our proposed ESI algorithm. Based on our experiments (result is provided in the supplementary material) and previous studies [31], [32], [33], we finally set $p = 0.8$. An advanced approach to handle this problem is to utilize the algorithm unrolling theory [52] to build a deep neural network (DNN) using the iteration framework of VSSI-GGD. Algorithm unrolling is a powerful tool to connect existing iterative algorithms with DNN, where all parameters in the original algorithms can be learned through end-to-end training. Our future work will combine the Bayesian update rules and the algorithm unrolling technique to derive a full data-driven interpretable neural network for ESI. Hence, both the value of p and λ can be determined in a fully data-driven way. Moreover, by changing the iterative methods into DNN form, we can derive the output at the millisecond level, which is crucial in real-time source imaging applications.

VI. CONCLUSION

This paper presents a novel algorithm for EEG extended source imaging within the Bayesian framework. Our method exhibits robustness against noise and interference from numerical simulation perspectives, delivering high spatial resolution results for reconstructing extended sources, which outperforms other benchmark algorithms. Moreover, the results of the human epilepsy dataset also suggest the potential of our method in real human EEG processing. In addition, in future work, our method can be further elucidated in neural networks through algorithm unrolling techniques to generate fully data-driven interpretable neural networks.

REFERENCES

[1] M. Seeber, L.-M. Cantonas, M. Hoevels, T. Sesia, V. Visser-Vandewalle, and C. M. Michel, "Subcortical electrophysiological activity is detectable with high-density EEG source imaging," *Nature Commun.*, vol. 10, no. 1, p. 753, Feb. 2019.

[2] O. Hauk, M. Stenroos, and M. S. Treder, "Towards an objective evaluation of EEG/MEG source estimation methods—The linear approach," *NeuroImage*, vol. 255, Jul. 2022, Art. no. 119177.

[3] B. He, A. Sohrabpour, E. Brown, and Z. Liu, "Electrophysiological source imaging: A noninvasive window to brain dynamics," *Annu. Rev. Biomed. Eng.*, vol. 20, no. 1, pp. 171–196, Jun. 2018.

[4] K. Liu, Z. L. Yu, W. Wu, Z. Gu, Y. Li, and S. Nagarajan, "Bayesian electromagnetic spatio-temporal imaging of extended sources with Markov random field and temporal basis expansion," *NeuroImage*, vol. 139, pp. 385–404, Oct. 2016.

[5] X. Bai, V. L. Towle, E. J. He, and B. He, "Evaluation of cortical current density imaging methods using intracranial electrocorticograms and functional MRI," *NeuroImage*, vol. 35, no. 2, pp. 598–608, Apr. 2007.

[6] J. X. Tao, A. Ray, S. Hawes-Ebersole, and J. S. Ebersole, "Intracranial EEG substrates of scalp EEG interictal spikes," *Epilepsia*, vol. 46, no. 5, pp. 669–676, May 2005.

[7] A. Sohrabpour and B. He, "Exploring the extent of source imaging: Recent advances in noninvasive electromagnetic brain imaging," *Current Opinion Biomed. Eng.*, vol. 18, Jun. 2021, Art. no. 100277.

[8] F. Vecchio et al., "Cortical connectivity from EEG data in acute stroke: A study via graph theory as a potential biomarker for functional recovery," *Int. J. Psychophysiol.*, vol. 146, pp. 133–138, Dec. 2019.

[9] R. Li, S. Li, J. Roh, C. Wang, and Y. Zhang, "Multimodal neuroimaging using concurrent EEG/fNIRS for poststroke recovery assessment: An exploratory study," *Neurorehabil. Neural Repair*, vol. 34, no. 12, pp. 1099–1110, Dec. 2020.

[10] J. Singh, J. S. Ebersole, and B. H. Brinkmann, "From theory to practical fundamentals of electroencephalographic source imaging in localizing the epileptogenic zone," *Epilepsia*, vol. 63, no. 10, pp. 2476–2490, Oct. 2022.

[11] M. Cao et al., "Virtual intracranial EEG signals reconstructed from MEG with potential for epilepsy surgery," *Nature Commun.*, vol. 13, no. 1, p. 994, Feb. 2022.

[12] T. Fang et al., "Decoding motor imagery tasks using ESI and hybrid feature CNN," *J. Neural Eng.*, vol. 19, no. 1, Feb. 2022, Art. no. 016022.

[13] S. Asadzadeh, T. Y. Rezaii, S. Beheshti, and S. Meshgini, "Accurate emotion recognition utilizing extracted EEG sources as graph neural network nodes," *Cognit. Comput.*, vol. 15, no. 1, pp. 176–189, Jan. 2023.

[14] B. He, L. Ding, and A. Sohrabpour, "Neural engineering," in *Electrophysiological Mapping and Source Imaging*. Cham, Switzerland: Springer, Sep. 2020, pp. 379–413, doi: 10.1007/978-3-030-43395-6_13.

[15] K. Liu, Z. Wang, Z. Yu, B. Xiao, H. Yu, and W. Wu, "WRA-MTSI: A robust extended source imaging algorithm based on multi-trial EEG," *IEEE Trans. Biomed. Eng.*, vol. 70, no. 10, pp. 2809–2821, Oct. 2023.

[16] F. Lucka, S. Pursiainen, M. Burger, and C. H. Wolters, "Hierarchical Bayesian inference for the EEG inverse problem using realistic FE head models: Depth localization and source separation for focal primary currents," *NeuroImage*, vol. 61, no. 4, pp. 1364–1382, Jul. 2012.

[17] S. Dattola, F. C. Morabito, N. Mammone, and F. La Foresta, "Findings about LORETA applied to high-density EEG—A review," *Electronics*, vol. 9, no. 4, p. 660, Apr. 2020.

[18] R. Sun, A. Sohrabpour, G. A. Worrell, and B. He, "Deep neural networks constrained by neural mass models improve electrophysiological source imaging of spatiotemporal brain dynamics," *Proc. Nat. Acad. Sci. USA*, vol. 119, no. 31, Aug. 2022, Art. no. e2201128119.

[19] L. Ding and B. He, "Sparse source imaging in electroencephalography with accurate field modeling," *Hum. Brain Mapping*, vol. 29, no. 9, pp. 1053–1067, Sep. 2008.

[20] F. Liu, L. Wang, Y. Lou, R.-C. Li, and P. L. Purdon, "Probabilistic structure learning for EEG/MEG source imaging with hierarchical graph priors," *IEEE Trans. Med. Imag.*, vol. 40, no. 1, pp. 321–334, Jan. 2021.

[21] D. Wipf and S. Nagarajan, "A unified Bayesian framework for MEG/EEG source imaging," *NeuroImage*, vol. 44, no. 3, pp. 947–966, Feb. 2009.

[22] A. Hashemi, C. Cai, G. Kutyniok, K.-R. Müller, S. S. Nagarajan, and S. Haufe, "Unification of sparse Bayesian learning algorithms for electromagnetic brain imaging with the majorization minimization framework," *NeuroImage*, vol. 239, Oct. 2021, Art. no. 118309.

[23] A. Hashemi et al., "Joint learning of full-structure noise in hierarchical Bayesian regression models," *IEEE Trans. Med. Imag.*, vol. 43, no. 2, pp. 610–624, Feb. 2024.

[24] C. Cai et al., "Bayesian adaptive beamformer for robust electromagnetic brain imaging of correlated sources in high spatial resolution," *IEEE Trans. Med. Imag.*, vol. 42, no. 9, pp. 2502–2512, Sep. 2023.

[25] K. Liu, Z. L. Yu, W. Wu, Z. Gu, and Y. Li, "Imaging brain extended sources from EEG/MEG based on variation sparsity using automatic relevance determination," *Neurocomputing*, vol. 389, pp. 132–145, May 2020.

[26] L. Ding, "Reconstructing cortical current density by exploring sparseness in the transform domain," *Phys. Med. Biol.*, vol. 54, no. 9, pp. 2683–2697, May 2009.

[27] A. Sohrabpour, Y. Lu, G. Worrell, and B. He, "Imaging brain source extent from EEG/MEG by means of an iteratively reweighted edge sparsity minimization (IRES) strategy," *NeuroImage*, vol. 142, pp. 27–42, Nov. 2016.

[28] H. Becker et al., "SISSY: An efficient and automatic algorithm for the analysis of EEG sources based on structured sparsity," *NeuroImage*, vol. 157, pp. 157–172, Aug. 2017.

[29] M. Zhu, W. Zhang, D. L. Dickens, and L. Ding, "Reconstructing spatially extended brain sources via enforcing multiple transform sparseness," *NeuroImage*, vol. 86, pp. 280–293, Feb. 2014.

[30] H. Becker, L. Albera, P. Comon, R. Gribonval, F. Wendling, and I. Merlet, "Brain-source imaging: From sparse to tensor models," *IEEE Signal Process. Mag.*, vol. 32, no. 6, pp. 100–112, Nov. 2015.

[31] H. Janati, T. Bazeille, B. Thirion, M. Cuturi, and A. Gramfort, "Multi-subject MEG/EEG source imaging with sparse multi-task regression," *NeuroImage*, vol. 220, Oct. 2020, Art. no. 116847.

[32] F. Wen, L. Pei, Y. Yang, W. Yu, and P. Liu, "Efficient and robust recovery of sparse signal and image using generalized nonconvex regularization," *IEEE Trans. Comput. Imag.*, vol. 3, no. 4, pp. 566–579, Dec. 2017.

[33] R. Chartrand and V. Staneva, "Restricted isometry properties and nonconvex compressive sensing," *Inverse Problems*, vol. 24, no. 3, Jun. 2008, Art. no. 035020.

[34] M. W. Seeger and D. P. Wipf, "Variational Bayesian inference techniques," *IEEE Signal Process. Mag.*, vol. 27, no. 6, pp. 81–91, Nov. 2010.

[35] D. P. Wipf, B. D. Rao, and S. Nagarajan, "Latent variable Bayesian models for promoting sparsity," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 6236–6255, Sep. 2011.

[36] D. P. Wipf, J. P. Owen, H. T. Attias, K. Sekihara, and S. S. Nagarajan, "Robust Bayesian estimation of the location, orientation, and time course of multiple correlated neural sources using MEG," *NeuroImage*, vol. 49, no. 1, pp. 641–655, Jan. 2010.

[37] W. Ou, M. S. Hämäläinen, and P. Golland, "A distributed spatio-temporal EEG/MEG inverse solver," *NeuroImage*, vol. 44, no. 3, pp. 932–946, Feb. 2009.

[38] Y. Wang, W. Yin, and J. Zeng, "Global convergence of ADMM in nonconvex nonsmooth optimization," *J. Sci. Comput.*, vol. 78, pp. 29–63, Jan. 2019.

[39] K. Liu, Z. L. Yu, W. Wu, Z. Gu, Y. Li, and S. Nagarajan, "Variation sparse source imaging based on conditional mean for electromagnetic extended sources," *Neurocomputing*, vol. 313, pp. 96–110, Nov. 2018.

[40] Y. Bazi, L. Bruzzone, and F. Melgani, "Image thresholding based on the EM algorithm and the generalized Gaussian distribution," *Pattern Recognit.*, vol. 40, no. 2, pp. 619–634, Feb. 2007.

[41] J. Palmer, K. Kreutz-Delgado, B. Rao, and D. Wipf, "Variational EM algorithms for non-Gaussian latent variable models," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 18, 2005, pp. 1–8.

[42] M. W. Seeger and H. Nickisch, "Large scale Bayesian inference and experimental design for sparse linear models," *SIAM J. Imag. Sci.*, vol. 4, no. 1, pp. 166–199, Jan. 2011.

[43] I. Daubechies, R. DeVore, M. Fornasier, and C. S. Güntürk, "Iteratively reweighted least squares minimization for sparse recovery," *Commun. Pure Appl. Math.*, vol. 63, no. 1, pp. 1–38, Jan. 2010.

[44] B. Wohlberg, "ADMM penalty parameter selection by residual balancing," 2017, *arXiv:1704.06209*.

[45] F. Tadel, S. Baillet, J. C. Mosher, D. Pantazis, and R. M. Leahy, "Brainstorm: A user-friendly application for MEG/EEG analysis," *Comput. Intell. Neurosci.*, vol. 2011, pp. 1–13, Apr. 2011.

[46] A. Gramfort, T. Papadopoulo, E. Olivi, and M. Clerc, "OpenMEEG: Opensource software for quasistatic bioelectromagnetics," *Biomed. Eng. Online*, vol. 9, no. 1, pp. 1–20, 2010.

[47] C. Grova, J. Daunizeau, J.-M. Lina, C. G. Bénar, H. Benali, and J. Gotman, "Evaluation of EEG localization methods using realistic simulations of interictal spikes," *NeuroImage*, vol. 29, no. 3, pp. 734–753, Feb. 2006.

[48] W.-T. Chang, A. Nummenmaa, J.-C. Hsieh, and F.-H. Lin, "Spatially sparse source cluster modeling by compressive neuromagnetic tomography," *NeuroImage*, vol. 53, no. 1, pp. 146–160, Oct. 2010.

[49] E. Mikulan et al., "Simultaneous human intracerebral stimulation and HD-EEG, ground-truth for source localization methods," *Scientific Data*, vol. 7, no. 1, p. 127, Apr. 2020.

[50] A. Gramfort et al., "MEG and EEG data analysis with MNE-Python," *Frontiers Neurosci.*, vol. 7, p. 267, Mar. 2013.

[51] M. Dümpelmann, T. Ball, and A. Schulze-Bonhage, "SLORETA allows reliable distributed source reconstruction based on subdural strip and grid recordings," *Hum. Brain Mapping*, vol. 33, no. 5, pp. 1172–1188, May 2012.

[52] V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," *IEEE Signal Process. Mag.*, vol. 38, no. 2, pp. 18–44, Mar. 2021.