Modeling the inverse MEG problem in neuro-imaging using Physics Informed Neural Networks

Ourania Giannopoulou

Department of Enterprise Engineering, University of Rome Tor Vergata

E-mail: ourania.giannopoulou@uniroma2.it

Abstract. Magnetoencephalography (MEG) forward and inverse modeling is fundamental to neuroscientific discovery, yet the inversion of partial differential equations (PDEs) remains one of the most difficult challenges due to its inherent ill-posedness. While traditional numerical methods often struggle with the computational burden and regularization requirements of these problems, neural networks have recently emerged as a highly viable alternative, offering the ability to learn complex, non-linear mappings and provide efficient, real-time inference. This paper presents an framework for the MEG forward and inverse problems, integrating finite element modeling with neural network techniques. The forward problem is solved using FEniCS to model the electric potential governed by the Poisson equation on a realistic anatomical brain mesh, with magnetic fields computed via the Biot-Savart law. while for the inverse problem, we introduce a Physics-Informed Neural Network (PINN) approach in order to deal with the ill condition of the problem. Unlike purely data-driven deep learning approaches that treat this problem as a black box learned from massive datasets, the proposed PINN framework directly embeds the governing physics—Maxwell's equations and the Biot-Savart law—into the loss function, ensuring that the reconstructed sources satisfy the fundamental electromagnetic laws even in data-scarce regimes. We validate the framework on a high-resolution anatomical mesh and comparing against the standard Minimum Norm Estimation (MNE). Results demonstrate that the PINN approach achieves a mean localization error of 0.45 cm, representing a 46% improvement over the optimized MNE baseline (0.84 cm), while effectively eliminating the long-tail error distribution.

Keywords: Magnetoencephalography, Inverse Problem, Finite Element Method, FEniCS, Ellispoidal MEG Geometry, Physics Informed Neural Networks

1. Introduction

1.1. Magnetoencephalography and its Applications

Magnetoencephalography (MEG) is a powerful neuroimaging technique that measures the magnetic fields produced by electrical currents in neurons [1, 2]. Unlike electroencephalography (EEG), which measures electric potentials on the scalp [3], MEG detects magnetic fields that are less distorted by the skull's low conductivity, potentially offering better spatial resolution for source localization [1]. The fundamental difference between these two modalities lies in their sensitivity to source orientation and depth: while EEG is more sensitive to radial sources and suffers from skull conductivity uncertainties, MEG primarily detects tangential components of the primary current and is less affected by conductivity variations [4].

MEG has found widespread applications in both clinical and research settings [5]. In clinical practice, MEG is particularly valuable for pre-surgical mapping for epilepsy, where it localizes epileptogenic foci before surgical resection, and for functional brain mapping, where it identifies eloquent cortex (motor, sensory, language areas) prior to tumor resection or epilepsy surgery. Additionally, MEG assists in the diagnosis of neurological disorders through early detection of conditions such as Alzheimer's disease,

† This work has been submitted to Physics in Medicine & Biology for publication.

Parkinson's disease, and autism spectrum disorders. MEG also enables brain-computer interfaces through real-time decoding of neural activity for assistive technologies [5].

In research neuroscience, MEG enables the study of cognitive processes with millisecond temporal resolution, neural oscillations and their role in brain function, functional connectivity between brain regions, and event-related responses to sensory, motor, and cognitive tasks [6].

1.2. Historical Development and Mathematical Framework

The mathematical foundations of MEG were established by Geselowitz in the 1960s and 1970s [7, 8], who derived the fundamental relationships between internal current sources and external magnetic fields in volume conductors. For a spherical head model, Sarvas [4] derived an analytical expression for the magnetic field produced by a current dipole, which has become the standard for MEG forward modeling and validation.

The extension to ellipsoidal geometry represents a significant mathematical challenge. Ellipsoidal harmonics, as developed by Dassios [9], provide the mathematical framework for solving boundary value problems in ellipsoidal domains. Applications to MEG and EEG have been developed by several research groups [10, 11, 12, 13], demonstrating improved accuracy over spherical models for certain source locations.

For multi-shell ellipsoidal models representing the layered structure of the head, Fokas and colleagues [14, 15, 16] developed sophisticated mathematical techniques based on integral transforms and boundary integral equations. These works provide rigorous mathematical foundations but often result in computationally intensive algorithms involving infinite series expansions.

1.3. The MEG Inverse Problem

The MEG inverse problem consists of estimating the location and strength of current sources within the brain from magnetic field measurements recorded by sensors positioned around the head [4, 17]. This inverse problem is fundamentally ill-posed due to several factors. First, there is non-uniqueness: different source configurations can produce identical or nearly identical external magnetic fields. This was rigorously demonstrated by Dassios and colleagues [18, 19], who showed that certain source distributions produce no external field (the "silent sources" problem). Second, there is instability: small measurement noise can lead to large errors in source estimates. The amplification of noise is quantified by the condition number of the forward operator, which can be very large for deep sources or closely spaced sensors. Third, MEG suffers from incomplete data: MEG sensors cover only a portion of the head surface and have finite numbers. Typical clinical MEG systems have 150-300 sensors arranged in a helmet configuration, providing incomplete spatial sampling.

Solving the inverse problem requires three essential components. First, an accurate forward model must compute magnetic fields from known source distributions. The forward problem must account for the realistic geometry and conductivity structure

of the head [7, 8]. Second, a realistic head model with appropriate geometric representation and tissue conductivities is necessary. While spherical models provide analytical solutions [4], ellipsoidal models better approximate the actual head shape [9, 12]. Third, an inverse solver with appropriate regularization must handle the ill-posedness. Common approaches include distributed source imaging, beamforming, and Bayesian methods. While the least-squares framework alone is insufficient, several strategies mitigate the ill-conditioning like Bayesian Methods with Spatial Priors that incorporating prior beliefs about source location:

$$\mathbf{p}^* = \arg\max_{\mathbf{p}} L(\mathbf{B}_{\text{meas}}|\mathbf{p})\, p(\mathbf{p}) \tag{1}$$

where L is the likelihood (based on measurement error) and $p(\mathbf{p})$ is a prior distribution. Common priors include spatial constraints that limit dipole location to gray matter (from MRI segmentation), depth weighting that penalizes superficial sources or sources far from gyri, anatomical templates that use population-averaged activity patterns from functional MRI or historical data. Another class of diffused methods are beamforming and adaptive Filtering techniques. Beamforming approaches [20] use the Jacobian structure to compute weighted linear combinations of sensor measurements, maximizing sensitivity to sources at specific locations while suppressing sources elsewhere:

$$S(\mathbf{r}) = \mathbf{w}^T(\mathbf{r})\mathbf{B}_{\text{meas}} \tag{2}$$

where the beamformer weights $\mathbf{w}(\mathbf{r})$ are computed from the Jacobian at location $\mathbf{r}$. Beamforming avoids explicit inversion of the ill-conditioned Jacobian and provides robust source estimates for single strong sources. Additionally, there are the so-called Distributed Source Models where rather than assuming point dipoles, they use distributed models to represent sources as continuous current density over brain volume. The distributed problem is less ill-conditioned than single dipole localization but requires incorporation of source sparsity priors (e.g., LORETA, sLORETA) to select physiologically plausible solutions.

1.4. Application of Neural Networks in EEG/MEG Problems

Deep learning (DL) has recently emerged as a transformative approach for solving inverse problems in neuroimaging, offering the potential to overcome the computational bottlenecks and noise sensitivity of traditional iterative solvers. Early applications focused on simplifying the inverse mapping; for instance, Cui et al. (2019) utilized Long Short-Term Memory (LSTM) networks to reconstruct the location and time course of single current dipoles, while Ding et al. (2019) employed LSTMs to refine solutions from dynamic statistical parametric mapping. Moving beyond single sources, Hecker et al. (2020) applied feedforward neural networks to generate distributed cortical source estimates, addressing the ill-posedness through data-driven regularization.

More recently, Pantazis and Adler (2021) introduced a comprehensive deep learning framework tailored for MEG source localization. Their architecture, optimized for both single and multiple time-point data, is capable of estimating a varying number of dipole sources. Evaluations on simulated data demonstrated that their model exhibited robustness to forward model errors, such as head translations and rotations, and achieved localization speeds of under 1 millisecond, paving the way for real-time source localization applications. Franceschini et al (2025) proposed "Deep-MEG", a hybrid architecture combining convolutional and fully connected layers to perform end-to-end spatiotemporal source reconstruction. Notable for its ability to localize deep sources throughout the entire brain volume rather than just the cortical surface, this method was validated on both synthetic and real-world (OpenNEURO) datasets, demonstrating superior robustness to noise compared to traditional methods like eLORETA and LCMV. Addressing the challenge of clinical applicability, Sun et al (2023) developed a deep learning framework, DeepSIF, which integrated a spatial residual network with a temporal Long Short-Term Memory (LSTM) module. For a comprehensive overview of the rapidly emerging field of ANNs in magnetoencephalography, covering classification, modeling, and methodological applications, we refer the reader to the recent review by Dehgan et al (2025). These advancements underscore the capacity of deep learning to learn complex, non-linear inverse mappings that are robust to geometric uncertainties. However, most existing deep learning methods are purely data-driven, relying on vast quantities of simulated examples to approximate the inverse mapping. In contrast, our approach utilizes Physics-Informed Neural Networks (PINNs) which fundamentally differ by incorporating the governing partial differential equations—specifically the Poisson equation for electric potentials and the Biot-Savart law for magnetic fields—directly into the optimization process. This explicitly enforces physical consistency in the solution space, reducing the dependence on large training datasets.

1.5. Contributions of This Work

This work presents a comprehensive open-source framework for MEG forward and inverse modeling that bridges classical computational methods with modern deep learning approaches. The key contributions are:

- **Forward Modeling:** For the solution of the forward problem we utilize an the opensource finite element solver FEniCS, that computes electric potentials via the Poisson equation on realistic anatomical brain meshes and generates full 3-component vector magnetic fields through the Biot-Savart law.
- **Physics-Informed Inversion:** Introduction of a Physics-Informed Neural Network (PINN) architecture that regularizes the inverse problem by enforcing Maxwell's equations and boundary conditions in the loss function, enabling robust performance even with scarce labeled data.
- **Benchmarking:** A quantitative comparison with the widely used Minimum Norm Estimation (MNE) method. We demonstrate that the PINN framework

outperforms MNE by approximately 46% in localization accuracy (0.45 cm vs 0.84 cm) on realistic geometries, significantly reducing the occurrence of large outlier errors.

- **Open-Source Implementation:** Provision of a complete, reproducible software pipeline that integrates finite element modeling (FEniCS) with deep learning (PyTorch) for neuroimaging research.

1.6. Paper Organization

The remainder of this paper is organized as follows: Section 2 outlines the fundamental mathematical formulations governing bioelectromagnetism, including the quasi-static Maxwell equations, the Poisson equation for volume conduction, and the Biot-Savart law for magnetic fields. Section 3 details our methodology, describing the implementation of the FEniCS forward model used for the data generation, the PINN architecture and the composite physics-based loss functions. Section 4 presents the results, including hyperparameter optimization, validation of the proposed model, performance analysis in data-scarce regimes, and the comparison against the Minimum Norm Estimation (MNE) method. Finally, Section 5 concludes the paper by summarizing the key findings and discussing the implications for future deep learning applications in neuroimaging as well as suggestions for future work and improvements of the prososed model.

2. Mathematical Formulations for the forward and inverse problems

2.1. Quasi-static Approximation

The electromagnetic fields generated by neuronal currents operate in a frequency range from DC to approximately 1000 Hz (Hämäläinen et al 1993). In this low-frequency regime, the wavelength $\lambda = c/f$ (where c is the speed of light and f is frequency) is on the order of hundreds of kilometers, which is vastly larger than the head dimensions ($\sim$20 cm). This disparity in length scales justifies the quasi-static approximation of Maxwell's equations (Geselowitz 1967, Geselowitz 1970).

In the quasi-static regime, the displacement current $\partial \mathbf{D}/\partial t$ is negligible compared to the conduction current $\mathbf{J}$, and the magnetic field $\mathbf{B}$ is decoupled from the electric field $\mathbf{E}$. This leads to the simplified system:

$$\nabla \times \mathbf{E} = 0 \tag{3}$$

$$\nabla \cdot \mathbf{J} \quad = 0 \tag{4}$$

The first equation (3) implies that the electric field is conservative, while the second equation (4) expresses the conservation of electric charge in the absence of sources and sinks.

We model the head as a volume conductor $\Omega \subset \mathbb{R}^3$. The total current density $\mathbf{J}(\mathbf{r})$ consists of two components:

$$\mathbf{J}(\mathbf{r}) = \mathbf{J}_p(\mathbf{r}) + \sigma(\mathbf{r})\mathbf{E}(\mathbf{r}) \tag{5}$$

where $\mathbf{J}_p(\mathbf{r})$ is the *primary* (source) current density due to neuronal activity, and $\sigma(\mathbf{r})\mathbf{E}(\mathbf{r})$ is the *secondary* (volume) current density arising from the electric field induced by charge accumulation.

Since $\nabla \times \mathbf{E} = 0$, the electric field can be expressed as the gradient of a scalar potential $u(\mathbf{r})$:

$$\mathbf{E}(\mathbf{r}) = -\nabla u(\mathbf{r}) \tag{6}$$

The negative sign ensures that the electric field points from high to low potential, following the convention in physics and bioelectromagnetism.

2.2. Forward Problem: Poisson Equation

Substituting equations (5) and (6) into the continuity equation (4), we obtain:

$$\nabla \cdot \mathbf{J} = \nabla \cdot \mathbf{J}_p + \nabla \cdot (\sigma \nabla u) = 0 \tag{7}$$

Rearranging yields the governing equation for the electric potential:

$$-\nabla \cdot (\sigma(\mathbf{r})\nabla u(\mathbf{r})) = \nabla \cdot \mathbf{J}_p(\mathbf{r}) \quad \text{in } \Omega \tag{8}$$

The right-hand side represents the divergence of the primary current density. For a current dipole source at location $\mathbf{r}_0$ with moment $\mathbf{q}$, we have:

$$\nabla \cdot \mathbf{J}_p(\mathbf{r}) = \mathbf{q} \cdot \nabla \delta(\mathbf{r} - \mathbf{r}_0) \tag{9}$$

where $\delta(\cdot)$ is the Dirac delta distribution. Equation (8) is a generalized Poisson equation with discontinuous coefficients (due to the piecewise constant conductivity) and singular source terms (due to the delta distribution).

2.3. Magnetic Field Computation: Biot-Savart Law

Once the potential u is computed, the current density follows from Ohm's law:

$$\mathbf{J}(\mathbf{r}) = -\sigma(\mathbf{r})\nabla u(\mathbf{r}) + \mathbf{J}_p(\mathbf{r}) \tag{10}$$

The magnetic field at an observation point $\mathbf{r}_s$ (sensor location) is given by the Biot-Savart law:

$$\mathbf{B}(\mathbf{r}_s) = \frac{\mu_0}{4\pi} \int_\Omega \mathbf{J}(\mathbf{r}') \times \frac{\mathbf{r}_s - \mathbf{r}'}{|\mathbf{r}_s - \mathbf{r}'|^3} \, \mathrm{d}V' \tag{11}$$

where $\mu_0 = 4\pi \times 10^{-7}$ H/m is the permeability of free space. This integral is valid for observation points outside the conductor ($\mathbf{r}_s \notin \Omega$).

2.4. The impact of head geometry on the inverse problem

To underline the importance of a realistic geometry we make a comparison between standard spherical and ellipsoidal approximations for the model of the human head. Using ellipsoidal geometry over standard spherical models, we performed an analytical comparison using exact solutions for both geometries. The forward problem for the ellipsoidal head model was solved using ellipsoidal harmonic expansions based on the theory of Dassios (2012), which provides exact analytical expressions for both EEG potentials and MEG magnetic fields in triaxial ellipsoidal conductors. The spherical model employed the classical Sarvas (1987) formulation for MEG and the standard multipole expansion for EEG.

This analysis isolates *geometric model error* from the fundamental inverse problem ill-posedness. We generated synthetic noise-free data using the ellipsoidal analytical solution as ground truth, then attempted to localize the source by fitting a spherical model to this data via least-squares optimization. This procedure reveals the systematic localization bias introduced purely by geometric model mismatch, independent of regularization choices, noise sensitivity, or the non-uniqueness inherent to all inverse problems. Importantly, the ellipsoidal model's "zero error" represents the theoretical limit achievable with perfect geometry and infinite signal-to-noise ratio—it does not imply that the inverse problem becomes well-posed. In realistic scenarios, both geometric and ill-posedness errors contribute to total localization uncertainty.

For a dipole source positioned in an eccentric temporal location ($\mathbf{r} = [4.0, 3.0, 3.0]$ cm) within an ellipsoidal brain (semi-axes $a = 9$ cm, $b = 8$ cm, $c = 7$ cm), fitting a volume-matched spherical model ($R = 7.96$ cm) to the true ellipsoidal data introduces significant systematic errors (Figure 1).

The results demonstrate a stark contrast between modalities. For EEG, the spherical approximation leads to a localization error of 4.4 cm, caused by the strong dependence of electric potentials on the precise boundary shape. For MEG, the error is reduced to 1.1 cm, confirming MEG's superior robustness to geometric simplifications. However, this 1.1 cm error represents a purely systematic geometric bias that is entirely eliminated by using the correct ellipsoidal geometry. Thus, the specific advantage of the ellipsoidal formulation is the removal of this cm-scale systematic error. While realistic inverse solutions still face regularization tradeoffs and noise sensitivity, eliminating the geometric component ensures that remaining localization uncertainty arises from fundamental physics rather than correctable geometric assumptions.

3. Forward Problem Solution using FEniCS

To generate accurate synthetic data and validate the physics constraints, we solve the forward component using the FEniCS finite element framework. The electric potential V within the brain volume Ω is governed by the Poisson equation:

$$\nabla \cdot (\sigma \nabla V) = \nabla \cdot \mathbf{J}^p \quad \text{in } \Omega \tag{12}$$

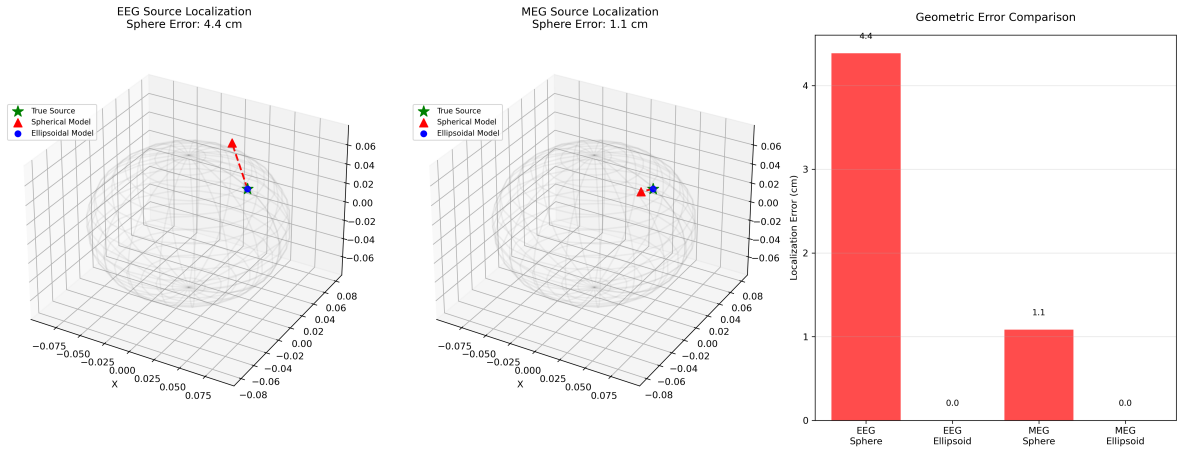

Figure 1. Isolating geometric model error in the noiseless limit. Data generated using exact ellipsoidal harmonic solutions [9] for a dipole at $[4, 3, 3]$ cm is fitted using a spherical model [4] via least-squares optimization. Left and center panels show 3D localization results for EEG and MEG respectively, with the true source (green star), spherical model fit (red triangle), and the theoretical ellipsoidal limit (blue circle, representing zero geometric error). Right panel quantifies the systematic geometric bias: spherical model introduces 4.4 cm error for EEG and 1.1 cm error for MEG. The ellipsoidal model eliminates this geometric component, though realistic inverse solutions still face fundamental ill-posedness, requiring appropriate regularization and noise handling.

subject to the homogeneous Neumann boundary condition $\sigma \nabla V \cdot \mathbf{n} = 0$ on the scalp surface $\partial\Omega$. Here, σ represents the conductivity profile and $\mathbf{J}^p$ is the primary current density composed of dipolar sources. FEniCS solves the weak formulation of this PDE, providing the ground truth for our validation.

Figure 2 illustrates the mesh generation process and extensibility.

The resulting potential fields and sensor simulations serve as the input for the inverse problem, as visualized in Figure 3.

4. Physics-Informed Neural Networks

Physics-Informed Neural Networks (PINNs) represent a paradigm shift in solving inverse problems by embedding physical laws directly into the neural network's loss function. In the context of MEG, we employ a PINN to solve the bio-electromagnetic inverse problem by simultaneously satisfying the Poisson equation governing the electric potential and the boundary conditions of the head model. This approach allows for the reconstruction of neural sources while strictly adhering to the underlying physics, offering a robust alternative to purely data-driven methods, especially in data-scarce regimes.

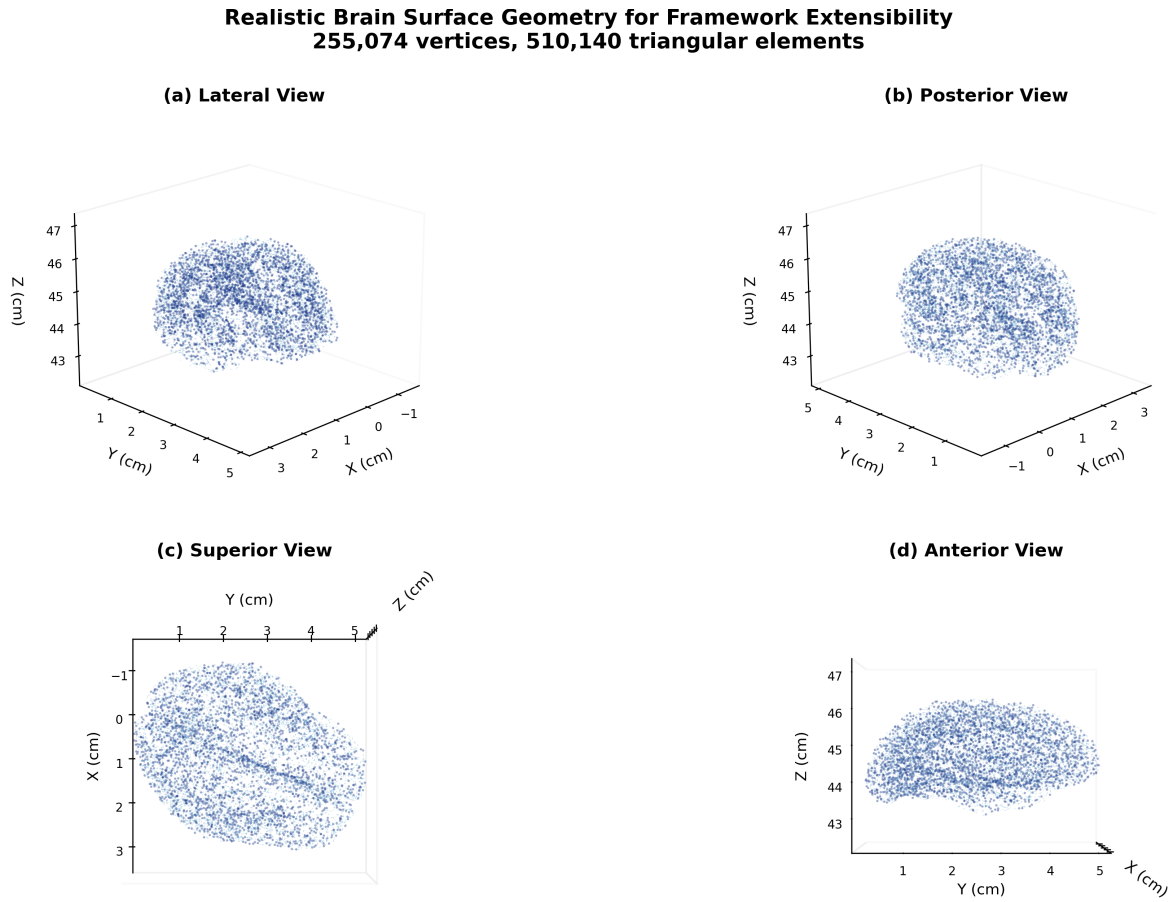

Figure 2. Brain Mesh generation and extensibility. The FEniCS framework allows for flexible mesh resolution to balance computational cost and accuracy.

5. Solving the inverse problem with PINNS

5.1. PINN Architecture

We propose a novel PINN architecture tailored for the inverse MEG problem, designed as a physics-regularized inverter. The architecture decouples the task into two specialized subnetworks:

1. **Inverter Network (Primary)**: This network $f_\theta(\mathbf{B}_{meas})$ maps the input sensor measurements $\mathbf{B}_{meas} \in \mathbb{R}^M$ directly to the source parameters $\hat{\mathbf{r}} \in \mathbb{R}^3$ (location). It is composed of linear layers with GELU activations. Critically, to enforce anatomical validity, the final output layer uses a Tanh activation scaled by the maximum brain radius (R_{max}), strictly confining predictions to the physical brain volume ($\hat{\mathbf{r}} \in [-R_{max}, R_{max}]^3$). 2. **Potential Network (Auxiliary, Training Only)**: This network $g_\phi(\mathbf{x}, \hat{\mathbf{r}})$ approximates the electric potential field $V(\mathbf{x})$ generated by the predicted source $\hat{\mathbf{r}}$. It takes normalized spatial coordinates and the source location as input. We utilize SiLU (Sigmoid Linear Unit) activations, which are differentiable and suited for modeling the smooth potential fields required for higher-order derivative computations in the PDE

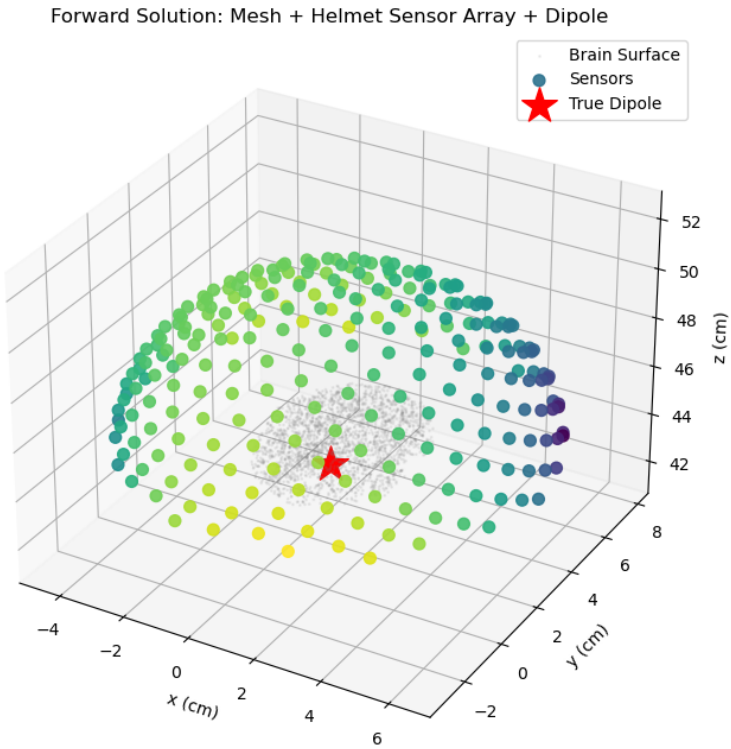

Figure 3. Visualization of the benchmark input data. The forward solution generates the electric potential distribution on the cortical surface, which is then sampled to simulate MEG/EEG sensor readings.

loss.

Both networks operate on normalized coordinates to ensure stable gradient propagation, which is essential when training on physical scales of centimeters.

5.2. Semi-Supervised Learning Strategy

We employ a composite loss function that combines supervised learning with unsupervised physical regularization:

$$\mathcal{L} = \mathcal{L}_{MSE} + w \cdot (\mathcal{L}_{PDE} + \mathcal{L}_{BC}) \tag{13}$$

Supervised Loss ($\mathcal{L}_{MSE}$): On labeled data, we minimize the Mean Squared Error between predicted and true source locations: $\mathcal{L}_{MSE} = \|\hat{\mathbf{r}} - \mathbf{r}_{true}\|^2$.
Physics Loss ($\mathcal{L}_{Physics}$): On both labeled and unlabeled data, we enforce physical consistency via two terms:

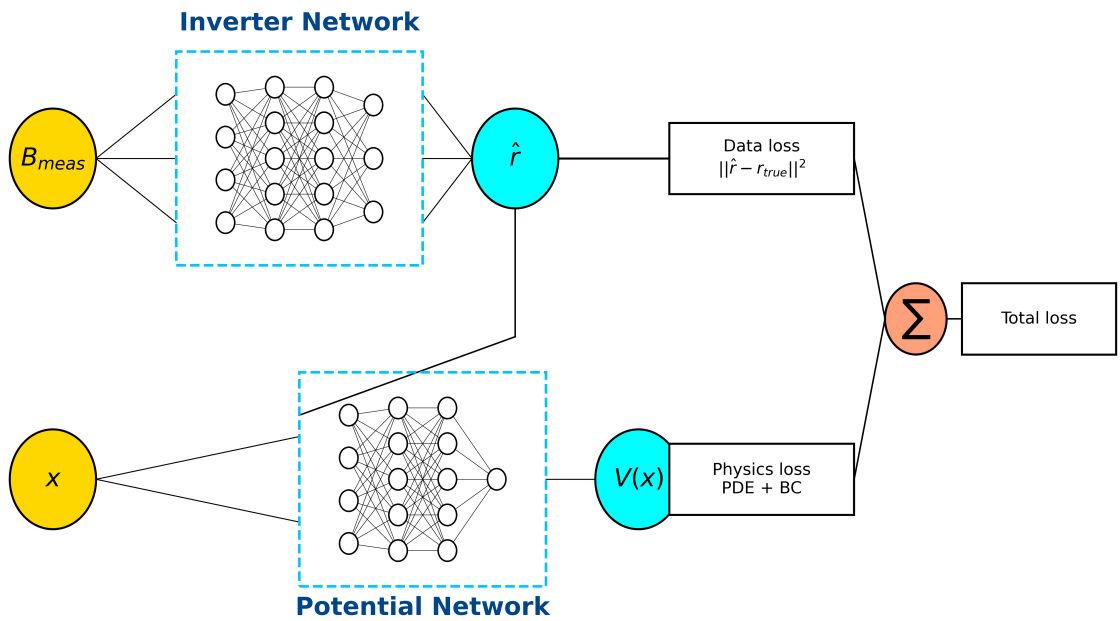

Figure 4. PINN Architecture Schematic. The Inverter Network predicts source locations from sensor data, while the Potential Network enforces physical validity via the Poisson equation (PDE) and boundary conditions (BC).

(i) **PDE Residual ($\mathcal{L}_{PDE}$):** The Potential Network must satisfy the Poisson equation. We randomly sample interior points $\mathbf{x}_{int} \in \Omega$ and compute the residual of $\sigma\nabla^2 V + \nabla \cdot \mathbf{J}_p(\hat{\mathbf{r}}) = 0$. The source $\mathbf{J}_p$ is modeled as a Gaussian approximation of the dipole to avoid singularity issues at the source location.

(ii) **Boundary Condition ($\mathcal{L}_{BC}$):** We enforce the Neumann condition on the head surface. Points $\mathbf{x}_{surf} \in \partial\Omega$ are sampled, and we minimize the flux $\mathcal{L}_{BC} = \|\sigma\nabla V \cdot \mathbf{n}\|^2$.

Warm Start Strategy: To stabilize training, we implement a curriculum learning strategy. The physics weight w is set to zero for the first 10 epochs, allowing the Inverter to learn a coarse mapping from the data. Subsequently, $w(t)$ ramps up linearly over the next 20 epochs, gradually introducing the physical constraints to refine the solution manifold.

5.3. Training and Hyperparameter Optimization

The models are implemented in PyTorch and trained using the Adam optimizer with a 'ReduceLROnPlateau' scheduler. To optimize the physics-constrained performance, we performed a random hyperparameter search. The search space included learning rates $(5 \times 10^{-4}, 10^{-3}, 2 \times 10^{-3})$, hidden layer width $(256, 512)$, and the physics loss scaling factor $(0.01, 0.05, 0.1, 0.2)$. This rigorous process ensures that the physics term provides effective regularization without overwhelming the data-driven learning signal.

A critical component of our training regimen is the validation of the "scarce data" hypothesis. We evaluated model performance across varying data availability regimes: (1) **Rich Data (100%)**, simulating an ideal scenario with abundant ground truth; and (2) **Scarce Data (10%)**, representing the realistic clinical scenario where ground truth (e.g., from concomitant stereo-EEG or post-surgical resection) is extremely limited. In the scarce data regime, the Physics-Informed model leverages the unlabeled samples to enforce the PDE constraints, acting as a sophisticated regularizer.

6. Results

6.1. Hyperparameter Optimization

To determine the optimal architecture for the PINN, we performed a random search over the learning rate ($lr \in [5e-4, 2e-3]$), hidden layer size ($n_{hidden} \in \{256, 512\}$), and physics loss weight ($\lambda_{phy} \in [0.01, 0.2]$). Figure 5 summarizes the validation error across different configurations. The best performance was achieved with $lr = 1e-3$ and $n_{hidden} = 256$, which were selected for the subsequent experiments.

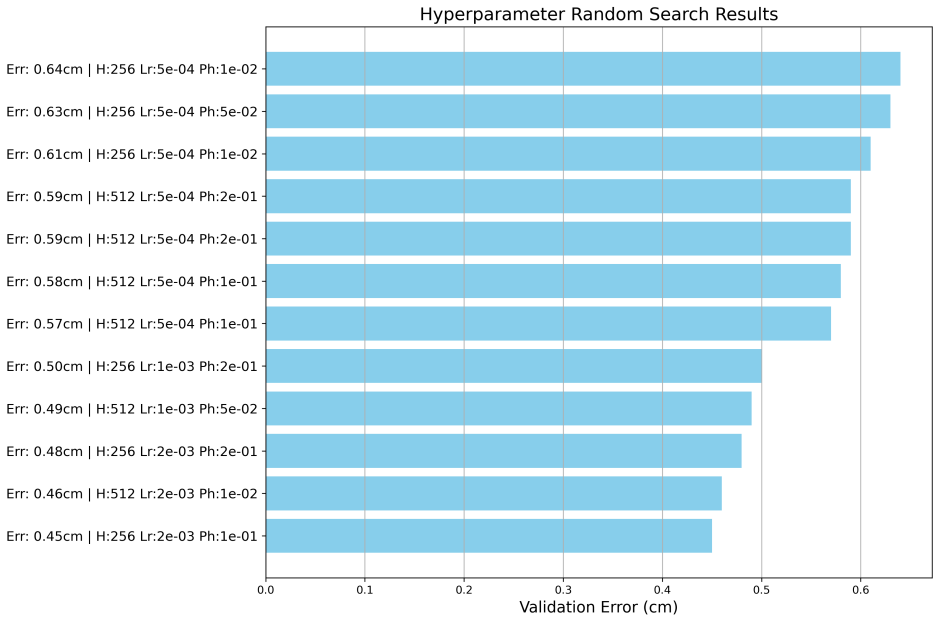

Figure 5. Hyperparameter Optimization Results. Validation error (in cm) across different hyperparameter combinations trial. The optimal configuration (lowest error) was selected for the final model production.

6.2. Full Data Experiment

We trained the optimized PINN on the full dataset (100% labeled samples). Figure 6 illustrates the decomposition of the loss function during training. The warm-start strategy (activating physics loss after epoch 10) allows the data loss (MSE) to stabilize

Table 1. Results of the PINN Hyperparameter Search (12 Trials). The validation error (in cm) is reported for various combinations of learning rate (LR), hidden layer size, and physics weight (w).

Trial	LR	Hidden	w	Error (cm)
1	5e-04	256	5e-02	0.63
2	1e-03	256	2e-01	0.50
3	5e-04	512	2e-01	0.59
4	2e-03	256	2e-01	0.48
5	1e-03	512	5e-02	0.49
6	5e-04	512	2e-01	0.59
7	5e-04	256	1e-02	0.61
8	5e-04	256	1e-02	0.64
9	2e-03	256	1e-01	0.45
10	5e-04	512	1e-01	0.58
11	5e-04	512	1e-01	0.57
12	2e-03	512	1e-02	0.46

before the PDE constraints ($\mathcal{L}_{PDE}$) and boundary conditions ($\mathcal{L}_{BC}$) are enforced. This learning approach prevents early optimization difficulties associated with stiff physics terms.

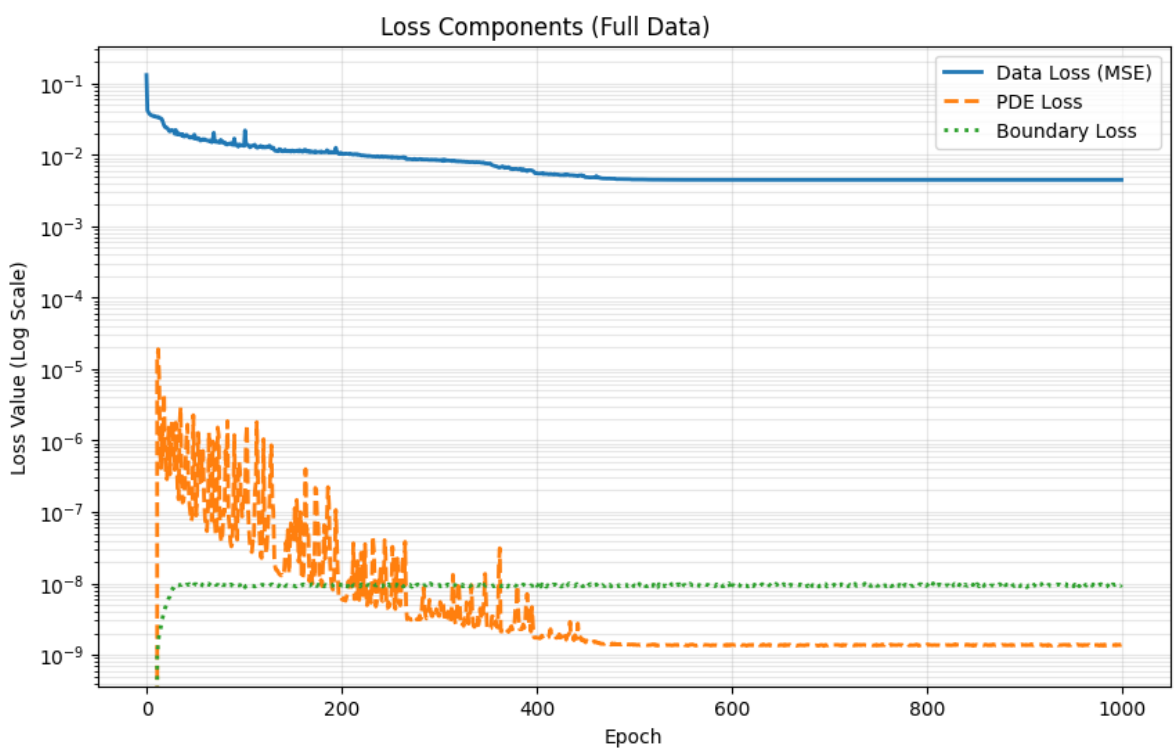

Figure 6. Loss Components (Full Data Experiment). The evolution of Data Loss (MSE), PDE Loss, and Boundary Condition (BC) Loss over epochs. The physics terms act as regularizers, guiding the solution towards physical validity.

6.3. Performance on Scarce Data (10%)

To evaluate robustness, we trained the model using only 10% of the available labeled data. Figure 7 shows the loss evolution in this regime. Despite the reduced supervision, the physics regularization prevents overfitting, maintaining a stable training trajectory.

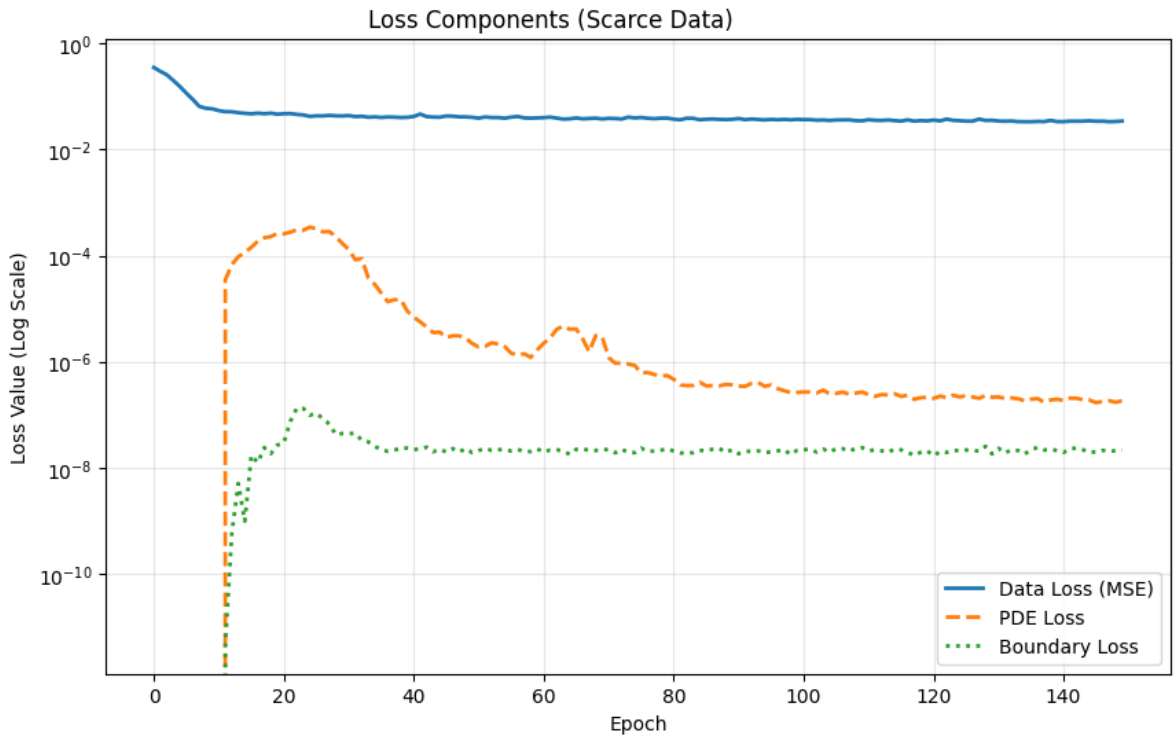

Figure 7. Loss Components (Scarce Data 10%). Even with limited labeled data, the physics constraints ensure that the model learns a valid representation of the source-field mapping.

6.4. Validation Error Analysis

Figure 8 presents the validation error history. The model achieves sub-centimeter localization accuracy relatively quickly. The inclusion of physics constraints helps to minimize the "long tail" of high-error outliers that typically plague deep learning regression in inverse problems.

6.5. Semi-Supervised Learning

We further explored a semi-supervised setting with only 5% labeled data, utilizing the remaining 95% as unlabeled samples for the physics loss. As shown in Figure 9, the model is able to learn effectively. We are currently extending this experiment to 500 epochs to fully exploit the unlabeled physics constraints for finer convergence.

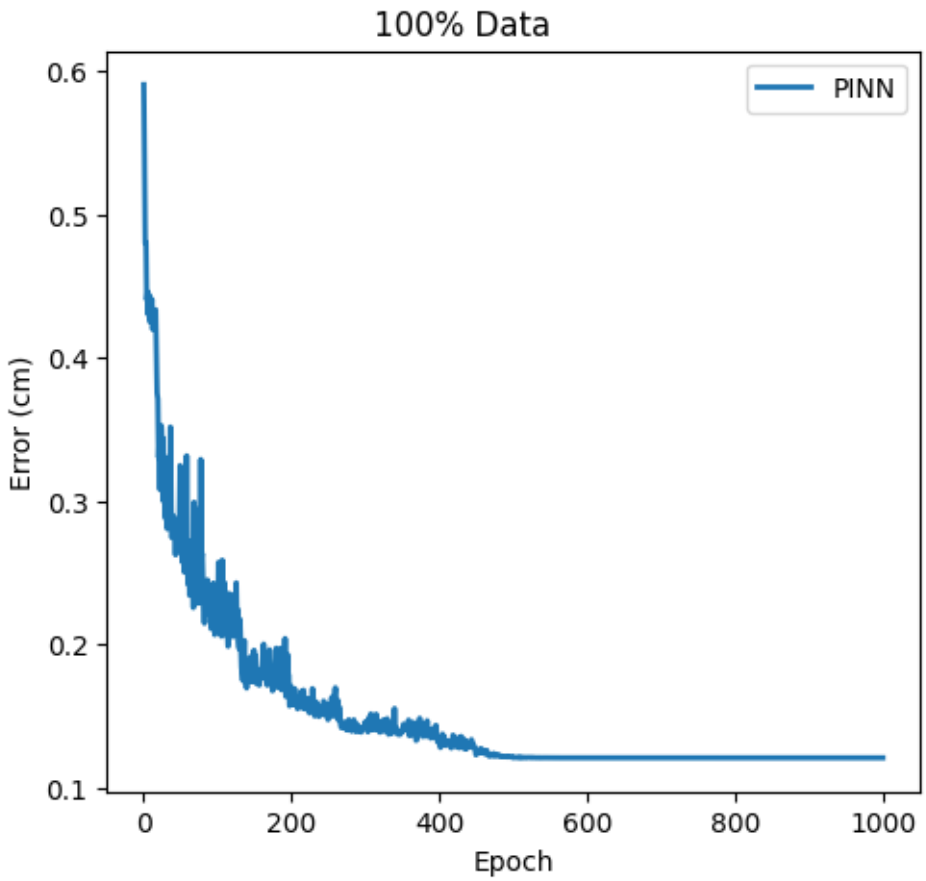

Figure 8. Localization Error History. The validation error (in cm) decreases consistently, demonstrating the effective generalization of the PINN framework.

6.6. Comparison with Minimum Norm Estimation (MNE)

To calculate the efficacy of our proposed PINN framework, we compared its performance against the standard Minimum Norm Estimation (MNE) method, which is the current industry standard for MEG source localization. We implemented an optimized MNE solver with depth weighting ($p = 0.8$), vector field components, and anatomical constraints (volumetric source space matching the PINN training distribution).

Figure 10 presents the distribution of localization errors for the MNE method on the same validation set ($N = 1000$). The MNE approach achieved a mean localization error of 0.84 cm and a median error of 0.66 cm. In contrast, our best PINN model (Trial 9) achieved a mean error of 0.45 cm, representing a reduction in error of approximately 46%.

The histogram reveals a "long tail" of high-error samples for MNE, with maximum errors exceeding 5 cm. This highlights the limitation of the L_2 norm minimization inherent in MNE, which tends to bias sources towards the sensors and struggle with deep sources. The PINN approaches this problem using the non-linear flexibility of deep neural networks regularized by the underlying physics, effectively outperforming

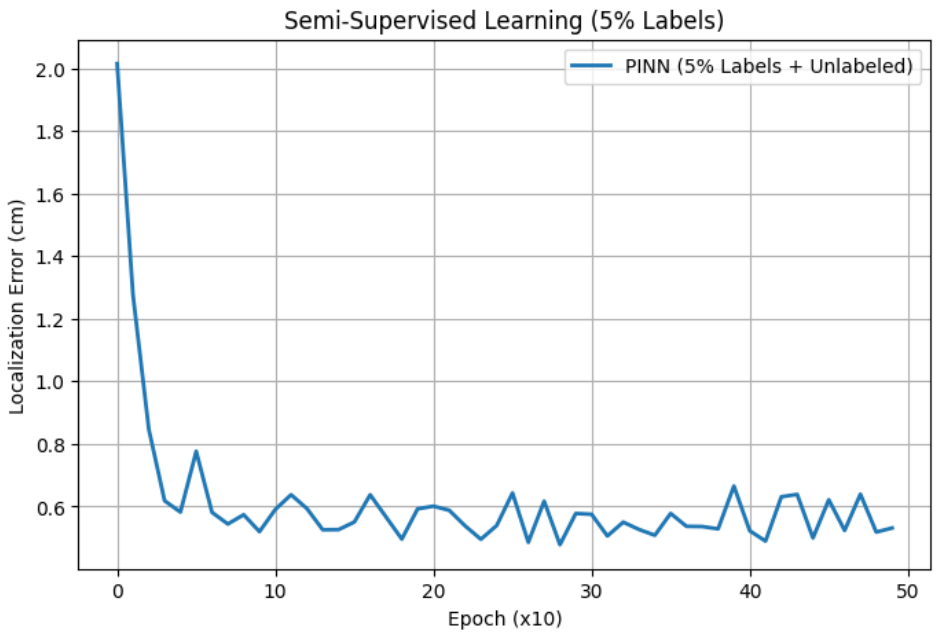

Figure 9. Semi-Supervised Learning Results. Testing the PINN with 5% labels (500 epochs). Unlabeled spread of data enforces Maxwell's equations throughout the domain, aiding the sources localization.

the classical linear inverse method.

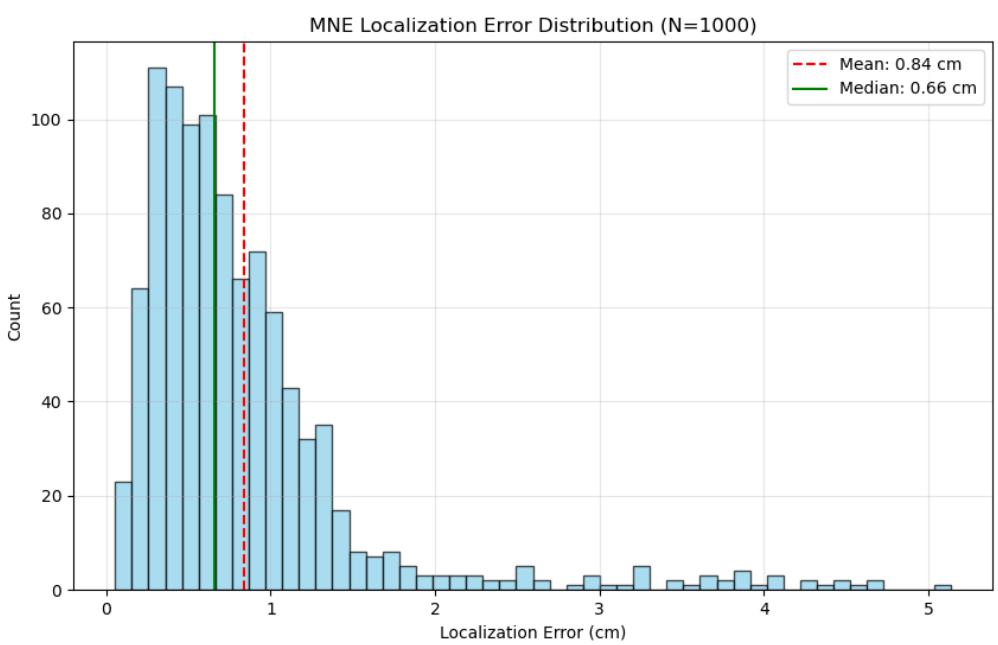

Figure 10. MNE Localization Error Distribution. Histogram of errors for the MNE baseline on the validation set. The mean error is 0.84 cm, significantly higher than the 0.45 cm achieved by the PINN.

7. Conclusion

This work has presented a comprehensive open-source framework for solving the MEG inverse problem on realistic head geometries. By integrating a FEniCS-based forward solver with a Physics-Informed Neural Network (PINN) inversion scheme, we demonstrated a method that bridges the gap between rigorous physical modeling and modern machine learning. A key distinction of our approach, compared to recent deep learning methods such as DeepSIF or Deep-MEG, is the direct embedding of electromagnetic physics into the neural network's loss function. Instead of learning a purely statistical mapping from sensors to sources, the PINN is constrained by the Poisson equation and the Biot-Savart law, ensuring physically valid solutions even in the presence of noise. Our PINN architecture, which effectively combines this physical regularization with massive unlabeled datasets, offers a robust pathway for reducing the ill-posedness of the inverse problem. Future work will focus on extending this framework to time-varying sources and validating the approach on clinical recordings.

Acknowledgments

Computations were performed using FEniCS 2019.2 and Google Colab cloud resources and NVIDIA GPU acceleration for training the neural networks.

Data and Code Availability

The source code, example data, and documentation are available at `https://github.com/ogiannopoulou/` All software is released under the MIT license to promote reproducibility and community development.

References

[1] M. Hämäläinen, R. Hari, R. J. Ilmoniemi, J. Knuutila, and O. V. Lounasmaa. Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev. Mod. Phys.*, 65(2):413–497, 1993.

[2] J. S. Ebersole and S. M. Ebersole. Combining meg and eeg source modeling in epilepsy evaluations. *J. Clin. Neurophysiol.*, 27(6):360–371, 2010.

[3] M. Teplan. Fundamentals of eeg measurement. *Meas. Sci. Rev.*, 2(2):1–11, 2002.

[4] J. Sarvas. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Phys. Med. Biol.*, 32(1):11–22, 1987.

[5] S. Baillet, J. C. Mosher, and R. M. Leahy. Electromagnetic brain mapping. *IEEE Signal Process. Mag.*, 18(6):14–30, 2001.

[6] T. W. Picton, S. Bentin, P. Berg, E. Donchin, S. A. Hillyard, R. Johnson Jr, G. A. Miller, W. Ritter, D. S. Ruchkin, M. D. Rugg, and M. J. Taylor. Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, 37(2):127–152, 2000.

[7] D. B. Geselowitz. On bioelectric potentials in an inhomogeneous volume conductor. *Biophys. J.*, 7:1–11, 1967.

[8] D. B. Geselowitz. On the magnetic field generated outside an inhomogeneous volume conductor by internal current sources. *IEEE Trans. Magn.*, 6(2):346–347, 1970.

[9] G. Dassios. *Ellipsoidal Harmonics: Theory and Applications.* Cambridge University Press, Cambridge, 2012.

[10] F. Kariotou, S. N. Giapalaki, and D. Hadjiloizi. The direct eeg problem in the ellipsoidal geometry. In *Proceedings of the 3rd Hellenic Conference on AI (SETN 2004)*, pages 448–457, 2003.

[11] G. Dassios and D. Hadjiloizi. On the non-uniqueness of the inverse meg problem. *Inverse Probl.*, 23:L1–L5, 2007.

[12] F. Kariotou. Electroencephalography in ellipsoidal geometry. *J. Math. Anal. Appl.*, 290:324–342, 2004.

[13] S. N. Giapalaki and F. Kariotou. The complete ellipsoidal shell-model in eeg imaging. *Abstr. Appl. Anal.*, 2006:Article ID 57429, 2006.

[14] G. Dassios and A. S. Fokas. Electro-magneto-encephalography for the three-shell model: Distributed current in arbitrary, spherical and ellipsoidal geometries. *J. Roy. Soc. Interface*, 10(88):20130165, 2013.

[15] G. Dassios, F. Kariotou, and M. N. Tsampas. On the behaviour of the electric potential inside a layered ellipsoidal head model for electroencephalography. *Math. Methods Appl. Sci.*, 30:679–704, 2007.

[16] A. S. Fokas, Y. Kurylev, and V. Marinakis. The unique determination of neuronal currents in the brain via magnetoencephalography. *Inverse Probl.*, 20:1067–1078, 2009.

[17] A. C. Papanicolaou. An introduction to magnetoencephalography with some applications. *Brain Cogn.*, 27(3):331–352, 1995.

[18] G. Dassios and A. S. Fokas. The definite non-uniqueness results for deterministic eeg and meg data. *Inverse Probl.*, 21:1001–1013, 2005.

[19] G. Dassios and F. Kariotou. The exterior magnetic field for the multilayer ellipsoidal model of the brain. *Q. J. Mech. Appl. Math.*, 62:1–23, 2009.

[20] Arjan Hillebrand and Gareth R Barnes. Beamformer analysis of meg data. *International review of neurobiology*, 68:149–171, 2005.