CANDLE: CORTICAL NULL-SPACE DECOMPOSITION FOR NONINVASIVE BRAIN SOURCE IMAGING

Anonymous authors
Paper under double-blind review

ABSTRACT

Electrophysiological source imaging (ESI) aims to estimate cortical source activity from noninvasive electrophysiological measurements such as electroencephalogram (EEG). However, ESI is fundamentally ill-posed because source activity is substantially higher-dimensional than sensor observations, resulting in non-unique solutions. Recent learning-based approaches address this ambiguity by learning data-driven source priors, yet they often struggle to generalize across subject-specific cortical geometries. To address this, we propose CANDLE, a learning-based ESI model that estimates source activity on subject-specific cortical geometries. CANDLE learns a prior over the null space induced by the source-to-sensor mapping derived from T1-weighted MRI, restricting learning to unobservable source components while preserving geometric constraints. To train CANDLE, we develop a whole-brain simulator spanning over 1,100 subject-specific cortical geometries with source configurations derived from over 26,000 statistical brain maps. Trained exclusively on simulated data, CANDLE outperformed prior ESI methods on simulated source activity estimation and generalized to two empirical tasks: (i) intracranial stimulation localization from simultaneously recorded scalp EEG and (ii) epileptogenic zone estimation from presurgical interictal EEG. Our project page is available at https://candle-esi.pages.dev.

1 INTRODUCTION

Monitoring whole-brain neural activity at both high temporal and spatial resolution is central to a broad range of applications, from the diagnosis of neurological disorders (Cao et al., 2022) to the development of brain-computer interfaces (BCIs) (Suzuki et al., 2026; Jayalath & Jones, 2026). This motivates the development of neuroimaging techniques capable of faithfully capturing such neural dynamics. Functional magnetic resonance imaging (fMRI) provides high spatial but limited temporal resolution and measures neural activity indirectly, whereas invasive electrophysiological techniques such as electrocorticography (ECoG) and stereoelectroencephalography (SEEG) provide high-fidelity recordings at fine temporal scales but require neurosurgical implantation and offer spatially restricted coverage (He & Liu, 2008; Momi et al., 2025). Noninvasive electrophysiological modalities such as electroencephalography (EEG) and magnetoencephalography (MEG) therefore constitute a promising avenue for whole-brain monitoring at high temporal resolution (Feng et al., 2025). Nevertheless, EEG/MEG exhibit limited spatial resolution at the sensor level, owing to relatively sparse sensor sampling and substantial spatial mixing of neural signals (i.e., volume conduction and field spread) (Wens, 2023).

Electrophysiological source imaging (ESI), which seeks to recover cortical source activity from EEG/MEG signals, has been extensively studied as a means of overcoming this limitation (Sun et al., 2022; Wang et al., 2026). By reconstructing neural activity in cortical source space, ESI holds the promise of whole-brain monitoring with both high temporal and spatial resolution. The utility of ESI has further been suggested for applications such as presurgical localization of epileptogenic zones (Corona et al., 2023) and improved decoding in BCIs (Gideoni et al., 2024).

ESI, however, is a fundamentally ill-posed problem: the source activity to be estimated is substantially higher-dimensional than the observed EEG/MEG, and the solution is consequently non-unique. Conventional ESI methods (Van Veen et al., 1997; Dale et al., 2000) therefore typically require careful tuning of regularization schemes and source-space constraints for individual applications. This reliance on manual constraints not only limits the broad applicability of ESI but also

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

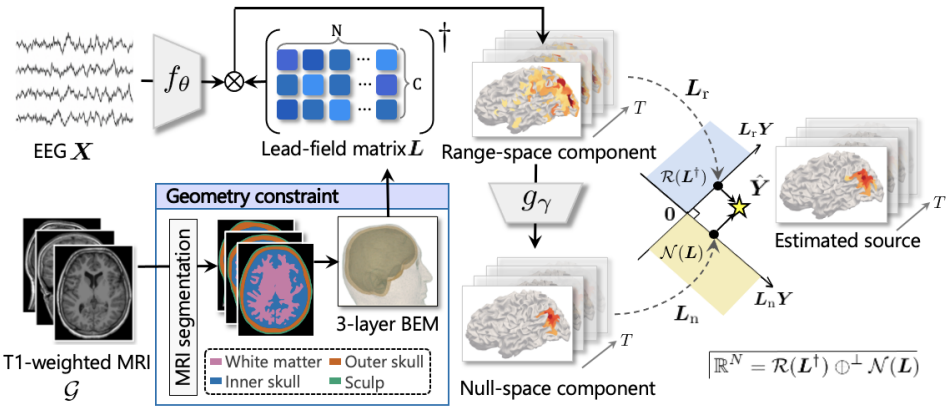

Figure 1: Overview of CANDLE. Given EEG $\boldsymbol{X}$ and the corresponding T1-weighted MRI $\mathcal{G}$, a subject-specific lead-field matrix $\boldsymbol{L}$ is derived as an anatomical constraint. The denoising model f_θ outputs denoised $\boldsymbol{X}$, from which the range-space component is analytically recovered, while g_γ estimates the unobservable null-space component. Their combination yields the source estimate $\hat{\boldsymbol{Y}}$.

undermines reproducibility. To alleviate these limitations, learning-based ESI models (Hecker et al., 2021; Sun et al., 2022) have recently been proposed to learn data-driven priors over source activity.

Despite this progress, most learning-based ESI models are trained in a subject-aggregated common cortical space, such as fsaverage (Fischl et al., 1999), and therefore cannot explicitly account for inter-individual differences in cortical geometry. Indeed, Sun et al. (2023) reported that applying learning-based ESI models to a particular subject requires subject-specific post-training.

To address this limitation, we propose CANDLE[1], a learning-based ESI model that directly estimates source activity on subject-specific cortical geometries from EEG. CANDLE learns the null-space component (Schwab et al., 2019) of source activity induced by a T1-derived source-to-sensor mapping. This allows the model to capture a prior over source activity while preserving subject-specific geometric constraints. Furthermore, we introduce a whole-brain simulator tailored to training CANDLE. Because ground-truth source activity is unobservable in real recordings, learning-based ESI models are typically trained on data synthesized by whole-brain simulators (Schirner et al., 2022; Ziaeemehr et al., 2025), and evaluated through sim-to-real validation on downstream tasks. Existing simulators, however, are not designed to jointly capture the anatomical variability and realistic source configurations required for training CANDLE. Our simulator addresses this gap by generating neurobiologically grounded source activity across diverse subject-specific cortical geometries.

Our contributions are summarized as follows:

- We propose CANDLE, a learning-based ESI model that incorporates subject-specific cortical geometry derived from each subject's T1-weighted MRI. CANDLE learns the null-space component of cortical source activity, thereby imposing a data-driven prior on the unobservable source subspace while preserving subject-specific geometric constraints.
- We develop a whole-brain simulator for training CANDLE that synthesizes realistic source activity across subject-specific cortical geometries from over 1,100 individuals and defines neurobiologically grounded source regions by clustering over 26,000 statistical brain maps.
- Without subject-specific tuning, CANDLE outperformed prior ESI methods on simulated source activity estimation and generalized to two relevant empirical tasks: (i) intracranial stimulation localization from simultaneously recorded scalp EEG and (ii) epileptogenic zone estimation from presurgical interictal EEG.

2 Methodology

We propose CANDLE, a learning-based electrophysiological source imaging (ESI) model that explicitly accounts for subject-specific cortical geometry. Figure 1 provides an overview of CANDLE.

[1]CANDLE: **C**ortical-geometry **A**ware **N**ull-space **D**ecomposition for **L**ocalized **ESI**

Given EEG $\boldsymbol{X} \in \mathbb{R}^{C\times T}$ and the corresponding T1-weighted MRI $\mathcal{G}$, the model estimates the underlying cortical source activity $\hat{\boldsymbol{Y}} \in \mathbb{R}^{N\times T}$. Here, C, N, and T denote the number of electrodes, the number of cortical source regions, and the sequence length, respectively.

Furthermore, to train CANDLE, we construct a whole-brain simulator illustrated in Figure 2. The simulator integrates subject-specific cortical geometries derived from more than 1,100 individuals with source configurations via clustering more than 26,000 statistical brain maps. In the following, Subsection 2.1 describes CANDLE, and Subsection 2.2 describes the construction of the simulator.

2.1 Proposed Method: CANDLE

Subject-specific cortical geometry constraint. In CANDLE, we derive a lead-field matrix from $\mathcal{G}$ that maps ground-truth cortical source activity $\boldsymbol{Y} \in \mathbb{R}^{N\times T}$ to $\boldsymbol{X}$ and serves as a subject-specific anatomical constraint. EEG arises from synchronized postsynaptic potentials of cortical pyramidal neuron populations and propagate through intervening tissues (e.g., cerebrospinal fluid) before being measured at the scalp (Hallez et al., 2007). This generating process is formulated as $\boldsymbol{X} = \boldsymbol{L}\boldsymbol{Y} + \boldsymbol{\epsilon}$, where $\boldsymbol{L} \in \mathbb{R}^{C\times N}$ denotes the lead-field matrix and $\boldsymbol{\epsilon}$ represents measurement noise.

Ideally, ESI could recover $\boldsymbol{Y}$ from $\boldsymbol{X}$ by inverting the operator $\boldsymbol{L}$; however, this inverse problem is inherently ill-posed because the cortical source space is typically much higher-dimensional than the EEG sensor space (i.e., $N \gg C$). Numerous source configurations can result in the same EEG observation. As described below, we address this by learning a data-driven prior over the distribution of $\boldsymbol{Y}$ within the source space decomposed according to $\boldsymbol{L}$.

Guided by prior neurophysiological insights into EEG generation (Tao et al., 2005; Birot et al., 2014), we derive $\boldsymbol{L}$ from $\mathcal{G}$ based on cortical regions defined by the Hagmann/Lausanne parcellation (Hagmann et al., 2008) and a three-layer boundary element model (Hallez et al., 2007) comprising the inner-skull, outer-skull, and scalp surfaces. Further details are provided in Appendix B.1.

Cortical source activity estimation. CANDLE estimates $\hat{\boldsymbol{Y}}$ while preserving the subject-specific geometric constraint by learning only the subspace of the source space that is unobservable from $\boldsymbol{X}$. We begin by introducing the range-null space decomposition (Schwab et al., 2019) induced by the linear operator $\boldsymbol{L}$. Let $\boldsymbol{L}^\dagger$ denote the pseudoinverse of $\boldsymbol{L}$, satisfying $\boldsymbol{L}\boldsymbol{L}^\dagger\boldsymbol{L} = \boldsymbol{L}$. We define two projection operators, $\boldsymbol{L}_\mathrm{r} \triangleq \boldsymbol{L}^\dagger\boldsymbol{L}$ and $\boldsymbol{L}_\mathrm{n} \triangleq \boldsymbol{I} - \boldsymbol{L}^\dagger\boldsymbol{L}$, which project a source-space sample $\boldsymbol{Y}$ onto the range space $\mathcal{R}(\boldsymbol{L}^\dagger)$ and the null space $\mathcal{N}(\boldsymbol{L})$, respectively. Since $\mathbb{R}^N = \mathcal{R}(\boldsymbol{L}^\dagger) \oplus^\perp \mathcal{N}(\boldsymbol{L})$, any source activity $\boldsymbol{Y}$ can be uniquely decomposed as

$$\boldsymbol{Y} = \boldsymbol{L}_\mathrm{r}\boldsymbol{Y} + \boldsymbol{L}_\mathrm{n}\boldsymbol{Y}. \tag{1}$$

Notably, since $\boldsymbol{L}\boldsymbol{Y} = \boldsymbol{L}\boldsymbol{L}_\mathrm{r}\boldsymbol{Y} + \boldsymbol{L}\boldsymbol{L}_\mathrm{n}\boldsymbol{Y} = \boldsymbol{L}\boldsymbol{Y} + \boldsymbol{0}$, these two components have fundamentally different observability: $\boldsymbol{L}_\mathrm{r}\boldsymbol{Y}$ can be interpreted as the component of $\boldsymbol{Y}$ observable in $\boldsymbol{X}$ through $\boldsymbol{L}$, whereas $\boldsymbol{L}_\mathrm{n}\boldsymbol{Y}$ is completely unobservable from $\boldsymbol{X}$. In the noiseless case, where $\boldsymbol{X} = \boldsymbol{L}\boldsymbol{Y}$, the range-space component can therefore be exactly recovered as $\boldsymbol{L}_\mathrm{r}\boldsymbol{Y} = \boldsymbol{L}^\dagger\boldsymbol{X}$, while only the unobservable null-space component $\boldsymbol{L}_\mathrm{n}\boldsymbol{Y}$ is learned. By restricting learning to degrees of freedom not determined by the measurements, this decomposition preserves source configurations consistent with $\boldsymbol{X}$ under $\boldsymbol{L}$ while enabling a plausible estimate of $\hat{\boldsymbol{Y}}$. This approach is closely related to prior methods that exploit range-null space decomposition for linear ill-posed problems in computer vision, such as image inpainting and super-resolution (Wang et al., 2023b; Jacome et al., 2025).

In practice, however, $\boldsymbol{X}$ often has a low signal-to-noise ratio due to instrumentation noise and motion artifacts. Consequently, directly recovering $\boldsymbol{L}_\mathrm{r}\boldsymbol{Y}$ as $\boldsymbol{L}^\dagger\boldsymbol{X}$ can propagate measurement noise into the source space, potentially yielding an unstable or overly diffuse estimate of $\hat{\boldsymbol{Y}}$.

To this end, we first denoise $\boldsymbol{X}$ before estimating $\hat{\boldsymbol{Y}}$ via range–null space decomposition as

$$\boldsymbol{X}_\mathrm{denoise} = f_\theta(\boldsymbol{X}), \qquad \hat{\boldsymbol{Y}} = \underbrace{\boldsymbol{L}^\dagger\boldsymbol{X}_\mathrm{denoise}}_{\text{Range-space component}} + \underbrace{\boldsymbol{L}_\mathrm{n} g_\gamma(\boldsymbol{L}^\dagger\boldsymbol{X}_\mathrm{denoise})}_{\text{Null-space component}}. \tag{2}$$

Here, f_θ and g_γ denote models for denoising $\boldsymbol{X}$ and estimating the null-space component of $\boldsymbol{Y}$, respectively. Under the above formulation, we define the training objective as

$$\mathcal{L} = \ell_2(\boldsymbol{Y}, \hat{\boldsymbol{Y}}) + \underbrace{\ell_2(\boldsymbol{L}\boldsymbol{Y}, \boldsymbol{X}_\mathrm{denoise})}_{\text{Denoising loss}} + \underbrace{\ell_2(\boldsymbol{L}_\mathrm{n}\boldsymbol{Y}, \boldsymbol{L}_\mathrm{n}\hat{\boldsymbol{Y}})}_{\text{Null-space loss}}, \tag{3}$$

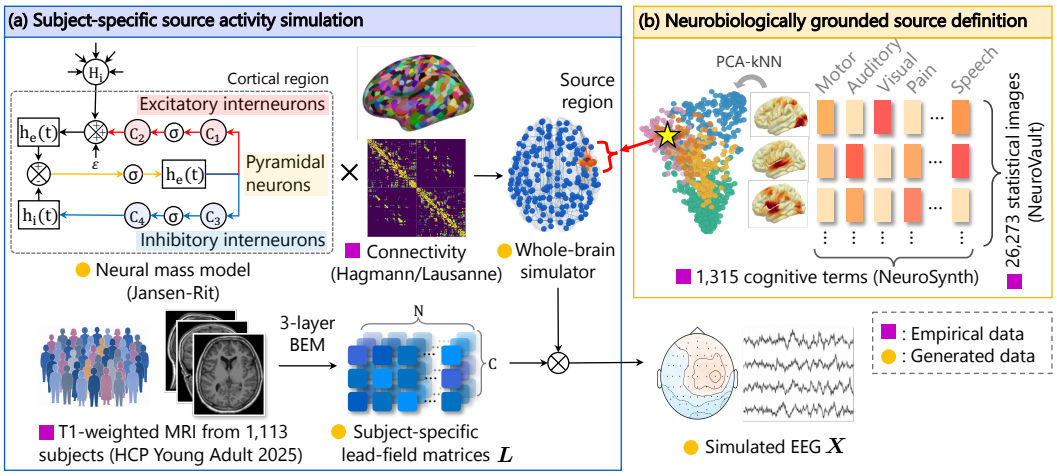

Figure 2: Whole-brain simulator for training CANDLE. (a) Subject-specific source activity simulation. Neural mass models assigned to N cortical regions are coupled via structural connectivity to generate source activity $\boldsymbol{Y}$, which is projected to EEG $\boldsymbol{X}$ through a subject-specific lead-field matrix $\boldsymbol{L}$ derived from the T1-weighted MRI of each of 1,113 subjects. (b) Neurobiologically grounded source definition. A total of 26,273 statistical maps are characterized by their correlations with 1,315 cognitive terms and clustered into representative source patterns. For each simulation, one pattern is sampled to modulate the excitatory gain of the corresponding cortical regions.

where $\ell_2(\cdot,\cdot)$ denotes the ℓ_2 loss. The first term directly supervises the final source estimate $\hat{\boldsymbol{Y}}$, whereas the second and third terms encourage f_θ and g_γ to be optimized for their respective roles of denoising and null-space component estimation. Further details of the model architecture and training configuration are provided in Appendix B.2.

2.2 Simulator Construction

Subject-specific source activity simulation. We synthesize cortical source activity and corresponding EEG by simulating a biophysical computational model on diverse subject-specific cortical geometries derived from a large-scale T1-weighted MRI database, as shown in Figure 2(a). Computational models of neural dynamics span multiple levels of granularity, from microscopic to mesoscopic and macroscopic scales (Breakspear, 2017), with the appropriate level depending on the phenomenon of interest. Because scalp EEG primarily reflects synchronized postsynaptic activity of cortical pyramidal neuron populations over spatially extended cortical patches (Tao et al., 2005), we employ a neural mass model (NMM) as a mesoscopic description of cortical dynamics. Specifically, we adopt the Jansen–Rit model (Jansen & Rit, 1995), a canonical NMM capable of reproducing characteristic EEG phenomena, including alpha oscillations and event-related potentials (Luck, 2014). Details of the Jansen–Rit model and its parameter settings are provided in Appendix B.3.

A single NMM, however, describes only the local dynamics of one cortical region and does not by itself simulate source activity distributed over the entire cortex of a given subject. We therefore construct a whole-brain simulator by assigning an NMM to each node of a cortical-region graph and coupling the nodes according to the structural connectivity defined by Hagmann et al. (2008), yielding $\boldsymbol{Y}$. Furthermore, to account for diverse cortical geometries, we instantiate the simulator across 1,113 subjects from the HCP Young Adult 2025 dataset (Van Essen et al., 2013), one of the largest publicly available MRI datasets. For each subject, $\boldsymbol{L}$ is derived from the corresponding $\mathcal{G}$ using the procedure described in Subsection 2.1 and Appendix B.1. The corresponding $\boldsymbol{X}$ is then synthesized as $\boldsymbol{X} = \boldsymbol{L}\boldsymbol{Y} + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon}$ is modeled as additive Gaussian noise. This protocol extends existing whole-brain simulators (Wang et al., 2025; Ziaeemehr et al., 2025) to ESI by incorporating subject-specific cortical geometries and their corresponding source-to-sensor mappings.

Neurobiologically grounded source definition. To synthesize realistic source activity, we define source regions from statistical brain maps derived from meta-analyses of fMRI studies and amplify the NMM activity of the corresponding cortical regions, as shown in Figure 2(b). To construct

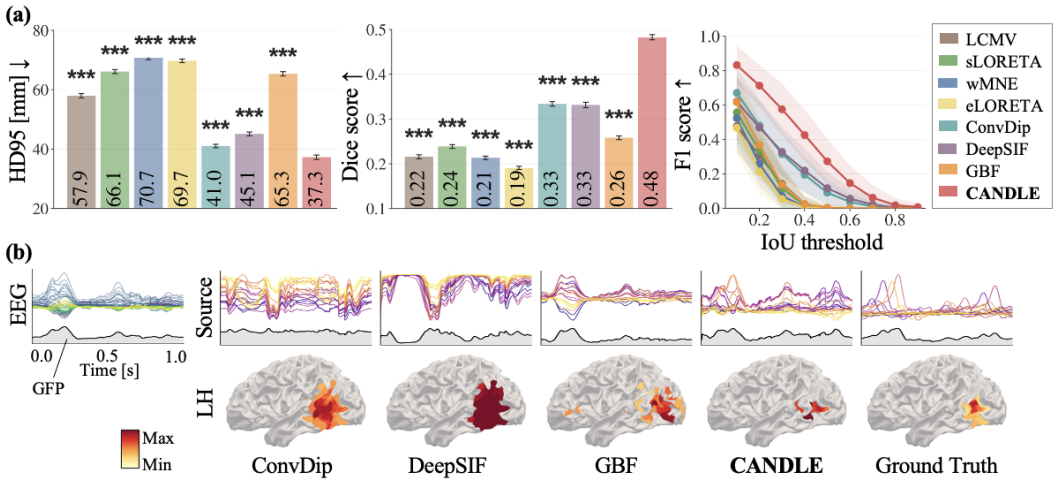

Figure 3: Simulated source activity estimation. (a) Quantitative comparison of ESI methods using HD95, Dice score, and F1 score across IoU thresholds. Error bars indicate the standard error of the mean (SEM). Statistical significance relative to CANDLE was assessed using two-sided Wilcoxon signed-rank tests (Holm corrected, *** $p < 0.001$). (b) Qualitative comparison for a representative sample, showing the EEG signals, estimated source activities, and their spatial distributions.

the source regions, we follow Suzuki & Yamashita (2021) and Wang et al. (2026) and first collect 26,273 statistical images from NeuroVault (Gorgolewski et al., 2015), a large-scale repository of statistical brain maps. Using Neurosynth (Yarkoni et al., 2011), a large-scale meta-analytic platform for fMRI studies, we then associate each image with its correlations to 1,315 cognitive terms (e.g., "motor", "auditory"), yielding a cognitive term vector for each image. These vectors are grouped into 200 clusters via PCA-kNN, and the images within each cluster are averaged to obtain a representative map per cluster, covering a broad range of cognitive terms. To further diversify the spatial distribution of the source regions, we randomly deactivate a subset of the cortical regions in each representative map according to a Bernoulli distribution ($p = 0.3$) and smooth the resulting map. Applying this perturbation four times to each representative map yields 800 distinct source regions. $\boldsymbol{Y}$ is then modulated by amplifying the excitatory postsynaptic potential gain of the NMM in each cortical region belonging to a given source configuration (see Appendix B.3), and the resulting activity is propagated to $\boldsymbol{X}$. For each simulation, a source configuration is sampled and instantiated on a subject-specific cortical geometry, yielding a single sample represented by the triplet $(\mathcal{G}, \boldsymbol{X}, \boldsymbol{Y})$.

3 Experiments

We evaluated CANDLE across three complementary settings of increasing realism. We first assessed source activity estimation on simulated data, where ground-truth cortical activity is available, and subsequently evaluated its zero-shot sim-to-real generalization on two empirical tasks: (i) intracranial stimulation localization and (ii) epileptogenic zone estimation. We further conducted ablation studies on both empirical tasks to assess the contributions of CANDLE's major components.

Across all experiments, we compared CANDLE with representative ESI methods, including LCMV (Van Veen et al., 1997), sLORETA (Pascual-Marqui, 2002), wMNE (Lin et al., 2006), eLORETA (Pascual-Marqui et al., 2011), ConvDip (Hecker et al., 2021), DeepSIF (Sun et al., 2022), and GBF (Wang et al., 2026). Among these, ConvDip and DeepSIF are learning-based ESI models and constitute the most direct baselines for CANDLE. For both empirical tasks, ConvDip, DeepSIF, and CANDLE were trained exclusively on simulated data and applied without further tuning. Unless otherwise stated, statistical significance relative to CANDLE was assessed using two-sided Wilcoxon signed-rank tests with Holm correction for multiple comparisons.

3.1 Simulated source activity estimation

Experimental setup. To evaluate ESI generalization to unseen cortical geometries, we synthesized 12,762 triplets $(\mathcal{G}, \boldsymbol{X}, \boldsymbol{Y})$ across 1,113 subject-specific cortical geometries following the pro-

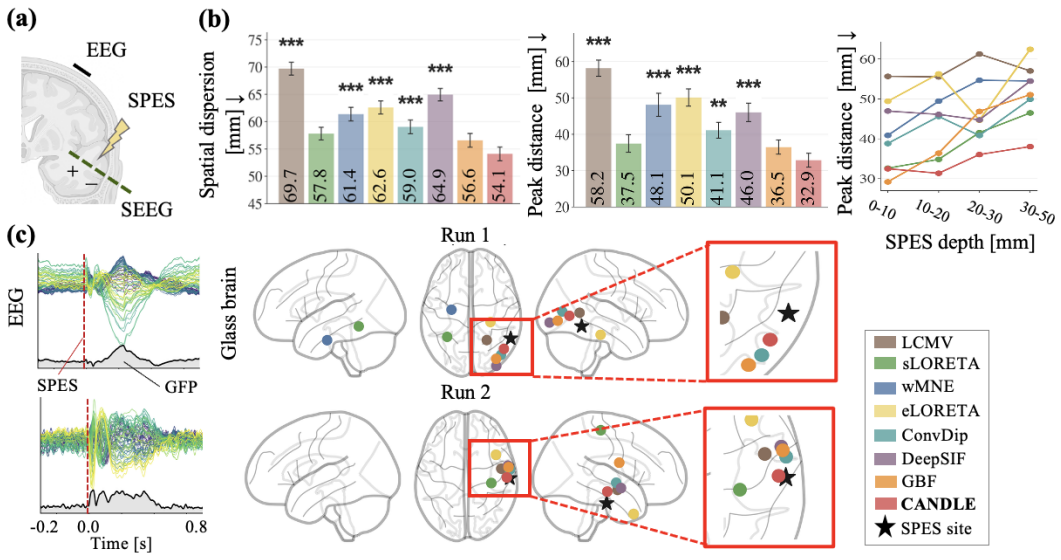

Figure 4: Intracranial stimulation localization. (a) Schematic of intracranial single-pulse electrical stimulation (SPES) site localization from simultaneous EEG. (b) Quantitative comparison using spatial dispersion and peak distance, with peak distance stratified by SPES site depth. Error bars indicate SEM. Statistical significance relative to CANDLE was assessed using two-sided Wilcoxon signed-rank tests (Holm corrected, ** $p < 0.01$, *** $p < 0.001$). (c) Qualitative comparison for two representative runs, showing EEG responses and estimated SPES locations on glass-brain maps.

cedure in Section 2.2. We then performed a random subject-wise split into training, validation, and test sets at a ratio of 0.70:0.15:0.15. For evaluation, we focused on spatial localization, a primary objective of ESI. Following prior ESI studies (Sun et al., 2022), we temporally pooled $\hat{\boldsymbol{Y}}$ to obtain a spatial source map and applied Otsu's thresholding (Otsu, 1979) to identify the estimated active region. We quantified localization performance using the 95th-percentile Hausdorff distance (HD95) and Dice score, which capture boundary discrepancy and spatial overlap between the estimated and ground-truth source regions, respectively. We further evaluated the F1 score at IoU thresholds from 0.1 to 0.9 in 0.1 increments. Definitions of evaluation metrics are provided in Appendix C.

Results. Figure 3 summarizes the results on simulated source activity estimation. As shown in Figure 3(a), CANDLE outperformed all baselines in HD95 and Dice score and consistently achieved the highest F1 scores across IoU thresholds. In particular, CANDLE achieved an HD95 of 37.3 mm and a Dice score of 0.48, improving upon the second-best baseline by 2.5 mm and 0.15 points, respectively. The improvements over all baseline methods were statistically significant ($p < 0.001$).

Furthermore, Figure 3(b) presents qualitative results for a representative sample, comparing CANDLE with the three best-performing baselines. The ground-truth source activity is spatially localized around the posterior temporal cortex and exhibits its maximum global field power (GFP) at approximately 0.2 s, while the corresponding EEG exhibits prominent GFP peaks at approximately 0.2 and 0.6 s. Given the EEG observation, the baseline methods estimate relatively diffuse source regions extending from the temporal to the occipital cortex, whereas CANDLE recovers a more localized source region around the temporal cortex. CANDLE also more faithfully reconstructs the temporal dynamics of the underlying source activity: ConvDip largely suppresses the dominant source-level GFP peak, while DeepSIF shifts its timing, whereas CANDLE recovers the peak near 0.2 s. Together, these results suggest that explicitly incorporating subject-specific cortical geometry enables CANDLE to generalize to unseen anatomies while preserving both the spatial localization and temporal dynamics of cortical source activity. Full numerical results and further qualitative results are reported in Table 2 and Figure 8 in Appendix D, respectively.

3.2 Intracranial Stimulation Localization

Experimental setup. To evaluate ESI under controlled perturbations with anatomically known source locations, we used the dataset of Parmigiani et al. (2022), which contains simultaneous scalp

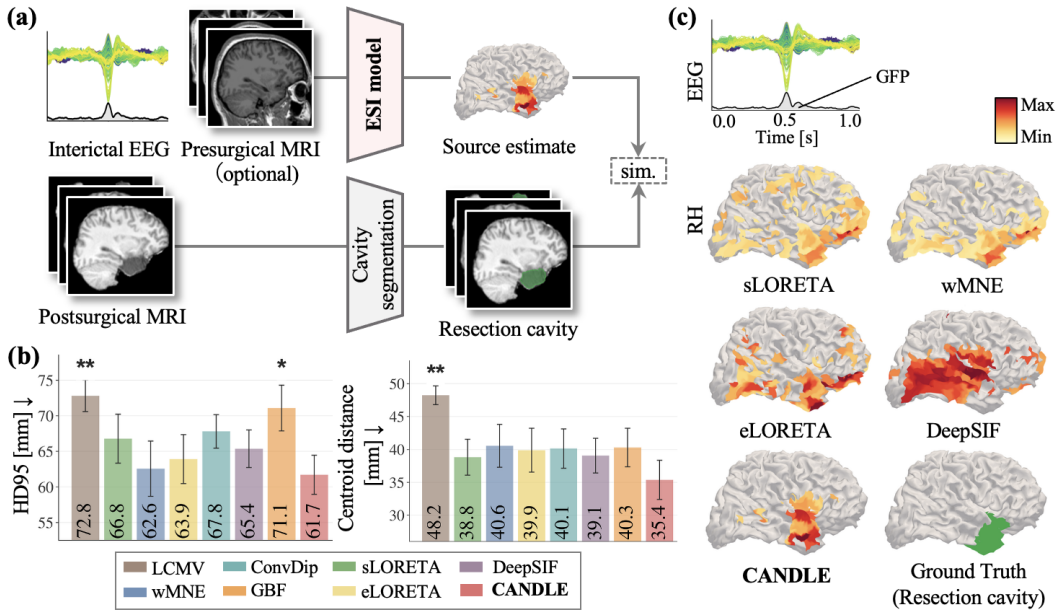

Figure 5: Epileptogenic zone (EZ) estimation. (a) Schematic of presurgical EZ localization from interictal EEG. (b) Quantitative comparison of ESI methods using HD95 and centroid distance. Error bars indicate SEM. Statistical significance relative to CANDLE was assessed using two-sided Wilcoxon signed-rank tests (Holm corrected, * $p < 0.05$, ** $p < 0.01$). (c) Qualitative comparison for a representative patient, showing the interictal EEG and estimated source distributions.

EEG and stereoelectroencephalography (SEEG) recordings during single-pulse electrical stimulation (SPES). Specifically, we evaluated whether the intracranial stimulation site could be localized from the corresponding scalp EEG, as illustrated in Figure 4(a). We included data from 32 patients with drug-resistant focal epilepsy, corresponding to 291 stimulation runs in total (9.1 per subject on average). Each run comprised an average of 31.5 SPES trials delivered at the same stimulation site. Following Wang et al. (2026), we averaged the EEG responses across trials within each run and used the resulting evoked response as input to each ESI method.

For evaluation, we considered the estimated source activity within a 10-ms window centered on SPES onset. This window excludes subsequent evoked responses (e.g., the N1 and N2 components (Hajnal et al., 2024)), which reflect propagation to remote regions rather than activity at the stimulation site. Localization performance was quantified using spatial dispersion and peak distance, which measure the spatial spread of the estimated source activity and the distance between its peak and the known SPES site, respectively. Both metrics were averaged across runs within each subject. Detailed definitions are provided in Appendix C.

Results. Figure 4(b) summarizes the quantitative results for intracranial stimulation localization. We evaluated spatial dispersion and peak distance across all stimulation runs and further stratified peak distance by SPES site depth. Across all runs, CANDLE achieved a spatial dispersion of 54.1 mm and a peak distance of 32.9 mm, the lowest errors among all methods and improvements of 2.5 mm and 3.6 mm over the second-best baseline, respectively. These improvements were statistically significant over five of the seven baselines ($p < 0.01$); the differences from sLORETA and GBF were not significant. When stratified by SPES-site depth, CANDLE achieved a peak distance of 32.48 mm for shallow stimulation sites at depths of 0–10 mm, achieving the second-best performance behind GBF (29.20 mm). Notably, the localization errors of the baseline methods generally increased as the stimulation site became deeper, whereas CANDLE remained comparatively robust to source depth and achieved the best performance in all depth ranges beyond 10 mm.

Figure 4(c) presents qualitative results for two representative stimulation runs, showing the EEG signals and the estimated SPES locations on glass-brain maps. In both cases, the ground-truth SPES sites (black stars) were located in the right temporal lobe, and the CANDLE estimates (red circles) were closest to the stimulation sites among all methods. Together, these results suggest that the

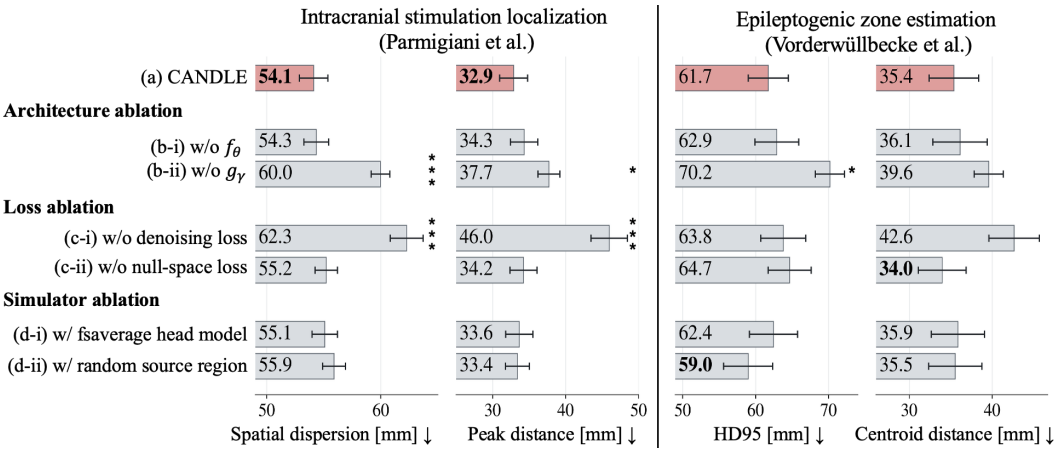

Figure 6: Ablation studies of CANDLE. Performance of the full model (a) and variants ablating the architecture (b-i, b-ii), training objective (c-i, c-ii), and simulator design (d-i, d-ii), evaluated on intracranial stimulation localization (spatial dispersion and peak distance) and epileptogenic zone estimation (HD95 and centroid distance). Error bars indicate SEM.

source activity prior learned by CANDLE exclusively from simulated data transfers to empirical EEG recordings. In particular, its robustness to SPES site depth suggests that the learned null-space prior can complement limited scalp information for deeper sources, enabling plausible localization under challenging source configurations. Complete quantitative results are listed in Table 3 in Appendix D, together with further representative runs in Figure 9.

3.3 Epileptogenic zone estimation

Experimental setup. To evaluate ESI in a clinically relevant setting, we considered epileptogenic zone (EZ) estimation from presurgical interictal EEG signals. The EZ is defined as the minimum cortical tissue whose resection or disconnection is necessary to achieve seizure freedom (Zijlmans et al., 2019). Although the EZ cannot be directly observed preoperatively, ESI of interictal epileptiform discharges (IEDs) observed in interictal EEG provides complementary information for estimating epileptogenic regions and is widely used in presurgical planning (Sun et al., 2023; 2022).

Accordingly, we evaluated CANDLE using the dataset of Vorderwülbecke et al. (2025), which provides presurgical EEG containing IEDs together with postsurgical T1-weighted MRI and surgical outcomes. Specifically, we evaluated whether the epileptogenic region could be localized from presurgical EEG together with the presurgical MRI for methods requiring subject-specific anatomy, as illustrated in Figure 5(a). We included 22 patients with favorable 12-month postsurgical outcomes, defined as International League Against Epilepsy (ILAE) (Wieser et al., 2001) classes 1 or 2. Because a favorable surgical outcome suggests that the clinically relevant EZ was encompassed by the resection, we used the postsurgical resection cavity as a reference for the EZ. The resection cavity was segmented from each postsurgical T1-weighted MRI using MELD-PostOp (Seo et al., 2026), a U-Net-based model for postoperative resection-cavity segmentation.

For evaluation, we obtained the source region by temporal pooling and Otsu's thresholding (Otsu, 1979), following Subsection 3.1. We evaluated localization using HD95 and centroid distance rather than overlap-based metrics (e.g., IoU), because surgical resections may extend beyond the true EZ to ensure complete removal of epileptogenic tissue. Detailed definitions are provided in Appendix C.

Results. Figure 5(b) presents the quantitative results for EZ estimation. CANDLE achieved the lowest errors among all methods for both HD95 and centroid distance. Specifically, CANDLE obtained an HD95 of 61.7 mm and a centroid distance of 35.4 mm, improving the second-best baseline by 0.9 mm and 3.4 mm, respectively. Compared with LCMV, CANDLE achieved significant improvements in both metrics ($p < 0.01$), and significantly improved HD95 over GBF ($p < 0.05$).

Figure 5(c) further presents qualitative results for a representative patient whose resection involved the anterior temporal lobe. Given the corresponding presurgical interictal EEG, sLORETA, wMNE,

and eLORETA produced spatially diffuse estimates extending beyond the resected region, while DeepSIF localized activity predominantly outside the resection. In contrast, although the CANDLE estimate showed some posterior extension along the temporal lobe, its dominant source region was largely concentrated around the resected anterior temporal region. Together, these results extend the zero-shot transfer observed under controlled stimulation in Section 3.2 to a more clinically realistic setting. Appendix D reports the corresponding full numerical results (Table 4) and qualitative examples for additional patients (Figure 10).

3.4 Ablation Studies

Figure 6 summarizes three ablation studies evaluated on intracranial stimulation localization (Parmigiani et al., 2022) and epileptogenic zone estimation (Vorderwülbecke et al., 2025). We use spatial dispersion and peak distance for the former, and HD95 and centroid distance for the latter.

Architecture ablation. To validate the architecture of CANDLE, we compared (a) the full model with variants removing (b-i) the denoising model f_θ or (b-ii) the null-space estimator g_γ. Both ablations degraded performance across four metrics, with peak distance increasing by 1.4 and 4.8 mm for Models (b-i) and (b-ii), respectively. The degradation was consistently larger for Model (b-ii), with statistically significant increases relative to the full model in spatial dispersion ($p < 0.001$), peak distance ($p < 0.05$), and HD95 ($p < 0.05$). These results suggest that both f_θ and g_γ contribute to CANDLE's performance, while the larger impact of removing g_γ highlights the importance of estimating source activity in the null space that is unobservable from scalp measurements.

Loss ablation. We next examined the training objective by comparing (a) the full model with variants excluding either (c-i) the denoising loss or (c-ii) the null-space loss in Equation (3). Model (c-i) degraded all four metrics, including a 13.1-mm increase in peak distance. The degradation in spatial dispersion and peak distance was significant ($p < 0.001$). By contrast, Model (c-ii) degraded three of the four metrics, except centroid distance, with peak distance increasing by 1.3 mm. These results suggest that both losses contribute to CANDLE's optimization, with the denoising loss playing a more critical role. This is likely because f_θ precedes the range–null space decomposition, such that inadequate denoising undermines the effectiveness of the entire subsequent source reconstruction.

Simulator ablation. Finally, we assessed the simulator design using two variants of the full simulator in Section 2.2: (d-i) a fixed fsaverage head model (Fischl et al., 1999) and (d-ii) random source regions of interest (ROIs). Model (d-i) degraded performance across all four metrics, whereas Model (d-ii) degraded three of the four metrics, with HD95 as the exception. In particular, peak distance increased by 0.7 and 0.5 mm for Models (d-i) and (d-ii), respectively. These results suggest that both anatomical diversity across subjects and neurobiologically grounded source regions contribute to effective training of CANDLE.

4 Conclusion and Limitations

Conclusion. In this study, we addressed electrophysiological source imaging (ESI). We proposed CANDLE, a learning-based ESI model that accounts for subject-specific cortical geometry by learning a prior over the unobservable null-space component of source activity. To train CANDLE, we further developed a whole-brain simulator that generates neurobiologically grounded source activity across more than 1,100 subject-specific cortical geometries. Without further tuning, CANDLE outperformed prior ESI methods on simulated source estimation and two empirical tasks: intracranial stimulation localization and epileptogenic zone estimation.

Limitations. CANDLE has several limitations that motivate future work. (i) The current implementation considers EEG recorded with a fixed montage (Oostenveld et al., 2001) and represents cortical source activity using the Hagmann/Lausanne parcellation (Hagmann et al., 2008). Extending the framework to accommodate diverse sensor configurations and source-space parcellations would broaden its applicability to recording settings and potentially other electrophysiological modalities, such as MEG. (ii) Our whole-brain simulator models cortical dynamics using a single canonical neural mass model, the Jansen–Rit model, although different neural dynamics models may better capture different physiological regimes. Future work could incorporate diverse neural dynamics models to learn more generalizable source priors and potentially improve sim-to-real transfer.

AI USE STATEMENT

In this work, we used generative AI tools to help develop theoretical models or conceptual frameworks, design or provide feedback on research methodology or experiments, and implement methods. Additionally, we used generative AI tools for sourcing and searching for information, identifying relevant literature, and editing the paper to improve readability. We did not use generative AI tools to generate synthetic datasets, formulate mathematical claims, provide critical ingredients for proving mathematical claims, assist in writing proofs, propose or refine hypotheses, assist with translation, clean or reformat datasets, support qualitative or thematic data analysis, or interpret results. AI-assisted implementations were reviewed and tested by the authors, and AI-assisted literature retrieval was independently verified against the original sources. We take responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This study involved only secondary analyses of previously collected human neuroimaging and electrophysiological datasets. We used data from the Human Connectome Project (Van Essen et al., 2013) and the datasets of Parmigiani et al. (2022) and Vorderwülbecke et al. (2025). No new data from human participants were collected as part of this work. All data were used in accordance with the applicable data-use terms and the ethical approvals reported in the original studies.

REPRODUCIBILITY STATEMENT

We provide the methodological details required to reproduce CANDLE and the whole-brain simulator in Section 2.1 and Appendix B.2, including the model architecture, training configuration, data preprocessing, head-model construction, and neural mass model parameters. The experimental protocols and evaluation metrics are described in Section 3 and Appendix C, respectively. The supplementary materials additionally include the implementation code and trained model checkpoints used in our experiments.

REFERENCES

Shiva Asadzadeh, Tohid Yousefi Rezaii, Soosan Beheshti, Azra Delpak, and Saeed Meshgini. A systematic review of EEG source localization techniques and their applications on diagnosis of brain abnormalities. *J. Neurosci. Methods*, 339:108740, 2020.

Gwénael Birot et al. Head model and electrical source imaging: A study of 38 epileptic patients. *NeuroImage: Clinical*, 5:77–83, 2014.

Michael Breakspear. Dynamic models of large-scale brain activity. *Nat. Neurosci.*, 20(3):340–352, 2017.

Miao Cao, Daniel Galvis, Simon J Vogrin, William P Woods, Sara Vogrin, Fan Wang, Wessel Woldman, John R Terry, Andre Peterson, Chris Plummer, et al. Virtual intracranial EEG signals reconstructed from MEG with potential for epilepsy surgery. *Nat. Commun.*, 13(1):994, 2022.

Ludovica Corona, Eleonora Tamilia, M Scott Perry, Joseph R Madsen, Jeffrey Bolton, Scellig SD Stone, Steve M Stufflebeam, Phillip L Pearl, and Christos Papadelis. Non-invasive mapping of epileptogenic networks predicts surgical outcome. *Brain*, 146(5):1916–1931, 2023.

Carlos Coronel-Oliveros, Fernando Lehue, Rubén Herzog, Iván Mindlin, Marilyn Gatica, Natalia Kowalczyk-Grkbska, Vicente Medel, Josephine Cruzat, Raul Gonzalez-Gomez, Hernán Hernandez, et al. A multi-frequency whole-brain neural mass model with homeostatic feedback inhibition. *PLoS Comput. Biol.*, 22(5):e1013463, 2026.

Anders M Dale, Arthur K Liu, Bruce R Fischl, Randy L Buckner, John W Belliveau, Jeffrey D Lewine, and Eric Halgren. Dynamic Statistical Parametric Mapping: Combining fMRI and MEG for High-Resolution Imaging of Cortical Activity. *Neuron*, 26(1):55–67, 2000.

Zhao Feng, Ioannis Kakkos, George K. Matsopoulos, Cuntai Guan, and Yu Sun. Explaining E/MEG Source Imaging and Beyond: An Updated Review. *IEEE JBHI*, 29(12):9271–9286, 2025.

Bruce Fischl. FreeSurfer. *NeuroImage*, 62(2):774–781, 2012.

Bruce Fischl, Martin I Sereno, Roger BH Tootell, and Anders M Dale. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum. Brain Mapp.*, 8(4):272–284, 1999.

Yonatan Gideoni, Ryan Charles Timms, and Oiwi Parker Jones. Non-invasive neural decoding in source reconstructed brain space. *ArXiv*, 2410.19838, 2024.

Krzysztof J Gorgolewski, Gael Varoquaux, Gabriel Rivera, Yannick Schwarz, Satrajit S Ghosh, Camille Maumet, Vanessa V Sochat, Thomas E Nichols, Russell A Poldrack, Jean-Baptiste Poline, et al. NeuroVault.org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain. *Front. Neuroinform.*, 9:8, 2015.

Patric Hagmann, Leila Cammoun, Xavier Gigandet, Reto Meuli, Christopher J Honey, Van J Wedeen, and Olaf Sporns. Mapping the structural core of human cerebral cortex. *PLoS biology*, 6(7):e159, 2008.

Boglárka Hajnal, Johanna Petra Szabó, Emília Tóth, Corey J Keller, Lucia Wittner, Ashesh D Mehta, Loránd Erőss, István Ulbert, Dániel Fabó, and László Entz. Intracortical mechanisms of single pulse electrical stimulation (SPES) evoked excitations and inhibitions in humans. *Sci. Rep.*, 14 (1):13784, 2024.

Hans Hallez, Bart Vanrumste, Roberta Grech, Joseph Muscat, Wim De Clercq, Anneleen Vergult, Yves D'Asseler, Kenneth P Camilleri, Simon G Fabri, Sabine Van Huffel, et al. Review on solving the forward problem in eeg source analysis. *J. NeuroEng. Rehabil.*, 4(1):46, 2007.

Bin He and Zhongming Liu. Multimodal Functional Neuroimaging: Integrating Functional MRI and EEG/MEG. *IEEE Rev. Biomed. Eng.*, 1:23–40, 2008.

Lukas Hecker, Rebekka Rupprecht, Ludger Tebartz Van Elst, and Jürgen Kornmeier. ConvDip: A convolutional neural network for better EEG source imaging. *Front. Neurosci.*, 15:569918, 2021.

Alan L Hodgkin and Andrew F Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.*, 117(4):500–544, 1952.

Weizhe Hua, Zihang Dai, Hanxiao Liu, and Quoc Le. Transformer Quality in Linear Time. In *ICML*, pp. 9099–9117, 2022.

Roman Jacome, Romario Gualdrón-Hurtado, León Suárez-Rodríguez, et al. NPN: Non-Linear Projections of the Null-Space for Imaging Inverse Problems. *ArXiv*, 2510.01608, 2025.

Ben H Jansen and Vincent G Rit. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biol. Cybern.*, 73(4):357–366, 1995.

Dulhan Jayalath and Oiwi Parker Jones. MEG-XL: Data-Efficient Brain-to-Text via Long-Context Pre-Training. In *ICML*, 2026.

Meng Jiao, Xiaochen Xian, Boyu Wang, Yu Zhang, Shihao Yang, Spencer Chen, Hai Sun, and Feng Liu. XDL-ESI: Electrophysiological sources imaging via explainable deep learning framework with validation on simultaneous EEG and IEEG. *NeuroImage*, 299:120802, 2024a.

Meng Jiao, Shihao Yang, Xiaochen Xian, Neel Fotedar, and Feng Liu. Multi-Modal Electrophysiological Source Imaging With Attention Neural Networks Based on Deep Fusion of EEG and MEG. *IEEE TNSRE*, 32:2492–2502, 2024b.

Taesung Kim, Jinhee Kim, Yunwon Tae, Cheonbok Park, Jang-Ho Choi, et al. Reversible Instance Normalization for Accurate Time-Series Forecasting against Distribution Shift. In *ICLR*, 2022.

Kimi Team. Kimi linear: An expressive, efficient attention architecture. *ArXiv*, 2510.26692, 2025.

Fa-Hsuan Lin, Thomas Witzel, Seppo P Ahlfors, Steven M Stufflebeam, John W Belliveau, and Matti S Hämäläinen. Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. *NeuroImage*, 31(1):160–171, 2006.

Yong Liu, Tengge Hu, Haoran Zhang, Haixu Wu, Shiyu Wang, Lintao Ma, and Mingsheng Long. iTransformer: Inverted Transformers Are Effective for Time Series Forecasting. In *ICLR*, 2024.

Steven J Luck. *An introduction to the event-related potential technique*. The MIT Press, 2 edition, 2014.

Christopher W Lynn and Danielle S Bassett. The physics of brain network structure, function and control. *Nat. Rev. Phys.*, 1(5):318–332, 2019.

Davide Momi, Zheng Wang, Sara Parmigiani, Ezequiel Mikulan, Sorenza P Bastiaens, Mohammad P Oveisi, Kevin Kadak, Gianluca Gaglioti, Allison C Waters, Sean Hill, et al. Stimulation mapping and whole-brain modeling reveal gradients of excitability and recurrence in cortical networks. *Nat. Commun.*, 16(1):3222, 2025.

J.C. Mosher, P.S. Lewis, and R.M. Leahy. Multiple dipole modeling and localization from spatio-temporal MEG data. *IEEE Trans. Biomed. Eng.*, 39(6):541–557, 1992.

Robert Oostenveld et al. The five percent electrode system for high-resolution EEG and ERP measurements. *Clin. Neurophysiol.*, 112(4):713–719, 2001.

Nobuyuki Otsu. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man Cybern.*, 9(1):62–66, 1979.

S Parmigiani, E Mikulan, S Russo, S Sarasso, FM Zauli, A Rubino, A Cattani, M Fecchio, D Giampiccolo, J Lanzone, et al. Simultaneous stereo-EEG and high-density scalp EEG recordings to study the effects of intracerebral stimulation parameters. *Brain Stimul.*, 15(3):664–675, 2022.

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

Roberto D Pascual-Marqui, Dietrich Lehmann, Martha Koukkou, Kieko Kochi, Peter Anderer, Bernd Saletu, Hideaki Tanaka, Koichi Hirata, E Roy John, Leslie Prichep, et al. Assessing interactions in the brain with exact low-resolution electromagnetic tomography. *Philos. Trans. R. Soc. A*, 369(1952):3768–3784, 2011.

Michael Schirner, Lia Domide, Dionysios Perdikis, Paul Triebkorn, Leon Stefanovski, Roopa Pai, Paula Prodan, Bogdan Valean, Jessica Palmer, Chloê Langford, et al. Brain simulation as a cloud service: The Virtual Brain on EBRAINS. *NeuroImage*, 251:118973, 2022.

Johannes Schwab, Stephan Antholzer, and Markus Haltmeier. Deep null space learning for inverse problems: convergence analysis and rates. *Inverse Problems*, 35(2):025008, 2019.

Jieun Seo, Mathilde Ripart, Helene Kaas, Cornelius Kronlage, Ben Sinclair, Lucy Vivash, Merran R Courtney, Terence J O'Brien, Siby Gopinath, Harilal Parasuram, et al. Automated segmentation of postsurgical resection cavities on magnetic resonance imaging in focal epilepsy: A Multicentre Epilepsy Lesion Detection study. *Epilepsia*, 2026.

Abbas Sohrabpour, Zhengxiang Cai, Shuai Ye, Benjamin Brinkmann, Gregory Worrell, and Bin He. Noninvasive electromagnetic source imaging of spatiotemporally distributed epileptogenic brain sources. *Nat. Commun.*, 11(1):1946, 2020.

Rui Sun, Abbas Sohrabpour, Gregory A Worrell, and Bin He. Deep neural networks constrained by neural mass models improve electrophysiological source imaging of spatiotemporal brain dynamics. *PNAS*, 119(31):e2201128119, 2022.

Rui Sun, Wenbo Zhang, Anto Bagić, and Bin He. Deep learning based source imaging provides strong sublobar localization of epileptogenic zone from MEG interictal spikes. *NeuroImage*, 281: 120366, 2023.

Keita Suzuki and Okito Yamashita. MEG current source reconstruction using a meta-analysis fMRI prior. *NeuroImage*, 236:118034, 2021.

Shuntaro Suzuki, Shunya Nagashima, and Komei Sugiura. Cortical-SSM: a deep state space model for motor imagery decoding from EEG signals. *J. Neural Eng.*, 23(4):046032, 2026.

James X Tao, Amit Ray, Susan Hawes-Ebersole, and John S Ebersole. Intracranial EEG substrates of scalp EEG interictal spikes. *Epilepsia*, 46(5):669–676, 2005.

David C. Van Essen, Stephen M. Smith, Deanna M. Barch, Timothy E.J. Behrens, Essa Yacoub, and Kamil Ugurbil. The WU-Minn Human Connectome Project: An overview. *NeuroImage*, 80: 62–79, 2013.

B.D. Van Veen, W. Van Drongelen, M. Yuchtman, and A. Suzuki. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans. Biomed. Eng.*, 44(9):867–880, 1997.

Bernd J Vorderwülbecke, Margherita Carboni, Sebastien Tourbier, Laurent Spinelli, Denis Brunet, Martin Seeber, Christian M Korff, Shahan Momjian, Maria Vargas, Margitta Seeck, et al. High-Density EEG Source Localisation of averaged interictal epileptic Discharges validated by surgical Outcome. *Sci. Data*, 12(1):1441, 2025.

Chaoming Wang, Tianqiu Zhang, Xiaoyu Chen, Sichao He, Shangyang Li, and Si Wu. BrainPy, a flexible, integrative, efficient, and extensible framework for general-purpose brain dynamics programming. *eLife*, 12:e86365, 2023a.

Huifang E Wang, Borana Dollomaja, Paul Triebkorn, Gian Marco Duma, Adam Williamson, Julia Makhalova, Jean-Didier Lemarechal, Fabrice Bartolomei, and Viktor Jirsa. Virtual brain twins for stimulation in epilepsy. *Nat. Comput. Sci.*, 5(9):754–768, 2025.

Song Wang, Kexin Lou, Chen Wei, Zhiyuan Sheng, Jiahao Tang, Kaining Peng, Xinke Shen, Shuhao Mei, Liang Chen, Dongfeng Gu, et al. A geometry aware framework enhances noninvasive mapping of whole human brain dynamics. *Nat. Biomed. Eng.*, 2026.

Yinhuai Wang, Jiwen Yu, and Jian Zhang. Zero-Shot Image Restoration Using Denoising Diffusion Null-Space Model. In *ICLR*, 2023b.

Fabrice Wendling, Fabrice Bartolomei, Jean J Bellanger, and Patrick Chauvel. Epileptic fast activity can be explained by a model of impaired GABAergic dendritic inhibition. *Eur. J. Neurosci.*, 15 (9):1499–1508, 2002.

Vincent Wens. Exploring the limits of MEG spatial resolution with multipolar expansions. *NeuroImage*, 270:119953, 2023.

HG Wieser, WT Blume, D Fish, E Goldensohn, A Hufnagel, D King, MR Sperling, and H Luders. Proposal for a new classification of outcome with respect to epileptic seizures following epilepsy surgery. *Epilepsia*, 42(2):282–286, 2001.

Hugh R Wilson and Jack D Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.*, 12(1):1–24, 1972.

Songlin Yang, Bailin Wang, Yu Zhang, Yikang Shen, et al. Parallelizing Linear Transformers with the Delta Rule over Sequence Length. In *NeurIPS*, volume 37, pp. 115491–115522, 2024.

Tal Yarkoni, Russell A Poldrack, Thomas E Nichols, David C Van Essen, et al. Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods*, 8(8):665–670, 2011.

Abolfazl Ziaeemehr, Marmaduke Woodman, Lia Domide, Spase Petkoski, Viktor Jirsa, and Meysam Hashemi. Virtual Brain Inference (VBI), a flexible and integrative toolkit for efficient probabilistic inference on whole-brain models. *eLife*, 14:RP106194, 2025.

Maeike Zijlmans, Willemiek Zweiphenning, and Nicole Van Klink. Changing concepts in presurgical assessment for epilepsy surgery. *Nat. Rev. Neurol.*, 15(10):594–606, 2019.

A Related Works

Electrophysiological source imaging. Electrophysiological source imaging (ESI) methodologies and their applications have been systematically reviewed by Asadzadeh et al. (2020) and Feng et al. (2025). Because ESI is severely ill-posed, conventional approaches address this non-uniqueness by imposing explicit spatial constraints on the cortical source space. Early methods such as MUSIC (Mosher et al., 1992) restrict source activity to one or a few equivalent current dipoles, whereas minimum-norm–based estimators such as sLORETA (Pascual-Marqui, 2002) and dSPM (Dale et al., 2000) impose depth-dependent weighting on distributed source estimates. More recent methods introduce richer structured priors over the source space. FAST-IRES (Sohrabpour et al., 2020) promotes source sparsity under a temporal basis derived from the observed EEG signals, yielding spatially localized source estimates. GBF (Wang et al., 2026) regresses the weights of subject-specific eigenmodes derived from cortical geometry to obtain spatially smooth source estimates.

In contrast to these explicitly constrained formulations, learning-based ESI models seek to acquire data-driven priors over source activity, thereby reducing reliance on manually specified spatial constraints. ConvDip (Hecker et al., 2021), an early representative, learns a CNN-based mapping from EEG topographies to source activity. DeepSIF (Sun et al., 2022) further demonstrates that a model trained exclusively on synthetic data generated using biophysical computational models can transfer to real recordings. Subsequent work has extended this paradigm in several directions, with XDL-ESI (Jiao et al., 2024a) introducing explainability into learning-based ESI, while MMDF-ANN (Jiao et al., 2024b) improves source-estimation accuracy by jointly exploiting multimodal EEG/MEG signals. Most existing learning-based methods, however, are trained in a common template source space such as fsaverage (Fischl et al., 1999), and their application to individual subjects may therefore require post-training adaptation to subject-specific anatomy (Sun et al., 2023).

Brain activity simulators. Mathematical models of brain activity span multiple spatial scales, from biophysically detailed single-neuron models such as Hodgkin–Huxley (Hodgkin & Huxley, 1952) to macroscopic neural field models of large-scale cortical dynamics (Breakspear, 2017). Between these scales, neural mass models (NMMs) describe the aggregate dynamics of interacting neuronal populations and are well suited to electrophysiological activity at the spatial scale relevant to ESI. Representative NMMs include the Wilson–Cowan (Wilson & Cowan, 1972) and Jansen–Rit (Jansen & Rit, 1995) models, with subsequent variants developed for specific regimes such as epileptiform activity (Wendling et al., 2002) and sleep-related dynamics (Coronel-Oliveros et al., 2026).

More recently, brain activity simulations have been extended to the whole-brain scale by embedding NMMs at the nodes of structural connectomes derived from diffusion MRI (Lynn & Bassett, 2019). Frameworks such as BrainPy (Wang et al., 2023a), The Virtual Brain (Schirner et al., 2022), and Virtual Brain Inference (Ziaeemehr et al., 2025) facilitate such simulations. However, these simulators are primarily designed to model brain dynamics and often abstract away anatomical variability and the spatial organization of cortical activity, limiting their suitability for applications requiring anatomically realistic simulations across subjects, such as training learning-based ESI models.

B Detailed Procedure of CANDLE

B.1 Lead-field matrix calculation

Among various head-modeling approaches for constructing the lead-field matrix $\boldsymbol{L}$, CANDLE derives $\boldsymbol{L}$ from the subject-specific anatomical image $\mathcal{G}$ by modeling volume conduction through the head tissues using a boundary element model (BEM) (Hallez et al., 2007). Specifically, we first segment the tissue boundaries relevant to EEG generation and volume conduction from $\mathcal{G}$ using a watershed-based segmentation procedure. These surfaces include the white-matter surface, which defines the cortical source space, as well as the inner-skull, outer-skull, and scalp surfaces through which the electrical potentials propagate before reaching the EEG electrodes. Each tissue boundary is subsequently tessellated to construct a surface-based head model, from which volume conduction is modeled using a three-layer BEM. We assign conductivities of 0.3, 0.006, and $0.3\ \mathrm{S/m}$ to the brain, skull, and scalp compartments, respectively.

BEM formulations can incorporate different numbers and configurations of tissue compartments. In this work, we adopt a three-layer model. Birot et al. (2014) showed that increasing the complexity of volume-conduction models has only a negligible effect on source localization performance for ESI with empirical EEG recordings. Based on this finding, we employ a three-layer BEM to reduce computational cost while preserving the anatomical detail relevant to EEG propagation.

We further discretize the cortical source space according to the Hagmann/Lausanne parcellation (Hagmann et al., 2008). Previous physiological studies have suggested that cortical activity must be synchronized over an area of approximately 6–10 cm^2 to generate scalp potentials of sufficient magnitude to be detectable by EEG (Tao et al., 2005). Motivated by this spatial scale, we model the cortical surface using the $N = 994$ regions of the Hagmann/Lausanne parcellation, each covering approximately $1.5\ \mathrm{cm}^2$ of cortical surface. The resulting subject-specific model defines $\boldsymbol{L}$, which maps cortical source activity to EEG measurements.

B.2 Additional Methodological Details

Model configuration. The architectures adopted for the denoising model f_θ and null-space estimation model g_γ are illustrated in Figure 7. As the backbone, we employ Kimi Delta Attention (Kimi Team, 2025), a cutting-edge architecture from the family of Delta Rule-based models (Yang et al., 2024) that have demonstrated strong sequence-modeling performance across diverse modalities. We extend this backbone with a gated attention unit-based module (Hua et al., 2022), which has also shown effectiveness in sequence modeling. We also adopt inverted normalization along the temporal dimension, following prior works (Kim et al., 2022; Liu et al., 2024) showing that this strategy improves robustness to non-stationarity in time-series modeling. We then stack M such blocks to construct the overall architecture.

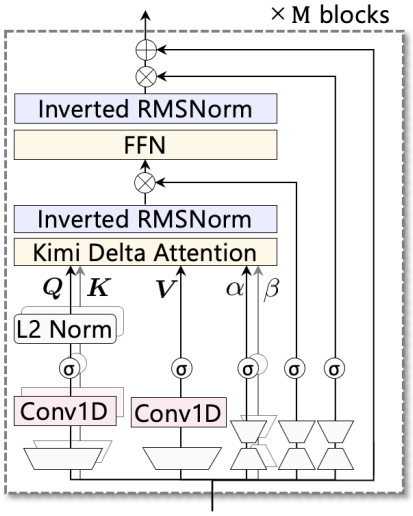

Figure 7: Architectures adopted for the denoising model f_θ and null-space estimation model g_γ in CANDLE.

We trained CANDLE for 20 epochs with a batch size of 64 using AdamW ($\beta_1 = 0.9,\ \beta_2 = 0.999$), a learning rate of 1.0×10^{-4}, and a weight decay of 1.0×10^{-2}. We set the sequence length to $T = 500$, corresponding to a one-second EEG segment sampled at 500 Hz, and used $M = 2$ stacked blocks. The number of EEG channels was fixed at $C = 64$, following the standard BioSemi 64-channel montage (Oostenveld et al., 2001), while the cortical source space comprised $N = 994$ regions based on the Hagmann–Lausanne parcellation (Hagmann et al., 2008), as detailed in Appendix B.1. Under this configuration, CANDLE contained approximately 11.6 M trainable parameters and required approximately 15.30 G FLOPs.

We split the simulated dataset into training, validation, and test sets. The training set was used for parameter optimization, the validation set for model selection, and the test set exclusively for final evaluation. After each training epoch, we evaluated the checkpoint on the validation set using the 95th-percentile Hausdorff distance (HD95), following the procedure described in Subsection 3.1 and Appendix C. Among the resulting checkpoints, we selected the model achieving the lowest HD95 on the validation set and used this checkpoint for evaluation on the test set.

All experiments were performed on an NVIDIA GeForce RTX 4090 with 24 GB of VRAM and an Intel Core i9-13900KF processor with 64 GB of RAM. Training CANDLE under the configuration described above required approximately 1.8 hours.

Data preprocessing. EEG signals were downsampled to 500 Hz and band-pass filtered between 0.5 and 45 Hz using a Butterworth filter. We subsequently applied common average referencing and z-score normalization. For subject-specific anatomical processing, T1-weighted MRIs were processed using FreeSurfer (Fischl, 2012) to obtain the cortical geometry required for constructing the subject-specific lead-field matrix.

Table 1: Parameters of the Jansen–Rit model.

Symbol	Value	Description
A_i	$3.25\ \mathrm{mV}$	Average amplitude of the excitatory postsynaptic potential
B	$22\ \mathrm{mV}$	Average amplitude of the inhibitory postsynaptic potential
a	$100\ \mathrm{s}^{-1}$	Time constant of the excitatory synaptic response
b	$50\ \mathrm{s}^{-1}$	Time constant of the inhibitory synaptic response
C_1, C_2	$135,\ 108$	Average pyramidal cell–excitatory interneuron synaptic connectivity
C_3, C_4	$33.75,\ 33.75$	Average pyramidal cell–inhibitory interneuron synaptic connectivity
$e_{\max}$	$5\ \mathrm{Hz}$	Maximum firing rate
v_0	$6\ \mathrm{mV}$	Membrane potential at half-maximum firing rate
r	$0.56\ \mathrm{mV}^{-1}$	Slope of the sigmoid function at v_0

B.3 Jansen–Rit model

We employ the Jansen–Rit neural mass model (Jansen & Rit, 1995) to generate mesoscopic cortical activity for each cortical region in the whole-brain simulator. The model describes the interactions among pyramidal neurons and excitatory and inhibitory interneurons through second-order synaptic dynamics. For cortical region i, the observable source activity $y_i(t)$ is defined as the difference between the excitatory and inhibitory postsynaptic potentials,

$$y_i(t) = y_{1,i}(t) - y_{2,i}(t), \tag{4}$$

where $y_{1,i}$ and $y_{2,i}$ denote the excitatory and inhibitory postsynaptic potentials received by the pyramidal population, respectively. The dynamics are given by

$$\ddot{y}_{0,i}(t) = A_i a\, \sigma(y_i(t)) - 2a\dot{y}_{0,i}(t) - a^2 y_{0,i}(t), \tag{5}$$

$$\ddot{y}_{1,i}(t) = A_i a\, [C_2 \sigma(C_1 y_{0,i}(t)) + H_i(t) + \xi_i(t)] - 2a\dot{y}_{1,i}(t) - a^2 y_{1,i}(t), \tag{6}$$

$$\ddot{y}_{2,i}(t) = Bb\, [C_4 \sigma(C_3 y_{0,i}(t))] - 2b\dot{y}_{2,i}(t) - b^2 y_{2,i}(t), \tag{7}$$

where $y_{0,i}$ denotes the excitatory postsynaptic potential generated by the pyramidal population, $H_i(t)$ denotes the interregional excitatory input from other cortical regions, $\xi_i(t)$ denotes the stochastic input, and $\sigma(\cdot)$ converts the average membrane potential into the population firing rate:

$$\sigma(v) = \frac{e_{\max}}{1 + \exp\{r(v_0 - v)\}}. \tag{8}$$

The remaining model parameters are summarized in Table 1.

To extend the local Jansen–Rit dynamics to the whole cortex, we instantiate one neural mass model at each of the N cortical regions and couple their pyramidal populations according to the structural connectivity of the Hagmann–Lausanne parcellation (Hagmann et al., 2008). The interregional input $H_i(t)$ is computed as

$$H_i(t) = \sum_{j=1}^{N} W_{ij} \sigma\left(y_j(t)\right), \tag{9}$$

where W_{ij} denotes the normalized structural connectivity from region j to region i, with $W_{ii} = 0$. The resulting regional activities are concatenated to form the whole-brain cortical source activity $\boldsymbol{Y}(t) = [y_1(t), \ldots, y_N(t)]^\top$.

Following Sun et al. (2022), the source configurations defined in Subsection 2.2 are implemented by locally modulating the excitatory synaptic gain A_i. For a sampled source configuration $\mathcal{S}$, we set

$$A_i = \begin{cases} 3.50\ \mathrm{mV}, & i \in \mathcal{S}, \\ 3.25\ \mathrm{mV}, & i \notin \mathcal{S}, \end{cases} \tag{10}$$

while keeping all other Jansen–Rit parameters unchanged. Thus, source activity is generated by selectively increasing the excitatory postsynaptic response of neurobiologically defined cortical regions rather than directly injecting an artificial waveform into the source space.

The stochastic differential equations were numerically integrated using the stochastic Heun method at a sampling rate of 2,000 Hz. Each simulation began with a 1-s burn-in period to reduce dependence on the initial state. This period was discarded, and the subsequent 10 s of simulated activity were used for analysis. We generated three independent trials for each source configuration under each of three noise conditions, resulting in 30 s of simulated activity per noise condition. The simulated cortical activity was subsequently downsampled from 2,000 to 500 Hz.

C Evaluation Metrics

We provide detailed definitions of the evaluation metrics used in Section 3. All spatial distances are computed in the subject-specific anatomical space.

95th-percentile Hausdorff distance (HD95). Let $\mathcal{S}$ and $\widehat{\mathcal{S}}$ denote the ground-truth and estimated source regions, respectively. For two sets of spatial points $\mathcal{A}$ and $\mathcal{B}$, the directed nearest-neighbor distances are defined as

$$D(\mathcal{A},\mathcal{B}) = \left\{ \min_{\boldsymbol{b}\in\mathcal{B}} \|\boldsymbol{a}-\boldsymbol{b}\|_2 \,\middle|\, \boldsymbol{a}\in\mathcal{A} \right\}. \tag{11}$$

The HD95 is then defined as follows, where $Q_{0.95}(\cdot)$ denotes the 95th percentile:

$$\mathrm{HD95}(\mathcal{S},\widehat{\mathcal{S}}) = \max\left\{ Q_{0.95}\Big(D(\mathcal{S},\widehat{\mathcal{S}})\Big), Q_{0.95}\Big(D(\widehat{\mathcal{S}},\mathcal{S})\Big) \right\}. \tag{12}$$

Dice score. The Dice score quantifies the spatial overlap between the estimated and ground-truth source regions and is defined as

$$\mathrm{Dice}(\mathcal{S},\widehat{\mathcal{S}}) = \frac{2|\mathcal{S}\cap\widehat{\mathcal{S}}|}{|\mathcal{S}|+|\widehat{\mathcal{S}}|}. \tag{13}$$

Spatial dispersion. For intracranial stimulation localization, we quantify the spatial spread of the estimated source activity relative to the stimulation site using spatial dispersion (SD). Let $\boldsymbol{r}_{\mathrm{stim}}$ denote the stimulation location, $\boldsymbol{r}_k$ the location of the k-th cortical source, and $d_k = \|\boldsymbol{r}_k - \boldsymbol{r}_{\mathrm{stim}}\|_2$. Following the definition of Vorderwülbecke et al. (2025), spatial dispersion is calculated as follows, where R_k denotes the estimated activity of the k-th source at the time point of maximum source activity:

$$\mathrm{SD} = \sqrt{\frac{\sum_{k=1}^{N} d_k^2 R_k^2}{\sum_{k=1}^{N} R_k^2}}. \tag{14}$$

F1 score across IoU thresholds. IoU measures the overlap between the estimated and reference source regions relative to their union. For each IoU threshold from 0.1 to 0.9 in increments of 0.1, an estimate was considered correct if its IoU with the ground-truth region exceeded the threshold, and the corresponding F1 score was computed.

Peak distance. Peak distance is the Euclidean distance between the known stimulation site and the cortical source with the maximum estimated activity.

Centroid distance. Centroid distance is the Euclidean distance between the spatial centroids of the estimated and ground-truth source regions.

D Additional experimental results

Additional qualitative results. Figures 8, 9, and 10 present additional qualitative results for simulated source activity estimation, intracranial stimulation localization, and epileptogenic zone estimation, respectively. Figure 8 shows two additional test samples, comparing all eight ESI methods. As in Figure 3(b), the conventional methods and the learning-based baselines produced spatially diffuse or displaced estimates, whereas CANDLE recovered a source region whose location and extent

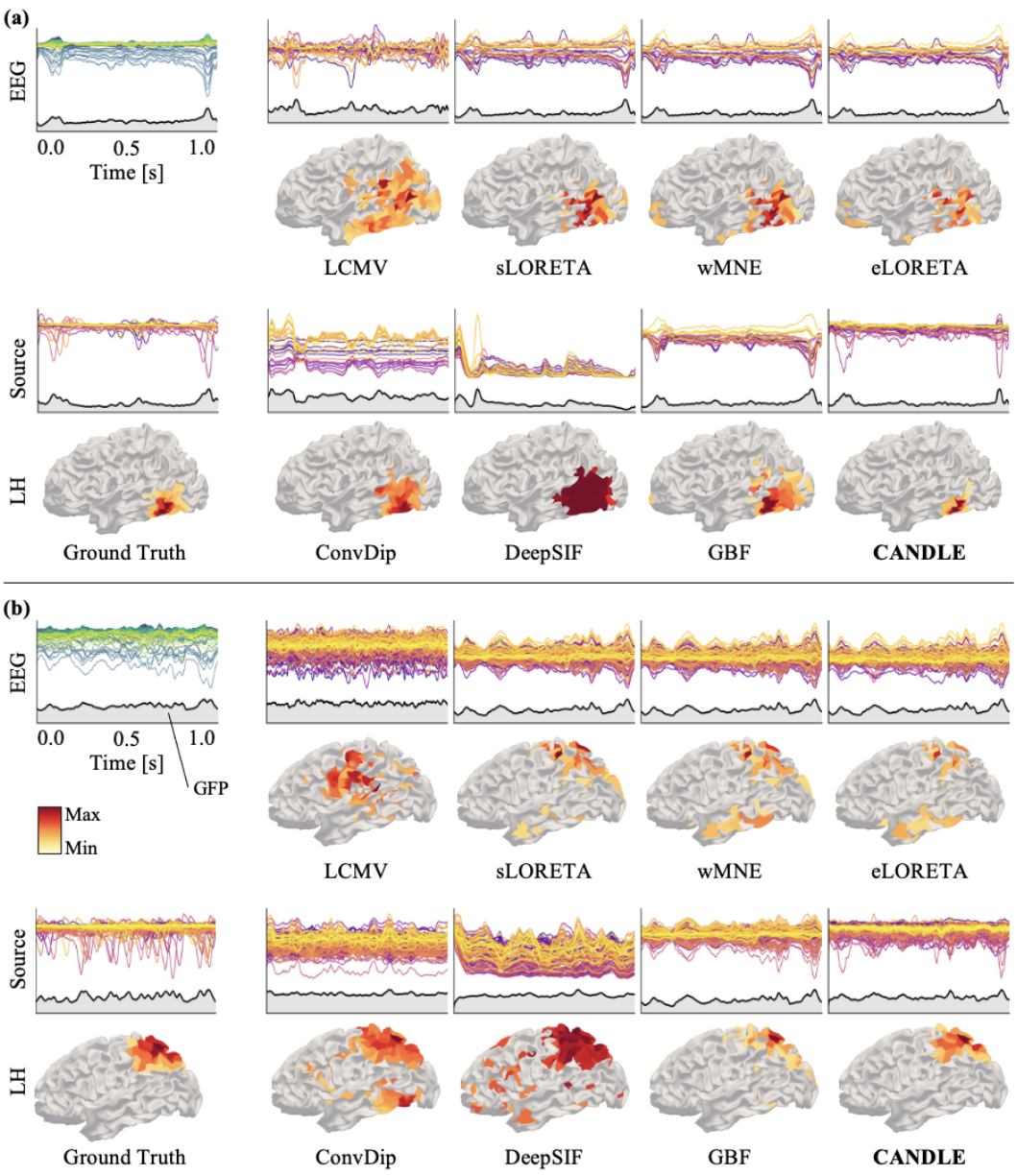

Figure 8: Additional qualitative results for simulated source activity estimation. Two representative test samples, (a) and (b). In each panel, the leftmost column shows the EEG input with its global field power (GFP), the ground-truth source activity, and its temporally pooled spatial map. The remaining columns show, for each of the eight ESI methods, the estimated source activity with its GFP and the corresponding temporally pooled spatial map.

closely matched the ground truth. Figure 9 shows four additional stimulation runs from different patients, with the estimated SPES locations displayed in left lateral, superior, and right lateral views. Figure 10 shows two additional patients with right- and left-hemisphere resections. In both cases, the dominant source region estimated by CANDLE lay within or adjacent to the resection cavity.

Full quantitative results. Tables 2, 3, and 4 report the complete numerical results underlying Figures 3(a), 4(b), and 5(b), respectively. Across all three tables, CANDLE achieved the best or second-best value in every metric, and the only case in which a baseline outperformed CANDLE was the shallowest SPES sites (0–10 mm), where GBF yield a lower peak distance.

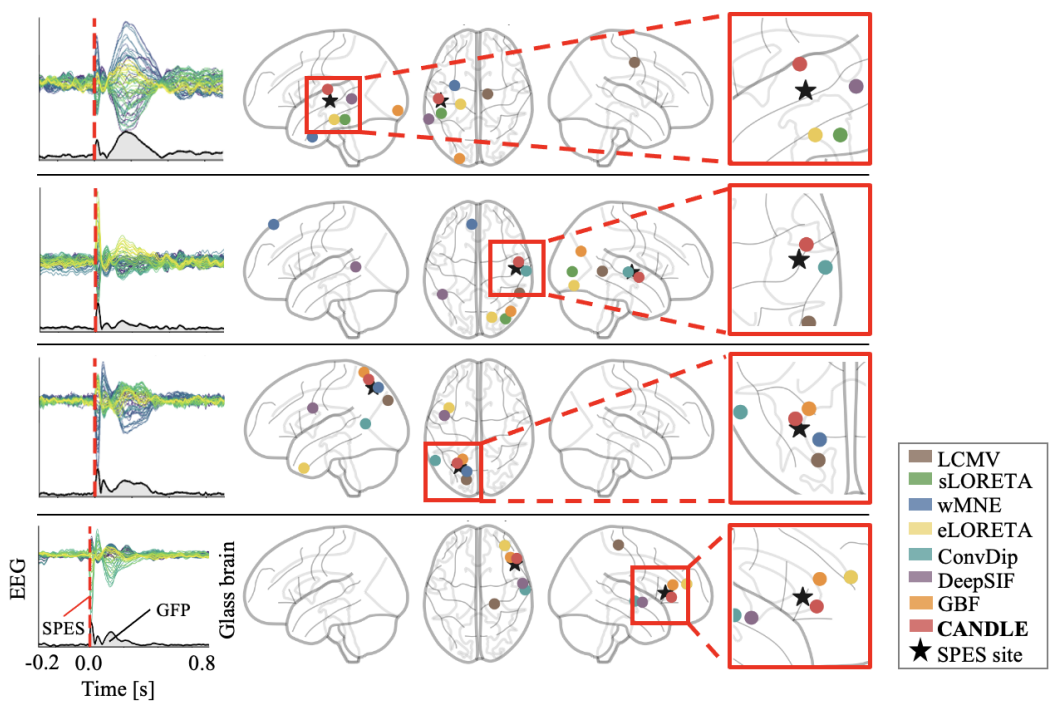

Figure 9: Additional qualitative results for intracranial stimulation localization. Four representative stimulation runs from different patients. In each row, the left panel shows the input EEG with its GFP, where the dashed line marks SPES onset. The remaining panels show the estimated SPES locations on glass-brain maps in left lateral, superior, and right lateral views.

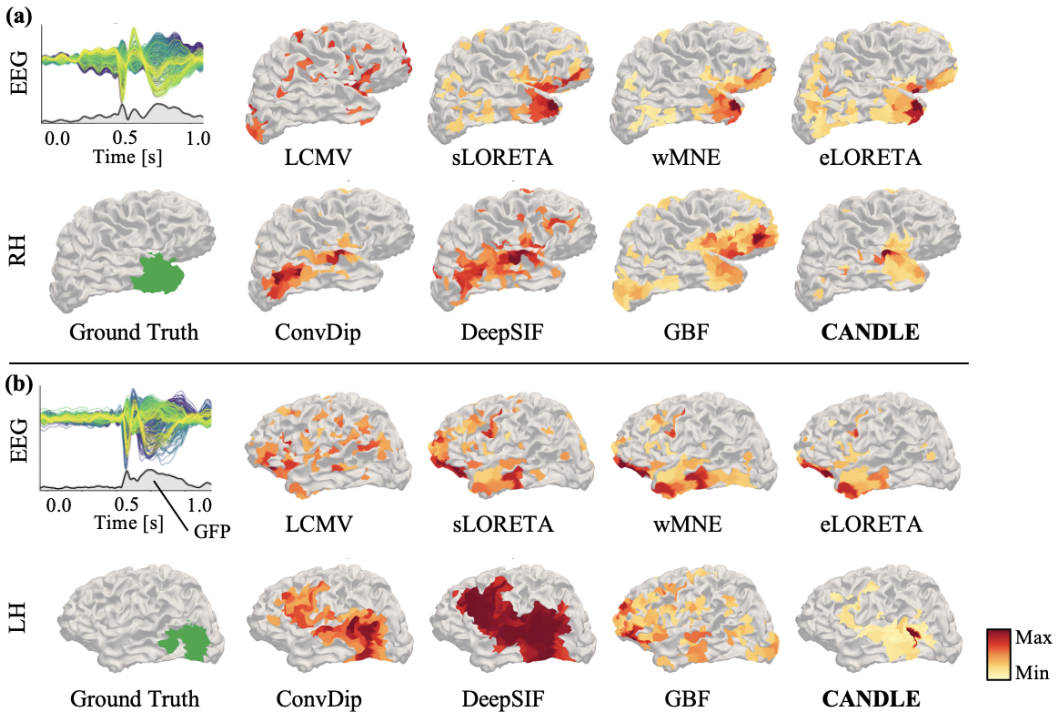

Figure 10: Additional qualitative results for epileptogenic zone estimation. Two representative patients with (a) a right-hemisphere (RH) and (b) a left-hemisphere (LH) resection. For each patient, the interictal EEG with its GFP is shown together with the temporally pooled source estimates of all eight ESI methods and the postsurgical resection cavity (green) used as the reference EZ.

Table 2: Full quantitative results for simulated source activity estimation. **Bold** and underlined values indicate the best and second-best performances, respectively. Statistical significance relative to CANDLE was assessed using two-sided Wilcoxon signed-rank tests (Holm corrected, † : $p < 0.05$, ‡ : $p < 0.01$, †† : $p < 0.001$). Corresponds to Figure 3(a).

Metric	LCMV	sLORETA	wMNE	eLORETA	ConvDip	DeepSIF	GBF	**CANDLE**
HD95 [mm] ↓	57.94†† ±9.07	66.08†† ±8.86	70.74†† ±8.58	69.71†† ±8.00	41.05†† ±7.35	45.12†† ±7.93	65.31†† ±9.27	**37.26** ±9.60
Dice score↑	0.22†† ±0.06	0.24†† ±0.05	0.21†† ±0.05	0.19†† ±0.05	0.33†† ±0.07	0.33†† ±0.08	0.26†† ±0.05	**0.48** ±0.07
F1@0.1 ↑	0.48†† ±0.15	0.56†† ±0.16	0.52†† ±0.15	0.47†† ±0.15	0.67†† ±0.14	0.62†† ±0.14	0.62†† ±0.16	**0.83** ±0.11
F1@0.2 ↑	0.31†† ±0.13	0.34†† ±0.14	0.26†† ±0.13	0.21†† ±0.13	0.48†† ±0.15	0.47†† ±0.16	0.37†† ±0.15	**0.71** ±0.12
F1@0.3 ↑	0.14†† ±0.11	0.12†† ±0.11	0.07†† ±0.08	0.05†† ±0.07	0.32†† ±0.13	0.33†† ±0.14	0.14†† ±0.10	**0.58** ±0.15
F1@0.4 ↑	0.03†† ±0.05	0.02†† ±0.04	0.01†† ±0.03	0.01†† ±0.02	0.19†† ±0.10	0.22†† ±0.13	0.03†† ±0.05	**0.43** ±0.15
F1@0.5 ↑	0.00†† ±0.01	0.00†† ±0.01	0.00†† ±0.00	0.00†† ±0.00	0.09†† ±0.08	0.12†† ±0.09	0.00†† ±0.02	**0.27** ±0.13
F1@0.6 ↑	0.00†† ±0.00	0.00†† ±0.00	0.00†† ±0.00	0.00†† ±0.00	0.04†† ±0.05	0.06†† ±0.07	0.00†† ±0.01	**0.15** ±0.10
F1@0.7 ↑	0.00†† ±0.00	0.00†† ±0.00	0.00†† ±0.00	0.00†† ±0.00	0.01†† ±0.03	0.02†† ±0.04	0.00†† ±0.00	**0.06** ±0.07
F1@0.8 ↑	0.00†† ±0.00	0.00†† ±0.00	0.00†† ±0.00	0.00†† ±0.00	0.00†† ±0.01	0.00†† ±0.02	0.00†† ±0.00	**0.02** ±0.04
F1@0.9 ↑	0.00‡ ±0.00	0.00‡ ±0.00	0.00‡ ±0.00	0.00‡ ±0.00	0.00‡ ±0.00	0.00‡ ±0.00	0.00‡ ±0.00	**0.01** ±0.02

Table 3: Full quantitative results for intracranial stimulation localization on the dataset of Parmigiani et al. (2022). Boldface, underlining, and significance markers follow the conventions in Table 2. Corresponds to Figure 4(b).

Metric	LCMV	sLORETA	wMNE	eLORETA	ConvDip	DeepSIF	GBF	**CANDLE**
Spatial dispersion [mm] ↓	69.71†† ±6.59	57.83 ±6.52	61.38†† ±7.04	62.62†† ±6.72	59.05†† ±7.11	64.95†† ±6.43	56.60 ±7.07	**54.11** ±7.09
Peak distance [mm] ↓								
Overall	58.18†† ±12.64	37.47 ±13.66	48.11†† ±17.90	50.09†† ±13.44	41.12‡ ±12.28	46.02†† ±14.39	36.47 ±10.98	**32.86** ±10.83
0–10 mm	55.65†† ±16.96	32.62 ±19.49	40.84 ±21.89	49.40† ±22.32	38.82 ±16.89	46.95†† ±19.47	**29.20** ±16.83	32.48 ±12.32
10–20 mm	55.52†† ±19.21	34.83 ±15.19	49.44† ±25.81	56.27†† ±21.66	45.59† ±20.76	46.10† ±19.09	36.39 ±13.20	**31.35** ±15.36
20–30 mm	61.25†† ±19.67	41.53 ±16.05	54.68† ±26.18	44.66 ±18.45	40.82 ±18.16	44.73 ±20.10	46.88 ±17.40	**36.04** ±11.80
30–45 mm	57.01 ±9.64	46.48 ±16.33	54.48 ±20.64	62.44† ±19.92	49.92 ±12.18	54.45 ±12.76	51.05‡ ±16.28	**38.04** ±14.56

Table 4: Full quantitative results for epileptogenic zone estimation on the dataset of (Vorderwülbecke et al., 2025). Boldface, underlining, and significance markers follow the conventions in Table 2. Corresponds to Figure 5(b).

Metric	LCMV	sLORETA	wMNE	eLORETA	ConvDip	DeepSIF	GBF	**CANDLE**
HD95 [mm] ↓	72.82‡ ±10.43	66.79 ±16.12	62.57 ±18.20	63.93 ±16.13	67.80 ±11.08	65.37 ±12.33	71.08† ±15.10	**61.72** ±12.88
Centroid distance [mm] ↓	48.25‡ ±6.69	38.82 ±12.81	40.56 ±15.27	39.89 ±15.61	40.14 ±13.99	39.07 ±12.36	40.29 ±13.67	**35.38** ±13.98