

Reconstructing whole-brain structure and dynamics using imaging data and personalized modeling

M. Fabbrizzi^{*1,2}, L.G. Amato^{*1,2}, L. Martinelli³, J. Carpaneto^{1,2}, E. Bartolini⁴, S. Calderoni^{4,5}, A. Retico⁶, A. A. Vergani^{1,2}, A. Mazzoni^{1,2}

¹The BioRobotics Institute, Sant'Anna School of Advanced Studies, Pisa, Italy

²Department of Excellence in Robotics and AI, Sant'Anna School of Advanced Studies, Pisa, Italy

³Department of Physics, University of Pisa, Pisa, Italy

⁴Department of Developmental Neuroscience, IRCCS Fondazione Stella Maris, Pisa, Italy

⁵Department of Clinical and Experimental Medicine, University of Pisa, Pisa, Italy

⁶National Institute for Nuclear Physics, Pisa, Division, Pisa, Italy

Abstract

Brain structure plays a pivotal role in shaping neural dynamics. Current models lack the anatomical and functional resolution needed to accurately capture both structural and dynamical features of the human brain. Here, we introduce the FEDE (high FidElity Digital brain modEl) pipeline, generating anatomically accurate brain digital twins from imaging data. Using advanced techniques of anatomical tissue segmentation and finite-element analysis, FEDE reconstructs brain structure with high spatial resolution, while also replicating whole-brain neural activity. We demonstrated its application by creating the first brain digital twin of a toddler with autism spectrum disorder (ASD). Through parameter optimization, FEDE replicated both time-frequency and spatial features of recorded neural activity. Notably, FEDE predicted patient-specific aberrant values of excitation to inhibition ratio, coherently with ASD pathophysiology. FEDE represents a significant leap forward in brain modeling, paving the way for more effective applications of digital twin in experimental and clinical settings.

Introduction:

Computational brain models are widely utilized to investigate the structural underpinning of brain activity. Bioinformatics platforms such as The Virtual Brain (TVB)^{1,2}, have emerged, enabling high-fidelity simulations of brain activity. Their use ranges from the identification of structural phenomena shaping brain dynamics in physiological conditions³⁻⁵, to clinical applications such as disease diagnosis⁶ and testing of therapeutic approaches^{7,8}. However, the significant simplifications required to simulate and analyze virtual brain data efficiently have limited current models to replicating neural activity dynamics without achieving full anatomical accuracy.

Over the past years, several advanced MRI preprocessing and postprocessing techniques have been developed, enabling accurate studies of brain structure and functions in patients⁹⁻¹⁴. These advancements include methods to quantify voxel-wise myelination levels^{11,15}, which are crucial for determining conduction velocity between brain regions¹⁶. Moreover, recent techniques take into

account detailed anatomical properties, such as tissue anisotropy^{17,18}, which significantly influence how brain activity is generated and propagated across brain tissues¹⁹.

Despite these advancements, there is currently no pipeline that successfully combines imaging analytic tools and computational models into a unified framework. Each technique requires specific software and packages (see Methods), resulting in a fragmented approach that limits integration. As a result, no existing model can achieve both an anatomically accurate reconstruction of the human brain and a high-fidelity replication of neural activity. This limits the potential to produce effective digital twins, i.e. virtual models that replicate the physiological and anatomical characteristics of the brain structure, capturing brain dynamics with the fidelity needed for comprehensive studies^{20,21}.

To address these challenges, we present FEDE (high FidELity Digital brain modEl), a pipeline incorporating state-of-the-art imaging analysis tools and computational techniques to reconstruct brain anatomy and replicate individual whole-brain activity with high precision. FEDE incorporates advanced software to reconstruct detailed structural features of the human brain, including local cortical connections, voxel-wise myelination levels and tissue-specific anisotropy and conductance properties, shaping neural activity generation from cortical sources.

We demonstrated FEDE's application by creating the high-fidelity digital brain twin model of a toddler (between 1y and 3y old) with autism spectrum disorder (ASD). Developing a digital twin of a toddler presents unique challenges, such as the incomplete myelination of frontal lobes¹⁶ and skull development, requiring tailored pipelines for young brain MRI analysis. Additional difficulties arise from the unclear etiology of ASD and the need to capture its multi-scale anomalies, ranging from synaptic dysfunction to whole-brain connectivity, without relying on predefined assumptions. ASD is a complex neurodevelopmental condition characterized by significant variability in its presentation and severity^{22,23}. Several abnormalities in brain structure and function have been reported in ASD, including synaptic imbalance^{24,25}, global and local connectivity alterations^{26,27}, and abnormal brain growth^{28,29}. Particularly, ASD is well documented to impair local connectivity within the brain^{26,30}, which is crucial for cognitive and sensory processes²⁷. Traditional models of brain activity often overlook the importance of this local connectivity, focusing instead on global measures of connectivity^{31,32} or synaptic transmission^{27,33,34}. Inter-subjective heterogeneity makes ASD an ideal candidate for the use of digital twins²⁰, as both structural and dynamical alterations contribute to the condition's pathophysiology.

FEDE replicated whole-brain neural activity of the patient (measured via EEG recordings) with high precision, accurately reproducing both time-frequency domain features, such as power spectral density (PSD), and spatial features, such as functional connectivity. Moreover, it identified aberrant values of excitation to inhibition ratio as the structural culprit of the experimental neural activity dynamics, consistent with ASD pathophysiology.

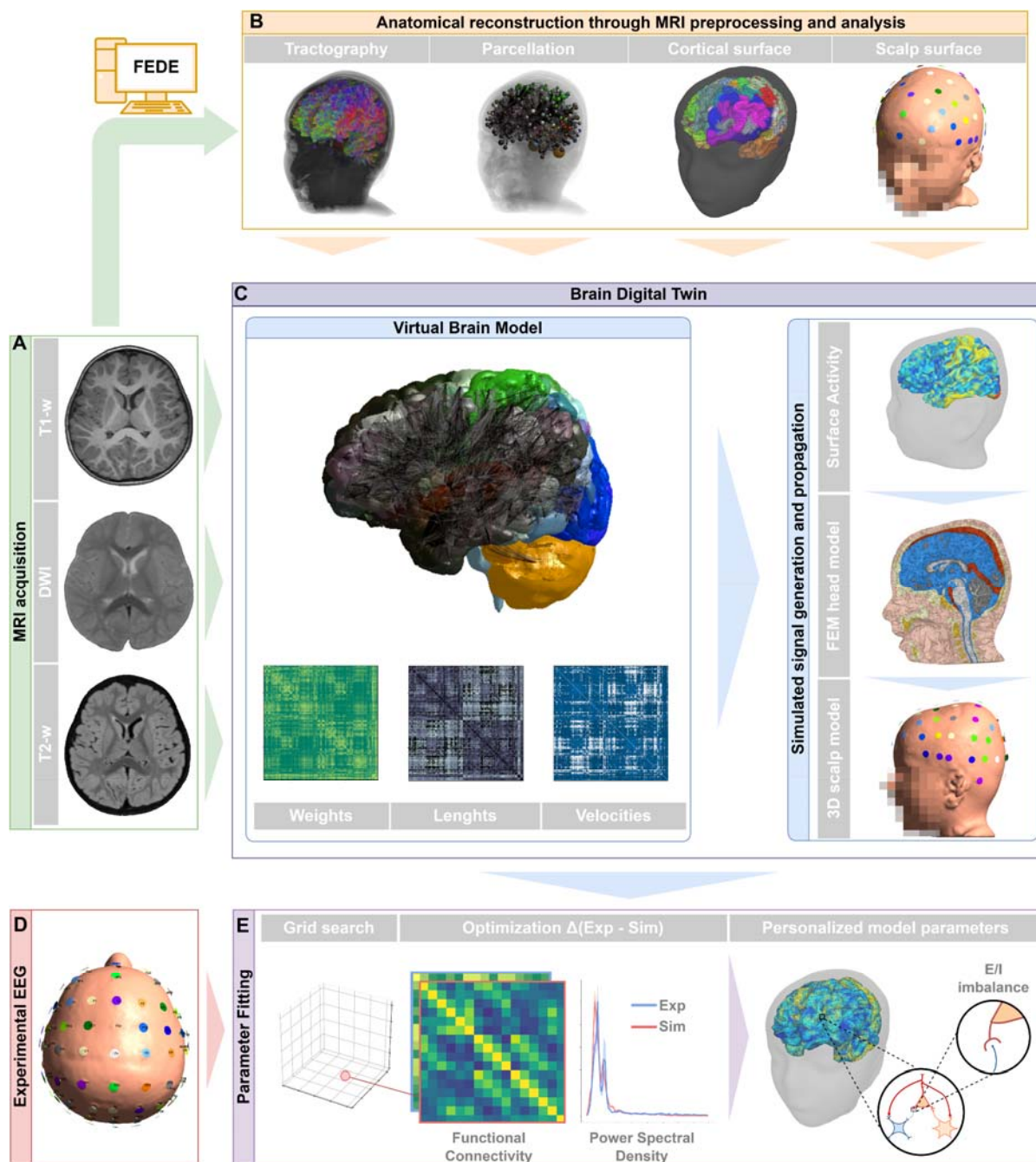


Figure 1: The FEDE (high Fidelity Digital brain modEl) pipeline: (A): MRI recordings, including (from top to bottom) T1-w, DWI and T2-w, were used to generate a 3D replica of the patient brain. **(B):** MRI processing steps: from left to right anatomical constrained tractography analysis, brain regions parcellation and segmentation according to the HCPMMP1 atlas. Right end: reconstructed 3D model of the patient's head, including a high-density mesh of the cortical surface and a scalp model with EEG electrode positions. **(C):** Bottom: the parcellation of brain areas allowed to compute connective weights, distances and conduction velocity maps for the whole brain structure, which were integrated in a virtual brain model (top, see Methods). Top right: Neural mass models were used to compute neural activity on the high-density cortical surface. Middle right: The activity was then projected to the scalp of the patient with an anatomically-accurate lead-field matrix, leveraging a FEM model of the patient's head and including tissue anisotropy of brain tissues. Bottom right: This allowed for the computation of EEG signals from simulated brain activity. **(D):**

EEG recordings were acquired during resting-state and EEG features such as power spectral distribution and functional connectivity are extracted. **(E)**: Optimal fit between experimental and simulated EEG led to the identification of structural parameters underlying patient's condition. Parameter inversion, based on a grid search on candidate parameters, allowed to infer personalized parameters that replicated at best the experimental EEG recordings of the patient.

Results

In this study, we present FEDE: a high-precision pipeline for the creation of personalized digital twin model leveraging comprehensive MRI data (Figure 1A), including T1-weighted (T1-w), T2-weighted (T2-w), and Diffusion-Weighted Imaging (DWI). Briefly, the model extracts MRI features (Figure 1B and Methods) to reconstruct brain properties (Figure 1C left and Methods), which are then used to simulate EEG signals (Figure 1C right and Methods). These simulated EEG signals can capture spatiotemporal organization of recorded EEG signals of the reconstructed brain (Figure 1E) and from this infer properties of the underlying structure and function (Figure 1E). The pipeline was tested on a toddler (aged between 1y and 3y) with autism spectrum disorder (ASD), to reconstruct an anatomically accurate digital twin replica of the patient's brain and to replicate with high fidelity the experimental EEG recordings.

Personalized conduction velocity map is needed for accurate reconstruction of neural connections.

The FEDE pipeline comprises the individual reconstruction of neural connections from MRI data. Brain regions were parcellated according to the HCPMMP1 atlas³⁵, encompassing 379 cortical and subcortical regions. Connections between regions were computed from DWI data (Supplementary Figure S1) and tract lengths (the average length of white matter fiber bundles connecting two regions) were determined from tractography DWI analysis. Crucially, alongside standard tractography, FEDE incorporates the computation of conduction velocities from brain areas, which considers voxel-wise levels of myelination and their role in determining velocity of signal transmission across brain areas.

In FEDE, voxel-wise levels of myelination are first computed from T1-w and T2-w imaging combined with apparent fiber density (obtained from DWI³⁶, see Methods) using Equation 1. Conduction velocity map is then computed from Equation 2, enabling a detailed analysis of signal transmission properties across the brain (Fig 2A). The conduction velocity map offers an improved method for understanding how signals propagate through different brain regions by considering the anatomical and microstructural properties of brain tissues, primarily myelination of white fibers. Notably, a lower conduction velocity can be observed in the frontal lobe (Fig 2A panel iv) due to the young age of the subject and the myelination process not completed. This introduces topographic differences in conduction velocities that would be overlooked using standard methodologies.

The model was then used to combine the voxel-wise map with the selected gray matter parcellation. This led to the computation of the patient-specific conduction velocity matrix $CV_{a,b}$. Each entry of the matrix encodes the conduction velocity between region a and region b in the adopted parcellation (Fig 2B). This matrix allows for the calculation of delays, dividing the tract lengths matrix by the conduction velocity matrix (TL/CV).

We compared the delays computed using FEDE with those derived from the standard methodology, in which delays are simply proportional to the tract lengths matrix. Significant discrepancies were found between the two methods (Fig 2C). Delays calculated with FEDE were significantly lower with respect to those computed with standard methodology (ratio of 0.79 ± 0.05 , Mann-Whitney test, $U=1.3e+11$, $p<0.00001$), suggesting that delays computed using standard methods tend to overestimate the actual values.

The overestimation of delay times by standard methodology was further analyzed by examining inter-hemispheric and intra-hemispheric connections separately (Fig 2E). For inter-hemispheric connections, delays were significantly larger when using the standard methodology (delays computed with FEDE were 0.72 ± 0.04 of those computed with standard methodology, $U=5e+9$, $p<0.00001$). Intra-hemispheric connections also showed statistically significant differences, though

less pronounced compared to inter-hemispheric connections (ratio was 0.87 ± 0.05 , $U=8e+9$, $p < 0.00001$).

The overestimation of delay times with standard methodology was confirmed by a pairwise comparison between delay times computed with FEDE and with standard methodology (i.e., from the length of connective tracts between regions). Delay times computed with FEDE strongly correlated with tract lengths ($r=0.989$, $p < 0.00001$, Fig 2F), with relevant differences from the exact proportionality postulated by standard methodology, (orange line in Fig 2F). Delay times computed with FEDE were found to be smaller than values expected from the standard methodology (dots below the orange line in Fig 2F) in 96% of the cases. The significance of the pairwise differences in delay times was confirmed by a χ^2 test (test= $7.8e+5$, $p < 0.00001$). A distribution analysis between FEDE and standard methodology also confirmed statistically significant differences (Kolmogorov-Smirnov test= 0.086 , $p < 0.00001$). This consistent overestimation introduces a significant bias in models built on standard delay values, potentially affecting the accuracy of any subsequent analysis on functional connections and brain activity.

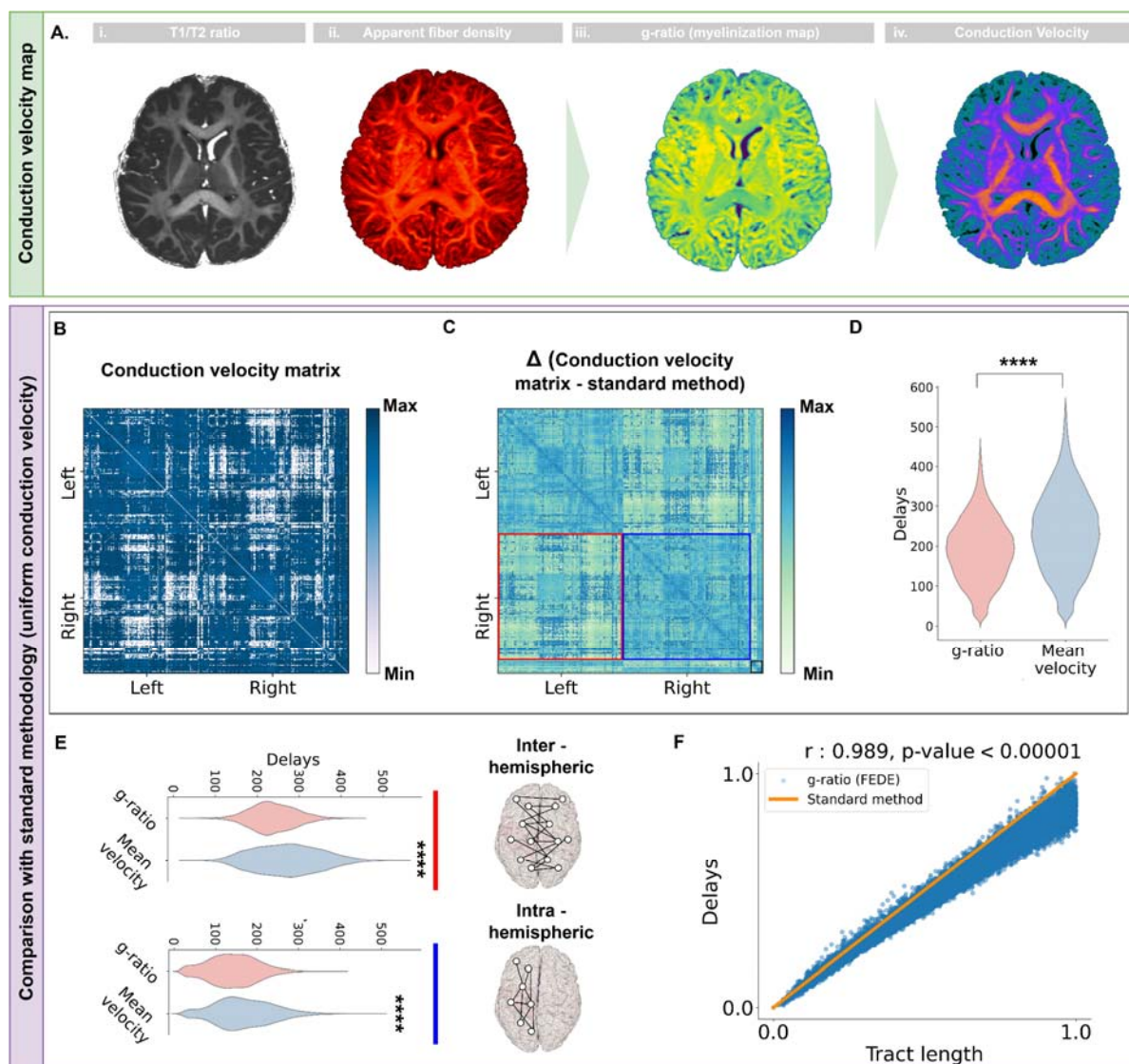


Figure 2: Delays in neural activity propagation are significantly overestimated by standard methodologies. (A): Conduction velocity map is computed from the ratio between T1-w and T2-w (i), combined with apparent fiber density (ii) obtained from DWI (see Methods). It allows for the computation of axonal myelin levels, defined as the g-ratio (iii), which leads to a map representing voxel-wise levels of conduction velocity of brain signals (iv). Orange indicates high values, while low values of conduction velocities are indicated in blue. **(B):** Conduction velocity matrix $CV_{a,b}$, whose entries encode the conduction velocity from region a to region b . **(C):** Difference between delays computed dividing the tract lengths matrix by the CV matrix (TL/CV) and delays computed with standard methodology (i.e. from the tract lengths matrix and a mean conduction velocity value). Red (blue) square represents inter-hemispheric (intra-hemispheric) connections, while black box represents subcortical regions. **(D):** Violin plot of delay values computed with the FEDE g-ratio method (red) and standard method (cyan). Delay values computed from a mean conduction velocity value tend to overestimate with high significance actual delay values computed from MRI analysis. **(E):** Violin plot of delay values computed with the two methods, separately for inter-hemispheric connections (top) and intra-hemispheric connections (bottom). Inter-hemispheric delays are significantly overestimated with standard methodology (see main text). **(F):** Scatter plot of tract length values and delays computed using the CV matrix obtained with FEDE. Orange line represents the expected delays according to standard methodology. Significance notation: **** stands for $p < 0.00001$.

Anatomically-accurate lead field matrix allows to precisely deduce cortical sources of simulated EEG signals.

To achieve a high-fidelity replication of individual neural activity, we computed an anatomically-accurate forward-solution utilizing personalized tissue reconstruction from MRI data (Fig 3A). In FEDE, neural activity is simulated using the Jansen-Rit neural mass model³⁷ on a high density cortical mesh of 20,484 vertices. The simulated cortical activity is then projected onto the scalp surface via a lead-field matrix (LFM) meticulously derived from the patient's anatomical data (see below and Methods). This matrix ensures that the projection of neural signals to the EEG electrodes accurately reflects the patient's unique brain anatomy³⁸.

To generate an anatomically-accurate LFM, FEDE constructs a finite element model (FEM) of anatomical tissues from MRI data using the SIMNIBS software (see Methods). This approach allows the LFM computation to account for ten distinct tissue types (plus two for electrodes and saline solution used on the patient scalp), each with its specific conductivity and anisotropic properties (see Supplementary Fig. S2).

We compared this anatomically-accurate LFM to one derived from standard boundary element methods (see Methods). For each electrode, we calculated the cortical contributions to the LF matrices, organizing cortical vertices by their distance from the electrode (Fig 3B). A Kolmogorov-Smirnov test revealed significant differences in the distributions of LFM values between the two matrices for all electrodes ($p < 0.00001$). We analyzed LFM values channel-wise, to assess if standard methodology introduced a significant bias in determining LFM contributions for different electrodes. Significant differences were found in 11/16 electrodes. The standard method overestimated contributions for 13/16 electrodes, with significant overestimation ($p < 0.05$) in 9 electrodes. For the remaining 3 electrodes, standard methodology caused an underestimation though this difference was statistically significant only in two of them. Notably, the bias introduced by standard methodology was strongly lateralized, with 10 of the 11 electrodes that presented statistically relevant differences located in the central or left portion of the scalp (Fig 3C).

Additionally, we conducted a separate analysis for vertices near to the electrode position (defined as the 20% closest vertices for each electrode), to assess whether standard methodology

overestimated their LFM contributions. Quantitative analysis confirmed this hypothesis, with the standard method significantly overestimating contributions from near vertices (0.32 ± 0.08 for FEM vs. 0.50 ± 0.10 for standard method, Fig 3D), with a strong statistical significance ($U=57$, $p=0.008$). On the contrary, differences between the two methods were non-significant for vertices far from the electrodes (defined as the top 20% farthest vertices for each electrode, $U=113$, $p=0.58$). The interaction between the distance from the electrode and the differences in LFM contribution between standard method and FEDE was confirmed by a two-way ANOVA ($f=3.49$, $p=0.066$).

In summary, the standard method introduced substantial bias in estimating cortical contributions to EEG signals, particularly by underestimating the effects of volume conduction and overestimating cortical source specificity.

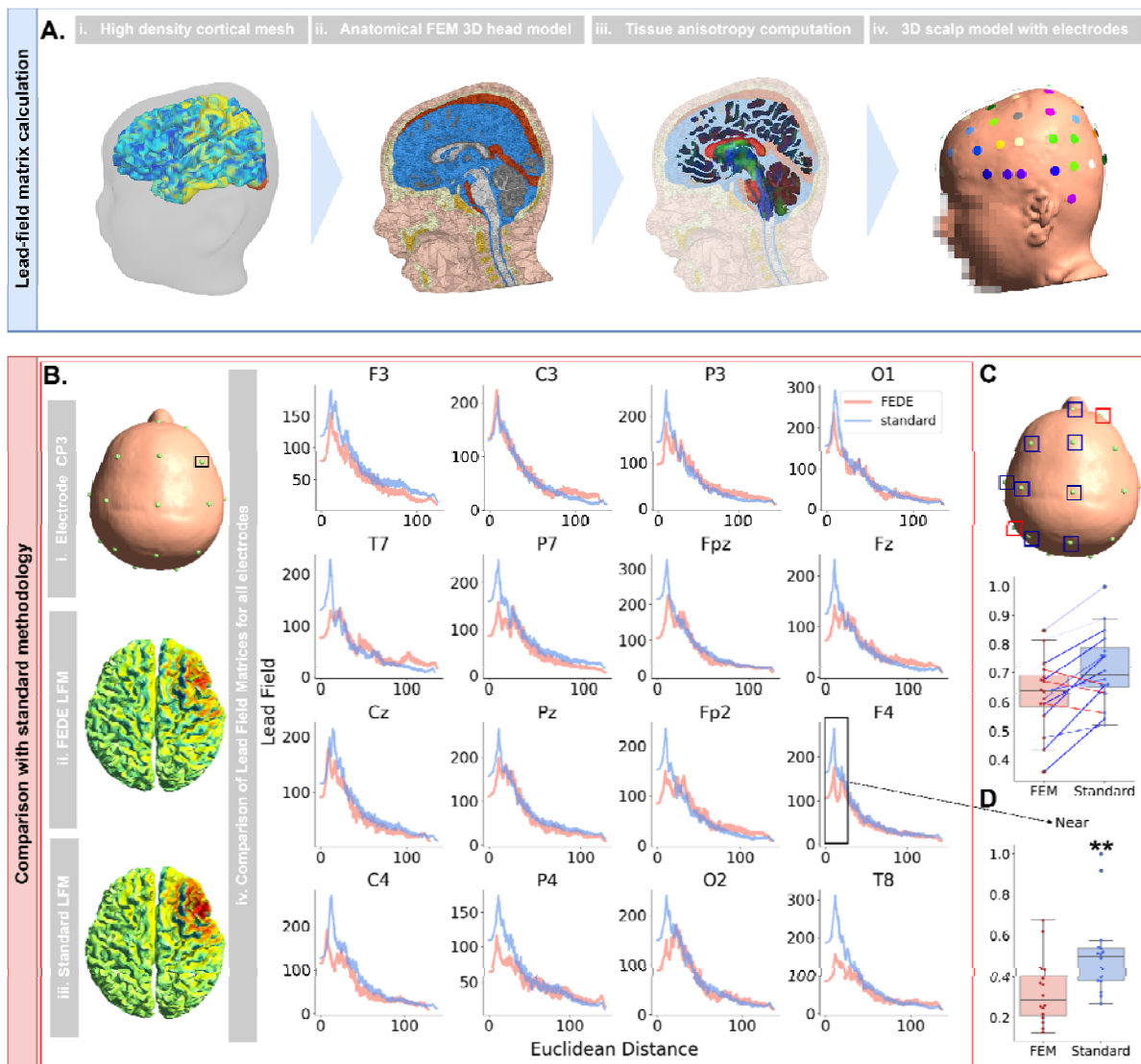


Figure 3: Anatomically-accurate forward-solution with personalized tissue reconstruction identifies correct cortical sources of each EEG channel. (A): From left to right: (i) Cortical activity computed on the high-density cortical mesh; (ii) anatomical tissues, each with its own conductance value, were reconstructed with SIMNIBS software, creating a 3D FEM model of the patient's head, which was combined with tissue anisotropy computed from DWI data (iii). This allowed for the precise propagation of the electric field onto the scalp surface, and thus to the

electrode grid (iv). This procedure results in the reconstruction of the cortical contribution to the activity of each electrode. **(B)**: Different LFMs for the F4 electrode (i) plotted as an example, with FEDE model LFM (ii) and standard method LFM (iii); In each EEG channel, LFM contribution as a function of the distance to the electrode (iv) presented strong differences between FEDE model (salmon) and standard method (light blue). Only channels selected for experimental analysis are shown. **(C)**: Scalp view and boxplot of LFM values according to both standard methodology and FEDE. Statistical tests revealed that standard method resulted in an overestimation of LFM values in most EEG channels (indicated by blue color in the boxplot and in the scalp plot), while it resulted in an underestimation only in two channels (indicated by red color in the boxplot and in the scalp plot). Only channels selected for experimental analysis are shown, and channels with significant differences across methods were highlighted. Whiskers represent interquartile ranges. **(D)**: Standard methodology significantly overestimates contributions from near vertices. Near vertices are defined (for each channel) as the 20% of closest vertices; in panel (B) is reported the example of LFM values of near vertices for the F4 electrode. Notation is the same as in (C). Significance notation: ** stands for $p < 0.01$, Bonferroni correction.

FEDE replicates accurately the spatiotemporal structure of experimental EEG recordings

Leveraging its anatomically accurate 3D cortex and head model, FEDE allowed to simulate individual brain activity with high precision. We validated the model by reproducing the previously acquired experimental recordings of the toddler with ASD (see Methods).

We determined the combination of model parameters leading to the optimal replicated the experimental EEG signals. We first analyzed the PSD of the digital twin model, finding high resemblance to the experimental one (Fig 4A). The analysis of different EEG frequency bands (Fig 4B) further demonstrated that the relative power of the digital twin's PSD correctly replicated the theta and delta dominant rhythm observed in the experimental recording supporting the model's reliability in simulating distinct oscillatory behaviors. Linear regression between simulated and experimental PSD values confirmed the high accuracy of the digital twin, with a correlation value $r = 0.92$, $r^2 = 0.84$ $p < 0.00001$.

Additionally, the model reproduced the complex topography observed in experimental recording (Fig 4C), maintaining consistent correlations with experimental EEG across all channels (see Fig 4D for differences between the two). The average correlation between simulated and experimental single channel PSDs was $r = 0.81 \pm 0.11$, and a mean squared error analysis yielded a 1-MSE value of 0.84 ± 0.10 (Fig 4E). These results indicated that the FEDE not only captured individual channel characteristics but also preserved a high degree of accuracy in simulating spatial EEG patterns. Channel-wise comparisons between simulated and experimental PSDs can be found in Supplementary Fig. S3A (an example channel is found in Fig 4F). We also conducted a single channel analysis for each frequency band, to assess the ability of the digital twin in replicating the experimental relative powers in each band for each channel. Kolmogorov-Smirnov test conducted between simulated and experimental distributions of relative powers revealed no difference in the EEG frequency bands (Supplementary Table 2, distributions are found in Supplementary Figure S3B).

The digital twin also correctly replicated experimental FC. When comparing functional connectivity matrices derived from both simulated (Fig 4G) and experimental recordings (Fig 4H), the digital twin produced FCs with strong correlation to experimental data (Fig 4I, $r = 0.80$, $r^2 = 0.64$, $p < 0.00001$). A Kolmogorov-Smirnov test comparing the distributions of FC values from the two data sets found no statistically significant difference (test = 0.13, $p = 0.58$), indicating that the model successfully replicated the complex patterns of brain connectivity observed in the experimental recordings (Fig 4J). A channel-by-channel comparison further corroborated the absence of relevant

differences between simulated and experimental signals in FC values (Mann-Whitney U test, $U=159.0$, $p=0.25$, not shown). FC replication was also repeated with metrics less affected by volume conduction such as phase slope index³⁹, obtaining similar results ($r=0.53$, $p<0.00001$).

Despite the generally good performance in replicating experimental EEG features, only few parameter combinations yielded precise replication of experimental PSD (Fig 4K), while many combinations allowed for an accurate replication of experimental FC (Fig 4L). This finding may be due to the influence of anatomical connections in shaping FC patterns (see Discussion) which highlights once more the importance of the digital twin model used as a scaffold to simulate brain activity.

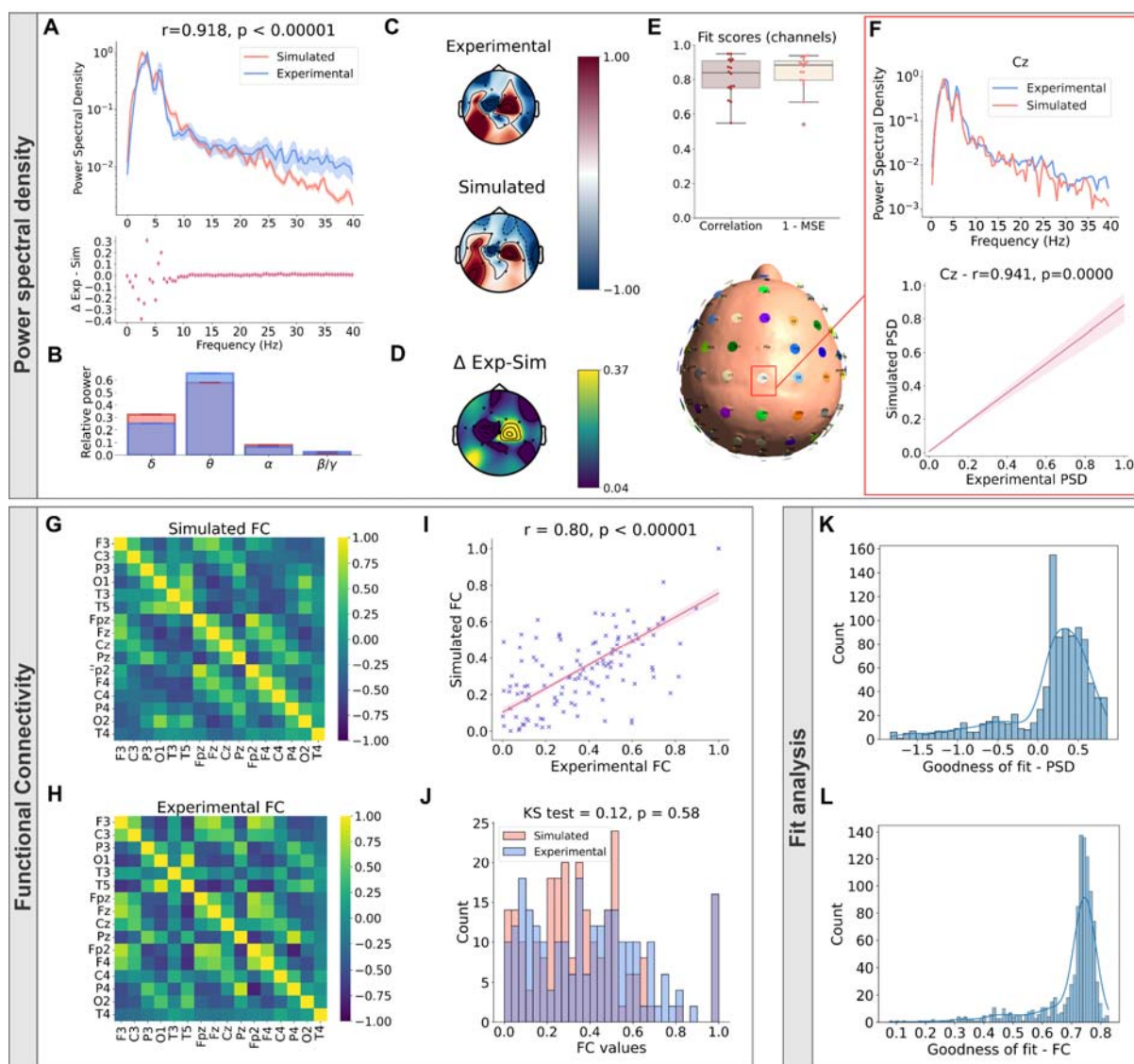


Figure 4: FEDE spatiotemporal accuracy in replicating experimental EEG signals (A): Patient's average PSD (top plot, light blue) superimposed to the average PSD of the digital twin (salmon). Bottom inset shows residual plot. The r -value between simulated and experimental data was of 0.918. **(B):** Relative power in different EEG bands for the patients and the digital twin. Notation is the same as in (A). **(C):** Topographic plot of scalp EEG activity in the patient and in the digital twin, indicating that the model can reproduce the non-trivial topographies encountered in experimental recordings. Signals were normalized between -1 and 1. **(D):** Topographic plot of the

(absolute) difference between experimental and simulated EEG activity reported in (C). **(E)**: Fit scores of the digital twin, reporting correlation values and accuracy (measured as $1 - \text{mean squared error}$) between simulated and experimental recordings. Each point represents a different EEG channel. Note how correlation was > 0.6 for all channels. **(F)**: Cz channel is reported as an example, with PSD (top plot) and regression line between experimental and simulated PSD values (bottom plot). Notation is the same as in (A). **(G)**: FC matrix computed from simulated recordings of the digital twin. **(H)**: FC matrix computed from patient's experimental recordings. **(I)**: Linear regression between simulated and experimental FCs, with correlation value reported. The digital twin's FC had a correlation of 0.80 with the experimental one. **(J)**: Histogram of FC values for the patient and the digital twin. Kolmogorov-Smirnov test highlighted no significant differences between the two distributions. Notation is the same as in (A). **(K)**: PSD Goodness of fit analysis for tested combination of model parameters (only values greater than 0 are reported). **(L)**: FC Goodness of fit analysis for tested combination of model parameters (only values greater than 0 are reported). While most combinations of model parameters produced FCs similar to the experimental one, only few combinations allowed to replicate the PSD.

We assessed the ability of the model constructed with FEDE to replicate EEG signals, by comparing it to a model constructed using a methodology not including the advancements introduced with FEDE. Specifically, we calculated the LFM using a standard 3-layer simplified Boundary Element Method (BEM) model of the head without anisotropy and with region-based simulations (see Methods). The structural connectivity matrix implemented in the standard model is the one obtained with the FEDE preprocessing pipeline, including anatomical constrained and probabilistic tractography. We conducted another parameter search to identify the optimal parameter combination for this standard approach. In terms of functional connectivity, the standard approach resulted in reduced fidelity with respect to FEDE (Fig 5A), with the correlation coefficient declining from $r=0.80$ to $r=0.75$. Regarding PSD, results indicated a notable decrease in fidelity for simulated PSD when compared to the FEDE model; the correlation (r -value) between experimental and simulated PSDs dropped from 0.92 to 0.60. Furthermore, this method produced simulated signals with a different dominant frequency than the experimental PSD, and it lacked key features such as the secondary peak in the PSD (Fig 5B).

To evaluate the robustness of both methods, we analyzed the top 100 parameter combinations for each, comparing their fit functions with experimental FCs (Fig 5C) and PSDs (Fig 5D). In both cases, simulations generated by the FEDE model significantly outperformed those generated by the standard method (Mann-Whitney U test: $U=9969$, $p<0.00001$ for FC; $U=7685$, $p<0.00001$ for PSD).

Notably, the similarity between experimental and simulated FC remained high even in the simplified model, probably reflecting the high anatomical accuracy of the connectome reconstruction obtained with FEDE (considering that both models utilized the FEDE-derived connectome). However, the absence of anatomical details in the LFM and in the 3D head replica of the standard model significantly decreased its ability to capture finer details of the experimental PSD such as the second peak at low alpha frequency.

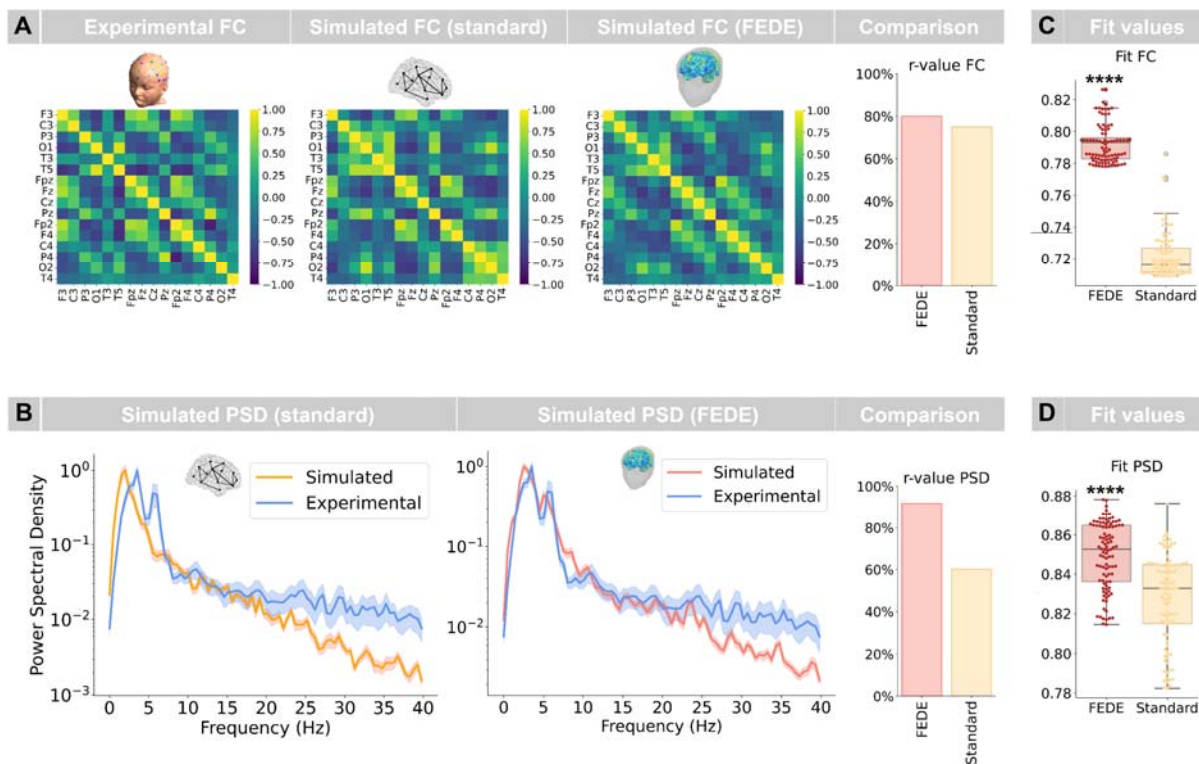


Figure 5: FEDE outperforms standard digital twin models in replicating experimental EEG features. Model parameters selected through parameter exploration were utilized in a region-based simulation, using as scaffold the 379-regions HCPMMP1 atlas employed for gray matter parcellation. No conduction velocity map was included, and both Lead-field matrix and 3D head model were reconstructed using standard methodology (see Methods). Results were compared with both the experimental values and with surface-based simulations. **(A):** FC matrices computed from standard digital twin showed different patterns with respect to those observed in experimental FC. For FC also, standard digital twin simulation presented a smaller Pearson r-value with experimental FC compared to the value obtained with FEDE. **(B):** PSD computed from region-based simulations present reduced similarity with experimental signals, presenting different dominant frequency and no second peak in low-alpha band. This caused the standard digital twin simulation to present a smaller Pearson r-value with experimental PSD compared to the value obtained with FEDE. **(C):** Fit values between experimental and simulated FC for 100 best standard and FEDE simulations. FEDE presented higher fit values in almost all simulations. **(D):** Fit values between experimental and simulated PSD for 100 best standard and FEDE simulations. FEDE presented higher fit values with high significance. Significance notation: **** stands for $p < 0.00001$.

Hierarchical parameter fitting to reconstruct personalized structural parameters of the patient's brain

To determine the best combination of model parameter that replicated at best experimental EEG signals (see Fig 4), we simulated EEG signals from various parameter combinations (see Methods). We then analyzed their performance in replicating key metrics such as power spectral density and functional connectivity. These simulated signals were compared to the actual EEG recording of the patient, using the hierarchical parameter fitting procedure implemented in FEDE, based on the loss functions reported in Equations 3-4 (see Methods).

In each step of the procedure (Fig 6A), FEDE operates a grid search over a range of candidate parameters, identifying the optimal parameter combination by minimizing the loss functions. In the first step, FEDE determines the optimal values of parameters describing connections between brain areas. These parameters include the global connectivity coupling, the area and the strength of the local connectivity gaussian kernel (Step 1 in Fig 6A), and the conduction velocity proportionality constant. The second step comprises the identification of parameters determining the frequency output of the Jansen-Rit model, including the post synaptic potentials, noise and time constants of excitatory and inhibitory subpopulations, as well as the mean input firing rate. The ratio between inhibitory and excitatory time constants allows to determine also the excitation to inhibition ratio (EI ratio) of the brain. The last step involves the tuning of Jansen-Rit parameters that determine finer effects on output activity, such as the total number of connections in the neural mass, and the magnitude of links between single subpopulations and the sigmoid term transforming membrane potential into firing rate³⁷. At each step, the initial point of the grid is determined by the best match identified in the previous step.

In our patient, this procedure allowed to increase at each step the precision of both simulated FC (Fig 6B) and PSD (Fig 6C). FEDE also comprises a module for the analysis of fitted parameters, allowing to identify the structural brain parameters playing the most relevant role in the agreement with experimental EEG activity of the patients (see Methods and Fig 6D). A SHAP analysis (see Methods) is implemented to quantitatively identify the model parameters that allow to replicate at best experimental EEG activity. The analysis revealed that only two parameters presented relatively high SHAP values, being EI ratio (0.0033 ± 0.008) and noise (0.0041 ± 0.024), while all other parameters presented values < 0.00001 .

Since each model parameter describes biophysical quantity closely related to microscopical mechanisms of neural connectivity and synaptic transmission, the optimal values identified by the procedure can be compared with physiological values to support diagnosis of pathological conditions. In our patient, this approach led to the identification of aberrant values of background noise (whose optimal value was 100 times greater than the standard one^{2,37}, Fig 6E) and EI ratio (found to be five times greater than the standard healthy value³⁷, see Fig 6E), which were essential in correctly replicating the experimental EEG activity. Aberrant values of both parameters are consistent with current hypotheses on ASD pathophysiology^{24,25,40,41}.

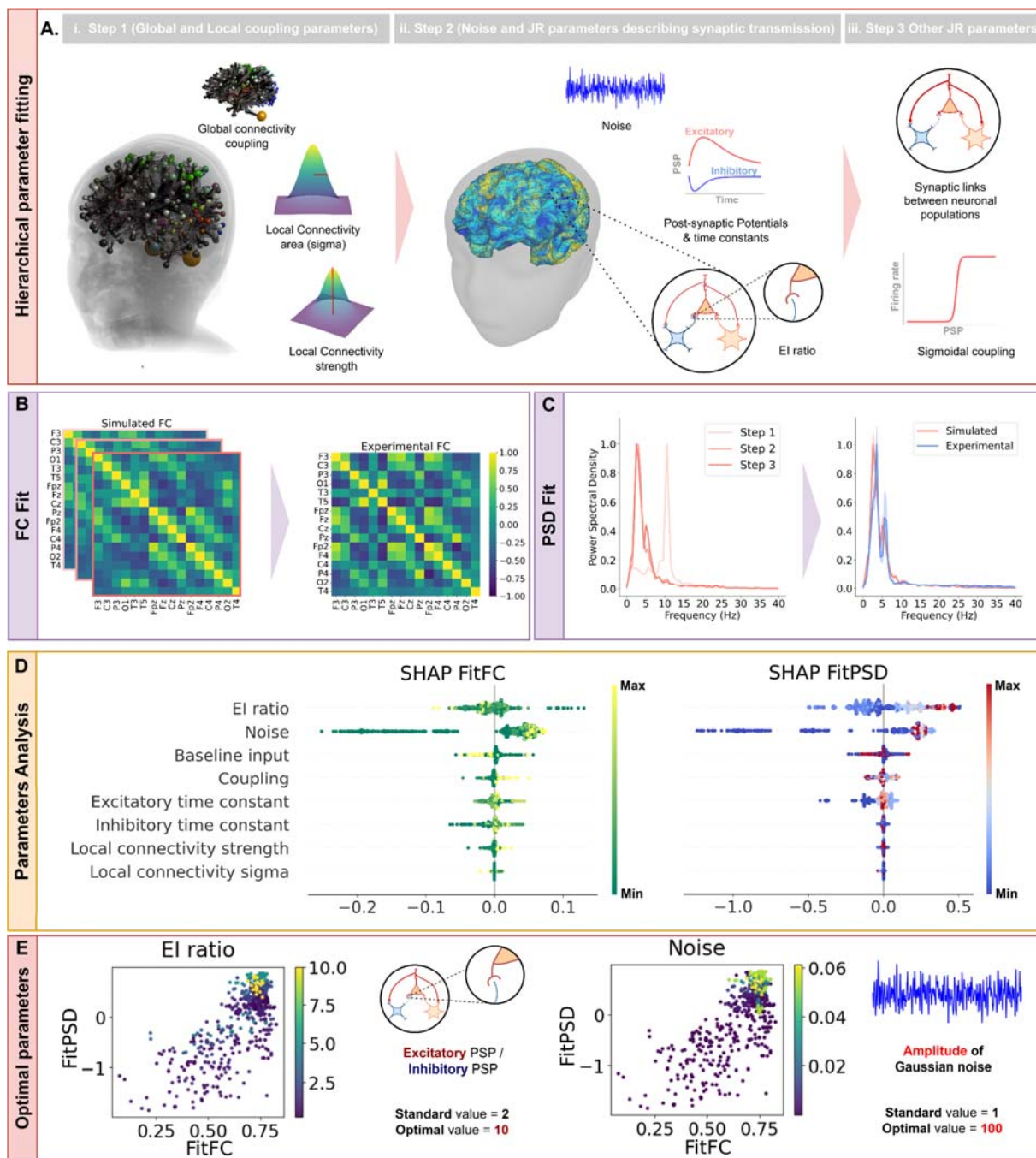


Figure 6: Hierarchical parameter fitting was employed to replicate EEG features and reconstruct underlying structural parameters. (A): The parameter fitting procedure followed three steps. At each step, the procedure identified a combination of model parameters, that was used as starting point for the next step of the parameter fitting procedure. The first step (i) identifies values of structural parameters of global and local connectivity, as well as velocity proportionality constant. The second step (ii) identifies Jansen-Rit parameters that dictate the frequency output of the model and mean input firing rate. The third step (iii) identifies Jansen-Rit parameters that govern finer features of the simulated output (see for example the second peak of the PSD). The last step focused on the sigmoid transforming membrane potential into firing rate. **(B-C):** At each step, the fidelity in replicating both FC and PSD increases, thanks to the identification of more precise parameters. **(D):** The parameter analysis module allows to investigate the importance of each parameter in correctly replicating experimental EEG activity, including statistical tools like

SHAP analysis (in the labels Baseline input represents the total input to excitatory interneurons as sum of mean input firing rate, noise, global and local coupling outputs). **(E)**: The pipeline allows to identify the structural parameters that more robustly determine EEG activity, that in our patient were aberrant values of the EI ratio and of background noise. FEDE includes tools for the visualization of fit values in relation to model parameters values. In the two scatter plots each dot is a different simulation, higher values of the corresponding parameter (respectively EI ratio and noise) are reported in yellow. See how higher values of both parameters allowed to obtain high fit values for both PSD and FC.

Discussion:

The FEDE pipeline generated a high-fidelity digital twin of the brain of a with ASD and to replicate experimental EEG signal with unprecedented precision, while concomitantly identifying structural culprits of relevant EEG features. This was achieved by integrating state-of-the-art methods for MRI analysis, brain modelling and EEG simulation, and represents a substantial leap forward in the use of computational brain models in clinical settings, particularly in the context of personalized medicine.

The use of a detailed conduction velocity map derived from T1-w, T2-w, and DWI data provides a more accurate and nuanced understanding of neural signal transmission delays. It corrects the biases introduced by the traditional approach, ensuring that the physiological variability in conduction velocity across white matter fibers is properly accounted for. In Newman et al.¹⁶ a similar procedure based on the derivation of a conduction velocity map from MRI recordings highlighted significant differences between healthy and ASD subjects. Our approach allows not only to replicate the personalized determination of conduction velocity maps, but also to assess the role played by this quantity in shaping whole-brain activity.

A critical component of our methodology was the precise computation of the lead-field matrix, which defines the contribution of electrical activity from different cortical regions to the signals recorded by each EEG electrode. Our high-resolution cortical model allowed for an accurate calculation of this matrix, minimizing potential errors in source localization. This accuracy was reflected in the detailed comparison of the simulated and experimental EEG data, where the spatial correlations of neural activity (the FC matrix) closely replicated real-life recordings. In current computational brain models, the LFM is derived from an approximated boundary element method based on the compartments between brain, skull and scalp^{7,8,42,43}. However, this approach neglects several aspects of brain anatomy, such as orientation of cortical sources and presence of several intermediate tissues between sources and electrodes with different conductivity values.

In the FEDE pipeline the EEG activity was computed from local sources on a high density (20,484 vertices) cortical mesh. In Wang et al.⁸ neural activity is simulated from a high density cortical mesh of more than 200,000 vertices. However, this procedure required to perform model inversion with a neural mass model-based simulation of 180 nodes, with parameters that were then translated to the surface simulation on the cortical mesh, in which neural activity was computed with a neural field model approach. In our study, we found a decrease in the fidelity of simulated neural signals (see Supplementary Fig S4 for an example). Our approach includes the parameter fitting procedure directly on the activity computed on the cortical mesh, enabling a more precise replication and localization of neural activity, crucial for both clinical and research applications. In clinical contexts, such precision can enhance the diagnosis and monitoring of neurological conditions, where understanding the exact origins of abnormal neural activity is critical.

FEDE allowed to replicate both spatial and time-frequency features of EEG activity. Several work have attempted the replication of PSD^{7,42,44} computed from electrophysiological recordings, no one reporting the high fidelity obtained with FEDE. The replication of experimental EEG functional connectivity with personalized brain modeling has instead never been attempted before to the best of our knowledge. Several studies focused on the replication of FC matrices deduced from functional MRI (fMRI) recordings⁴⁵⁻⁴⁹. Modeling fMRI functional connectivity is more feasible than reproducing functional connections in EEG, as the high spatial resolution of functional MRI allows for precise localization of cortical sources and reduced volume conduction effects, which heavily affect EEG FC metrics³⁹.

The introduction of a module for parameter analysis led to the identification of structural culprit of experimental EEG features. This module could be crucial for the translational use of FEDE, as it allows to reconstruct from EEG recordings key parameters related to structural brain mechanisms

like synaptic transmission and noisiness of brain activity. This holds great promise for the clinical use of the model. In the case of our patient, FEDE identified as key structural parameters aberrant values of EI ratio and background noise. (both quantities are currently hypothesized to concur in determining the ASD pathophysiology^{24,40}). FEDE allows to test multiple hypothesis about ASD underpinning neural mechanisms for the specific patient, which could potentially be different from one patient to another due to the heterogeneity of ASD; the model was prepared also to test hypothesis about conduction velocity and local/global connectivity impairments, however the lack of a healthy control group for comparison did not allow to test them (this being our next focus). It is also important to consider that ASD could potentially be the result of multiple concurrent impairments due to a non-typical neurodevelopment, and this can be tested using the current pipeline (see for instance concurrent combination of high background noise and high EI ratio found in the current subject); this can also be generalized to a multitude of conditions and recording techniques. FEDE offers a unique method for investigating the presence of these alterations from non-invasive recordings, nowadays unfeasible with current methods.

The choice of testing our pipeline with the creation of the first digital brain twin of a toddler patient affected by ASD also holds important implication. This complex task has never been attempted before, as both modelling a toddler brain and capturing the complexities of ASD present unique challenges yet to be addressed by current models. Working with very young subjects presents unique challenges, particularly in the acquisition and preprocessing of MRI and EEG data, affected by subject restlessness, causing several motion artefacts. By incorporating local connectivity into our models and developing a detailed representation of cortical surface activity, we have provided a novel framework for exploring the structural and functional alterations associated with ASD. The realization of a high-fidelity digital twin replica of a toddler with ASD underscores the potential of personalized precision medicine to revolutionize the diagnosis and treatment of neurodevelopmental disorders, paving the way for more effective and individualized therapeutic strategies.

Our study presents several limitations that should be addressed in future research. Firstly, the primary limitation is that our work was conducted on a single patient diagnosed with ASD. While this allowed us to demonstrate the feasibility of creating a digital brain twin in such a complex case, it limits the generalizability of our findings. The absence of a control group or additional subjects further restricts our ability to draw definitive conclusions about the broader applicability of the pipeline. A control subject or group would have provided a valuable baseline to compare the digital brain twin's performance, enhancing the robustness of our results. This will be key in future activities. Additionally, our computational pipeline has yet to be validated across a diverse range of patients with varying conditions. The complexities of brain development and ASD are unique, and while our approach seems to be promising, it remains uncertain whether it will perform equally well in other populations, such as older children or adults with ASD, or those with other neurological conditions. Future studies with larger and more diverse cohorts are essential for validating our pipeline and assessing its broader utility in personalized medicine. Furthermore, while in principle it is possible to model several virtual versions of real-life methods of imaging and recordings, FEDE has been tested only for the replication of EEG signals.

Methods:

Subject recruiting, ethical approval and data acquisition

The proband is a male toddler aged between 1y and 3y with a diagnosis of ASD formulated by a multi-disciplinary team according to the Diagnostic and Statistical Manual of Mental Disorders-Fifth Edition⁵⁰ criteria, and supported the administration of the ADOS-2 Toddler Module⁵¹. The child's developmental level was measured by an experienced psychologist through the Griffiths Scales of Child Development 3rd Edition⁵², and the following scores were obtained: scale A (Foundations of Learning) 99; scale B (Language and Communication) 60; scale C (Eye and Hand Coordination) 105; scale D (Personal–Social–Emotional) 88; scale E (Gross Motor) 105. Moreover, the Childhood Autism Rating Scale second edition (CARS-2)⁵³ was administered by a qualified evaluator with clinical experience in ASD, and resulted in an overall score of 36, which indicates a level of moderate autism. He underwent both neuroimaging and neurophysiological investigations for clinical purposes at a tertiary care university hospital, the IRCCS Fondazione Stella Maris (Pisa, Italy), in order to exclude brain anomalies. The electroencephalogram (EEG) (Micromed, Mogliano Veneto, Italy) study was performed according to the 10-20 International System. Brain magnetic resonance imaging (MRI) was obtained with a GE 3T scanner (Signa HDx, GE-Medical-Systems, Milwaukee, United States) using a protocol optimized for neurodevelopmental disorders which included three-dimensional (3D) T1-weighted (T1W) fast-spoiled gradient echo, 3D T2 fluid-attenuated inversion recovery, two-dimensional (2D) T2-weighted (T2W) fast spin echo, and diffusion weighted imaging. This study was conducted in accordance with the Declaration of Helsinki and approved by the Regional Ethical Committee of Meyer Hospital (Florence, Italy), number 131/2024. Informed consent for the study was provided by the patient's legal guardians.

Pipeline Overview:

T1-weighted (T1-w), T2-weighted (T2-w) and Diffusion-Weighted Imaging (DWI) acquisitions from a single pediatric subject (aged between 1y and 3y), provided by IRCCS Stella Maris, were processed to create the digital brain model. Due to the clinical data not meeting the mandatory requirements of the 'TVB Image Processing Pipeline' — specifically lacking B0 Reverse Phase Encoding¹² (RPE), having only one non-zero B-value, and missing FLASH acquisition¹⁴ — a tailored pipeline was developed. This customized pipeline integrates multiple tools and software packages to achieve the necessary preprocessing and analysis:

- **FSL and ANTs:** Employed for preprocessing steps such as denoising, unwarping, removal of eddy currents, bias correction, and brain mask generation.
- **MRtrix3:** Utilized for tractography, constructing the connectome, computing structural connectivity (SC) weights, mean tract lengths and conduction velocity.
- **FreeSurfer:** Used for anatomical reconstruction, cortical mesh creation, and segmentation/parcellation of the brain.
- **SimNIBS:** Applied to create the Head Model and the EEG Forward Leadfield Matrix
- **MNE (MNE-Python):** Implemented to generate inputs for **The Virtual Brain** (TVB).

This integrated approach ensures that the subject-specific data, despite its limitations, can be effectively processed and used as inputs for advanced brain modeling and analysis within TVB.

MRI acquisition:

The subject data were acquired in a clinical setting at IRCSS Stella Maris in Pisa (Italy). MRI was acquired with a 3T General Electric scanner; T1 with voxel size of 1.0mm by 0.4297 mm by 0.4297 mm, TR=2471.42ms, TE=3.824ms; T2 with voxel size of 0.799805mm by 0.5 mm by 0.5 mm; DTI with voxel size of 0.9375mm by 0.9375mm by 5.0mm 30 gradient encoding directions and B=1000s/mm² TR=3100ms, TE=59.6ms.

MRI Preprocessing:

The preprocessing phase involved several critical steps to prepare the imaging data for further analysis^{9,54}. Initially, the DWI data underwent denoising to remove noise artifacts and removal of Gibbs' ringing artifacts, followed by unwarping to correct for geometric distortions⁵⁵. Concerning this latter, since clinical data were lacking the Reverse Phase Encoding, a synthetic RPE was created using SynB0-DISCO^{12,56}.

Eddy current-induced distortions were also corrected⁵⁷, and bias correction was applied to mitigate intensity inhomogeneities⁹. A brain mask was generated to isolate the brain from non-brain tissues. Preprocessing steps were performed using FSL⁵⁸ and ANTs⁵⁹ softwares. A possible different segmentation of subcortical structures (by using Freesurfer Infants⁶⁰) was also tested with the current subject, with no relevant differences in segmentation results. Images were subsequently preprocessed following the preprocessing pipeline outlined in Andy's Brain Book (Jahn, 2022. doi:10.5281/zenodo.5879293).

Constrained Spherical Deconvolution:

To characterize the white matter structure, a constrained spherical deconvolution (CSD) was performed¹⁰. This process included estimating the white matter, cerebrospinal fluid and gray matter response functions for the subject, generating fiber orientation densities⁶¹ (FODs) and normalizing them. Due to the limited diffusion data available (only one gradient direction with B=1000s/mm²), the response functions were generated using MRtrix3Tissue (<https://3Tissue.github.io>), a fork of MRtrix3⁶².

Tissue Boundaries and Coregistration:

Following deconvolution, tissue boundaries were created, and the diffusion images were co-registered with the anatomical T1-weighted images⁶³. The boundary interface between gray matter and white matter was identified and used as a seed region for streamline generation through Anatomically Constrained Tractography¹³ (ACT). This step was performed to ensure that the tractography results were anatomically accurate, providing a robust foundation for subsequent connectivity analysis and brain modeling.

Probabilistic Tractography:

Probabilistic tractography⁶⁴ was performed using Constrained Spherical Deconvolution¹⁰ coupled with Anatomically Constrained Tractography¹³ and Spherical-deconvolution Informed Filtering of Tractograms⁶⁵ (SIFT2). Dynamic seeding was also employed to ensure robust streamline generation. The maximum length of fibers was set to 250 mm, following the guidelines of the TVB Image Processing Pipeline⁶⁶. We chose a number of streamlines proportional to the numerosity of

regions considered in the parcellation (1000 streamlines for each pair of regions). Specifically, this resulted in a number of 71.631 million streamlines for the adopted HCPMMP1³⁵ atlas.

Anatomical Reconstruction:

Anatomical reconstruction was conducted using FreeSurfer's `recon-all` command, which performed cortical mesh creation and segmentation/parcellation using HCPMMP1 atlas. The `-autorecon1` command executed the initial five steps, with tailored inputs to improve subsequent skull extraction. Brainmask was checked and small manual adjustments performed. After the `-autorecon2` the white matter mask was checked and slightly edited followed by autorecon2-wm; at last, the `-autorecon3` command completed the remaining steps of the anatomical reconstruction process.

Structural Connectivity Weights, Mean Tract Length:

We employed MRtrix3⁶² to generate the connectome and calculate structural connectivity (SC) weights, intended as weighted (since SIFT2 was used) number of streamlines connecting two regions, and mean tract length

Myelin volume fraction, g-ratio and conduction velocity:

We conducted a comprehensive preprocessing and analysis of MRI data to derive conduction and myelination properties⁶⁷. It has been followed the methodology from Newman et al.¹⁶. Initially, the T2-w was co-registered to T1-w.

Brain extraction was performed on the T1-w and T2-w images using the bet function from FSL with specific fractional intensity and gradient threshold settings to generate brain-only images and their masks. The intensity values of these images were then standardized. Minimum and maximum intensity values of T1 and T2 images were computed using fsstats, and the images were rescaled to a 0-100 range via fslmaths.

Subsequently, a T1/T2 ratio map was generated. We then extracted fixels ("fiber population within a voxel") from the normalized wfod using MRtrix3 fod2fixel and computed the sum of the Apparent Fiber Density (AFD) using fixel2voxel.

Coregistration and regridding was achieved by aligning the fixel-derived apparent fiber density³⁶ (AFD) sum map with the T1/T2 map using the flirt command, followed by transformation matrix conversion (transformconvert) and regridding (mrtransform, mrgrid).

Calculation of the myelin volume fraction¹¹ (MVF) and axon volume fraction (AVF), are derived by the method described by Mohammadi and Callaghan⁶⁸, computing MVF/AVF ratios and generating a voxel-wise g-ratio map using fslmaths. The equation for the g-ratio is:

$$g = \frac{d}{D} = \sqrt{1 - \frac{MVF}{MVF - AVF}} \quad (1)$$

where (d) is the axon diameter and (D) is the external myelinated fiber diameter.

This g-ratio map was subsequently upper-thresholded. We further computed the voxel-aggregated conduction velocity (CV) based on the derived g-ratio and AVF using the model by Berman, Filo, and Mezer¹⁵, according to the equation:

$$CV = AVF \sqrt{-\ln(g)} \quad (2)$$

The CV map was then coregistered with diffusion-weighted imaging (DWI) to align with streamlines from tractography.

Finally, the CV map was sampled along the streamlines of tractography to extract mean velocity per streamline (*tcksample*) and from this latter the CV matrix was calculated (*tck2connectome*).

Cortex Surface Processing:

In order to reduce computational time, the cortical mesh was downsampled from FreeSurfer's original fsaverage(7) with 327680 triangles to the fsaverage5, which has 20,484 vertices and 40,960 triangles (for comparison, the TVB default cortical mesh² uses 16,384 vertices and 32,760 triangles).

Reconstruction of Head Model and Calculation of anatomically-accurate EEG Forward Leadfield Matrix:

The replication of EEG signals requires projecting neural activity from the high-density cortical mesh to the electrodes on the scalp surface, computing a mathematical quantity called lead-field matrix. The lead-field matrix is a $n_{vertices} \times n_{electrodes}$ matrix, whose entries encode how the activity of each vertex in the cortical mesh can influence the signal in each EEG electrode. We implemented a pipeline to determine patient-specific lead-field matrix based on patient anatomy. Anatomical tissues between cortex and scalp electrodes were reconstructed with the SIMNIBS software (see Methods).

In order to calculate the EEG Forward Problem Leadfield Matrix, it has been first reconstructed the Head of the subject with 12 tissues using SimNIBS4.1.0 : the T1-w and coregistered/regridded T2-w were used as input to *charm* command together with the Freesurfer anatomical reconstruction folder and the initial transformations (the SimNIBS template was coregistered to the T1-w using Freesurfer to find the Transformation matrix used to initialize the affine registration of the SimNIBS template to the subject MR), then also the DWI image was used with the *dwi2cond* command (based on FSL *dtifit*), in order to calculate the tensors necessary to compute the anisotropic properties of conductivity for GM and WM (see Supplementary Figure S5) using the "Volume Normalized" algorithm as shown by Opitz et al¹⁷

It has been chosen an EEG cap from Neuroelectrics following the 10-10 system⁶⁹ and with 64 channels (see Figure 1) using electrodes with circular shape (10mm diameter and 4mm of saline gel). This latter, together with the Head Model, the conductivities shown in Supplementary Figure S2, and the calculated anisotropic properties, allowed to calculate the EEG Leadfield Matrix using the *TDCSLEADFIELD* algorithm.

TVB Format Conversion

The next step involved the use of MNE to rearrange the cortical mesh and prepare inputs for The Virtual Brain (TVB). A Python script, derived from `convert2TVB_format.py`⁶⁶, was employed to achieve this.

The script generates for each vertex the pertaining region, hemisphere, and assigns a flag indicating the vertex to be either in a cortical or in a non-cortical area. We computed a mid-thickness cortical surface by averaging the white and pial surfaces (this in order to match one to one the SimNIBS cortex), with a dummy region for non-cortical areas⁷⁰. Also it writes the 'Leadfield' Matrix in TVB format, convert the vertices coordinates from Freesurfer system (ras-*tkr*) to image coordinates (Subject World) and write the cortex file in zipped TVB format (normals, triangles, vertices), then converts electrodes coordinates (again from FS *ras-tkr* to image coordinates "Subject World") and write the EEG sensors locations in TVB format, write in zipped TVB format the SC weight, Mean Tract Length, Mean Velocity, Region Centres, Average Orientation (mean normal of vertices for a triangle).

As last step, the local connectivity matrix of the cortical surface, providing the weight of the lateral connectivity between cortical columns (assumed as instantaneous), was computed using a gaussian distribution (where amplitude and sigma were chosen after the parameter fitting) applied to the geodesic distance on the mid-thickness cortex. Also, conversion to h5 format was performed in order to use the visualizer of TVB.

To compare the anatomically-accurate forward solution with current standard methodology, we created the BEM model of the patient's with Brainstorm, importing the result of SimNIBS *charm* and then considering only the interfaces between brain tissue, skull and scalp. Default conductivity values (measured as conductance over length, S/m) were assigned to different volumes: 0.465 S/m for brain compartment, 0.008 S/m for skull compartment, and 0.465 S/m for scalp compartment⁷¹ (this because for BEM method it is the ratio between compartments driving the results, which is normally 1:1/80:1,¹⁸ however it was reduced in order to account for the young age of the subject). The leadfield matrix was computed with OpenMEEG⁴³. BEM surfaces and electrodes coordinates were converted from the FreeSurfer system (*ras-tkr*) to the image coordinate system (*ras-scanner*), and the lead field matrix was converted into TVB format with the same procedure adopted for the anatomically-accurate one.

Importing Input Files into The Virtual Brain and Creating Brain Simulation:

SC data, conduction velocity, Region mapping, cortical vertices, electrode positions and local connectivity were imported in TVB to perform the simulations (h5 were created for import and visualization purposes in TVB GUI). Activity on cortical vertices was simulated by using the Jansen-Rit model³⁷. The Jansen-Rit cortical was chosen for its capacity of replicating EEG signals³⁷. Subcortical structures were minimally segmented and modeled using the same Jansen-Rit model as the cortex. The simulation of cortical surface activity allows the modeling and analysis of both local and global connectivity², as both are reported to be aberrant in ASD^{26,30,72,73}. An anatomically-accurate forward solution was used to compute virtual EEG from simulated activity⁷⁴. As already mentioned, the Local connectivity was computed from geodesic brain distance by using a gaussian kernel, whose extension and strength was determined by parameter fitting.

Parameter Fitting with Experimental EEG Recording:

To tune model parameters using the experimental EEG recording, a hierarchical grid search procedure was implemented (see Results). Parameters varied including white matter speed, SC weights scaling, sigmoidal Jansen-Rit coupling, local connectivity, and Jansen-Rit model parameters. For each combination of parameters, a simulated EEG was generated, with a total of 1480 simulations.

Simulated signals were fitted to experimental recordings, based on a loss function that computes the difference in functional connectivity and power spectral densities. The two loss functions are:

$$Loss_{FC} = 1 - Corr(FC_{exp}, FC_{sim}) \quad (3)$$

$$Loss_{PSD} = 1 - [\log_{10}(PSD_{exp}) - \log_{10}(PSD_{sim})] \quad (4)$$

Where *Corr* is the standard Pearson correlation. In the $Loss_{PSD}$ function, the PSD were truncated at 20 Hz (this in order to improve the efficacy of the loss function in locating the best fits). The best match between simulated and actual EEG was computed using these metrics, thus identifying the optimal parameter values.

Experimental EEG acquisition and pre-processing:

EEG data were acquired using the Micromed System Plus Evolution with a sampling rate of 256 Hz, utilizing a 21-electrode sensor system adhering to the standard 10-20 layout. The recorded signals were imported into EEGLAB⁷⁵ for further processing. Initially, the resting state epoch was extracted from the continuous EEG recording to capture the subject's brain activity in a relaxed state without any task-specific stimuli. The EEG data were then filtered in the 1-40 Hz range using a Hamming windowed sinc filter, following an initial high-pass filtering at 1 Hz to remove slow drifts and baseline wander, ensuring the relevant frequency components were retained for analysis.

To enhance the quality of the EEG signals, the *cleanrawdata* plugin in EEGLAB was used to identify and remove bad channels and segments. Channels that were silent or flat for more than 5 seconds, exhibited a large amount of noise based on their standard deviation (with a rejection threshold of 4 standard deviations), or had low-frequency signals poorly correlated with nearby channels (using a threshold value of 0.8) were removed. This resulted in a final number of 16 channels selected for further experimental analysis. Additionally, bad portions of the data series were rejected using the Artifact Subspace Reconstruction⁷⁶ algorithm for regions exceeding 20 times the standard deviation of the calibration data, and further rejection was based on how many channels within a specified time range exceeded a standard deviation threshold, with a maximum out-of-bound channel percentage set at 25%.

The remaining data were re-referenced using the average reference method⁷⁷, averaging the signals across all electrodes and subtracting this mean from each electrode. Independent Component Analysis (ICA) was then performed using the infomax algorithm to decompose the EEG signals into independent components, which were classified using the ICLabel plugin⁷⁸. Components identified as artifacts (e.g., eye blinks, muscle activity, or electrical noise) were removed from the dataset, ensuring that only clean EEG signals were retained for further analysis. This preprocessing ensured a robust foundation for accurate characterization of neural activity and the validity of subsequent analyses.

EEG features: power spectral distribution:

We computed the power spectral distribution (PSD) using Welch's method, dividing the signal into overlapping segments, applying windowing, and taking the discrete Fourier transform to obtain power spectrum estimates. These estimates were averaged to reduce variance, producing a binned array of power spectral densities for each frequency.

Oscillations in neural data are embedded within aperiodic activity, typically following a 1/f distribution. Traditionally, this aperiodic component was disregarded or treated as noise. However, variations in aperiodic activity are now recognized as potential biological indicators for disease, aging, and development. Therefore, we parametrized PSDs into periodic and aperiodic components using the FOOOF (fitting oscillations and one-over f) algorithm⁷⁹.

FOOOF models the PSD as a combination of $1/f$ activity L and N periodic components G_n :

$$\log(\text{PSD}(f)) = L + \sum_n G_n$$

With periodic components G_n modeled with Gaussians and the aperiodic component L is modeled as:

$$L(f) = b - \log(k + f^\alpha)$$

where b is the broadband offset, α is the exponent, and k is the knee parameter.

FOOOF involves initial fitting of the aperiodic component, detrending the spectrum, detecting and fitting periodic components iteratively, and combining these fits to create the model, computing goodness-of-fit scores. If the knee parameter $k = 0$, aperiodic exponents are more comparable and interpretable, serving as potential biomarkers.

EEG features: functional connectivity:

Functional connectivity was computed with standard Pearson correlation, considering only significant ($p < 0.05$, Bonferroni correction) values. The choice of the Pearson correlation metric was motivated by the limited number of electrodes³⁹. Furthermore, we adopted this metric to prove that the FEDE pipeline is able of replicating also FC measures heavily affected by volume conduction⁸⁰.

Acknowledgments:

We thank Adriana di Martino, Ting Xu, Emiliano Ricciardi, Benjamin T Newman, Jan Paul Triebkorn, Jurek Timo Hofsähs and Marco Pagani for fruitful discussions on different aspects of the pipeline.

Fundings:

LGA, AAV and AM were supported by #NEXTGENERATIONEU (NGEU) and funded by the Ministry of University and Research (MUR), National Recovery and Resilience Plan (NRRP), project EBRAINS-Italy (IR0000011) - European Brain ReseArch INfrastructureS-Italy (DN. 101 16.06.2022), project MNESYS (PE0000006) – A Multiscale integrated approach to the study of the nervous system in health and disease (DN. 1553 11.10.2022), and project Fit4MedRob—Fit for Medical Robotics Grant (# PNC0000007).

LM and JC were supported by #NEXTGENERATIONEU (NGEU) and funded by the Ministry of University and Research (MUR), National Recovery and Resilience Plan (NRRP), project MNESYS (PE0000006) – A Multiscale integrated approach to the study of the nervous system in health and disease (DN. 1553 11.10.2022),

AR was partly funded by the National Institute for Nuclear Physics (INFN, CSN5) within the next_AIM project and by the European Commission under the NextGeneration EU with the PNRR M4C2 Inv. 1.3, PE00000013 FAIR project (Spoke 8). SC was partially supported by the Italian Ministry of Health (Grant Ricerca Corrente 2025).

EB and SC were partially supported by the Italian Ministry of Health (Grant Ricerca Corrente 2025).

Disclosures:

The authors declare they have no conflict of interest to declare.

Author Contributions:

MF conceived and performed the simulations and analysis and ideated the MRI preprocessing pipeline; LGA conceived and performed the simulations and analysis and wrote the paper; LM preprocessed experimental EEG data and performed the simulation and analysis; JC performed the simulation and analysis; EB and SC collected experimental MRI and EEG data; AR conceived the simulations and analysis; AAV conceived and performed the simulations and analysis; AM conceived the simulations and analysis and obtained the fundings that financed the project.

References:

1. Sanz Leon P, Knock SA, Woodman MM, et al. The Virtual Brain: a simulator of primate brain network dynamics. *Frontiers in neuroinformatics*. 2013;7:10.
2. Sanz-Leon P, Knock SA, Spiegler A, Jirsa VK. Mathematical framework for large-scale brain network modeling in The Virtual Brain. *Neuroimage*. 2015;111:385-430.
3. Koller DP, Schirner M, Ritter P. Human connectome topology directs cortical traveling waves and shapes frequency gradients. *Nat Commun*. 2024;15(1):3570. doi:10.1038/s41467-024-47860-x
4. Schirner M, Deco G, Ritter P. Learning how network structure shapes decision-making for bio-inspired computing. *Nat Commun*. 2023;14(1):2963. doi:10.1038/s41467-023-38626-y
5. Sip V, Hashemi M, Dickscheid T, Amunts K, Petkoski S, Jirsa V. Characterization of regional differences in resting-state fMRI with a data-driven network model of brain dynamics. *Science Advances*. 2023;9(11):eabq7547. doi:10.1126/sciadv.abq7547
6. Triebkorn P, Stefanovski L, Dhindsa K, et al. Brain simulation augments machine-learning-based classification of dementia. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*. 2022;8(1):e12303. doi:10.1002/trc2.12303
7. Stefanovski L, Triebkorn P, Spiegler A, et al. Linking Molecular Pathways and Large-Scale Computational Modeling to Assess Candidate Disease Mechanisms and Pharmacodynamics in Alzheimer's Disease. *Front Comput Neurosci*. 2019;13:54. doi:10.3389/fncom.2019.00054
8. Wang HE, Woodman M, Triebkorn P, et al. Delineating epileptogenic networks using brain imaging data and personalized modeling in drug-resistant epilepsy. *Science Translational Medicine*. 2023;15(680):eabp8982. doi:10.1126/scitranslmed.abp8982
9. Tax CMW, Bastiani M, Veraart J, Garyfallidis E, Okan Irfanoglu M. What's new and what's next in diffusion MRI preprocessing. *NeuroImage*. 2022;249:118830. doi:10.1016/j.neuroimage.2021.118830
10. Jeurissen B, Tournier JD, Dhollander T, Connelly A, Sijbers J. Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data. *NeuroImage*. 2014;103:411-426. doi:10.1016/j.neuroimage.2014.07.061
11. West KL, Kelm ND, Carson RP, Gochberg DF, Ess KC, Does MD. Myelin volume fraction imaging with MRI. *NeuroImage*. 2018;182:511-521. doi:10.1016/j.neuroimage.2016.12.067
12. Schilling KG, Blaber J, Hansen C, et al. Distortion correction of diffusion weighted MRI without reverse phase-encoding scans or field-maps. *PLOS ONE*. 2020;15(7):e0236418. doi:10.1371/journal.pone.0236418
13. Smith RE, Tournier JD, Calamante F, Connelly A. Anatomically-constrained tractography: Improved diffusion MRI streamlines tractography through effective use of anatomical information. *NeuroImage*. 2012;62(3):1924-1938. doi:10.1016/j.neuroimage.2012.06.005
14. Helms G, Dathe H, Kallenberg K, Dechent P. High-resolution maps of magnetization transfer with inherent correction for RF inhomogeneity and T1 relaxation obtained from 3D FLASH MRI. *Magnetic Resonance in Medicine*. 2008;60(6):1396-1407. doi:10.1002/mrm.21732

15. Berman S, Filo S, Mezer AA. Modeling conduction delays in the corpus callosum using MRI-measured g-ratio. *NeuroImage*. 2019;195:128-139. doi:10.1016/j.neuroimage.2019.03.025
16. Newman BT, Jacokes Z, Venkadesh S, et al. Conduction velocity, G-ratio, and extracellular water as microstructural characteristics of autism spectrum disorder. *PLOS ONE*. 2024;19(4):e0301964. doi:10.1371/journal.pone.0301964
17. Opitz A, Windhoff M, Heidemann RM, Turner R, Thielscher A. How the brain tissue shapes the electric field induced by transcranial magnetic stimulation. *NeuroImage*. 2011;58(3):849-859. doi:10.1016/j.neuroimage.2011.06.069
18. Mosher JC, Leahy RM, Lewis PS. EEG and MEG: forward solutions for inverse methods. *IEEE Transactions on Biomedical Engineering*. 1999;46(3):245-259. doi:10.1109/10.748978
19. Nielsen JD, Puonti O, Xue R, Thielscher A, Madsen KH. Evaluating the influence of anatomical accuracy and electrode positions on EEG forward solutions. *NeuroImage*. 2023;277:120259. doi:10.1016/j.neuroimage.2023.120259
20. D'Angelo E, Jirsa V. The quest for multiscale brain modeling. *Trends in Neurosciences*. 2022;45(10):777-790. doi:10.1016/j.tins.2022.06.007
21. Wang HE, Triebkorn P, Breyton M, et al. Virtual brain twins: from basic neuroscience to clinical use. *National Science Review*. 2024;11(5):nwae079.
22. Lord C, Brugha TS, Charman T, et al. Autism spectrum disorder. *Nat Rev Dis Primers*. 2020;6(1):1-23. doi:10.1038/s41572-019-0138-4
23. Sharma SR, Gonda X, Tarazi FI. Autism Spectrum Disorder: Classification, diagnosis and therapy. *Pharmacology & Therapeutics*. 2018;190:91-104. doi:10.1016/j.pharmthera.2018.05.007
24. Bozzi Y, Provenzano G, Casarosa S. Neurobiological bases of autism–epilepsy comorbidity: a focus on excitation/inhibition imbalance. *European Journal of Neuroscience*. 2018;47(6):534-548. doi:10.1111/ejn.13595
25. Rubenstein JLR, Merzenich MM. Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes, Brain and Behavior*. 2003;2(5):255-267. doi:10.1034/j.1601-183X.2003.00037.x
26. Dajani DR, Uddin LQ. Local brain connectivity across development in autism spectrum disorder: A cross-sectional investigation. *Autism Research*. 2016;9(1):43-54. doi:10.1002/aur.1494
27. Kana RK, Libero LE, Moore MS. Disrupted cortical connectivity theory as an explanatory model for autism spectrum disorders. *Physics of Life Reviews*. 2011;8(4):410-437. doi:10.1016/j.plrev.2011.10.001
28. Courchesne E. Brain development in autism: Early overgrowth followed by premature arrest of growth. *Mental Retardation and Developmental Disabilities Research Reviews*. 2004;10(2):106-111. doi:10.1002/mrdd.20020
29. Courchesne E, Campbell K, Solso S. Brain growth across the life span in autism: Age-specific changes in anatomical pathology. *Brain Research*. 2011;1380:138-145. doi:10.1016/j.brainres.2010.09.101
30. Rane P, Cochran D, Hodge SM, Haselgrove C, Kennedy DN, Frazier JA. Connectivity in Autism: A Review of MRI Connectivity Studies. *Harvard Review of Psychiatry*. 2015;23(4):223. doi:10.1097/HRP.0000000000000072

31. Takahashi Y, Murata S, Ueki M, Tomita H, Yamashita Y. Interaction between Functional Connectivity and Neural Excitability in Autism: A Novel Framework for Computational Modeling and Application to Biological Data. *Comput Psychiatr.* 7(1):14-29. doi:10.5334/cpsy.93
32. Askari E, Setarehdan SK, Sheikhan A, Mohammadi MR, Teshnehlab M. Computational model for detection of abnormal brain connections in children with autism. *JIN.* 2018;17(3):237-248. doi:10.31083/JIN-180075
33. Rosenberg A, Patterson JS, Angelaki DE. A computational perspective on autism. *Proceedings of the National Academy of Sciences.* 2015;112(30):9158-9165. doi:10.1073/pnas.1510583112
34. O'Laughlin C, Thagard P. Autism and Coherence: A Computational Model. *Mind & Language.* 2000;15(4):375-392. doi:10.1111/1468-0017.00140
35. Huang CC, Rolls ET, Feng J, Lin CP. An extended Human Connectome Project multimodal parcellation atlas of the human cortex and subcortical areas. *Brain Struct Funct.* 2022;227(3):763-778. doi:10.1007/s00429-021-02421-6
36. Raffelt D, Tournier JD, Rose S, et al. Apparent Fibre Density: A novel measure for the analysis of diffusion-weighted magnetic resonance images. *NeuroImage.* 2012;59(4):3976-3994. doi:10.1016/j.neuroimage.2011.10.045
37. Jansen BH, Rit VG. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns.
38. Windhoff M, Opitz A, Thielscher A. Electric field calculations in brain stimulation based on finite elements: An optimized processing pipeline for the generation and usage of accurate individual head models. *Human Brain Mapping.* 2013;34(4):923-935. doi:10.1002/hbm.21479
39. Bastos AM, Schoffelen JM. A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls. *Front Syst Neurosci.* 2016;9. doi:10.3389/fnsys.2015.00175
40. Sohal VS, Rubenstein JLR. Excitation-inhibition balance as a framework for investigating mechanisms in neuropsychiatric disorders. *Mol Psychiatry.* 2019;24(9):1248-1257. doi:10.1038/s41380-019-0426-0
41. Davis G, Plaisted-Grant K. Low endogenous neural noise in autism. *Autism.* 2015;19(3):351-362. doi:10.1177/1362361314552198
42. Sorrentino P, Pathak A, Ziaemehr A, et al. The virtual multiple sclerosis patient. *iScience.* 2024;27(7). doi:10.1016/j.isci.2024.110101
43. Gramfort A, Papadopoulo T, Olivi E, Clerc M. OpenMEEG: opensource software for quasistatic bioelectromagnetics. *Biomedical engineering online.* 2010;9:1-20.
44. Ruiz-Gómez SJ, Hornero R, Poza J, Maturana-Candelas A, Pinto N, Gómez C. Computational modeling of the effects of EEG volume conduction on functional connectivity metrics. Application to Alzheimer's disease continuum. *J Neural Eng.* 2019;16(6):066019. doi:10.1088/1741-2552/ab4024
45. Griffiths JD, Wang Z, Ather SH, et al. Deep Learning-Based Parameter Estimation for Neurophysiological Models of Neuroimaging Data. Published online May 19, 2022:2022.05.19.492664. doi:10.1101/2022.05.19.492664
46. Honey CJ, Sporns O, Cammoun L, et al. Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences.* 2009;106(6):2035-2040.

47. Cohen JD, Daw N, Engelhardt B, et al. Computational approaches to fMRI analysis. *Nat Neurosci*. 2017;20(3):304-313. doi:10.1038/nn.4499
48. Demirtaş M, Falcon C, Tucholka A, Gispert JD, Molinuevo JL, Deco G. A whole-brain computational modeling approach to explain the alterations in resting-state functional connectivity during progression of Alzheimer's disease. *NeuroImage: Clinical*. 2017;16:343-354. doi:10.1016/j.nicl.2017.08.006
49. Hansen ECA, Battaglia D, Spiegler A, Deco G, Jirsa VK. Functional connectivity dynamics: Modeling the switching behavior of the resting state. *NeuroImage*. 2015;105:525-535. doi:10.1016/j.neuroimage.2014.11.001
50. Association AP. Diagnostic and Statistical Manual of Mental Disorders. *Text Revision*. Published online 2000. Accessed December 6, 2024. <https://cir.nii.ac.jp/crid/1573950399819987840>
51. Luyster R, Gotham K, Guthrie W, et al. The Autism Diagnostic Observation Schedule—Toddler Module: A New Module of a Standardized Diagnostic Measure for Autism Spectrum Disorders. *J Autism Dev Disord*. 2009;39(9):1305-1320. doi:10.1007/s10803-009-0746-z
52. Green, E., Stroud, L., O'Connell, R., Bloomfield, S., Cronje, J., Foxcroft, C., Hurter, K., Lane, H., Marais, R., Marx, C., McAlinden, P., Paradise, R., & Venter, D. *Griffiths Scales of Child Development, 3rd Ed. Part II: Administration and Scoring*. Hogrefe; 2016.
53. Schopler E, Van Bourgondien ME, Wellman GJ, Love SR. *The Childhood Autism Rating Scale (CARS2) 2nd Ed*. Western Psychological Services; 2010.
54. Tustison NJ, Cook PA, Klein A, et al. Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements. *NeuroImage*. 2014;99:166-179. doi:10.1016/j.neuroimage.2014.05.044
55. Kybic J, Thevenaz P, Nirkko A, Unser M. Unwarping of unidirectionally distorted EPI images. *IEEE Transactions on Medical Imaging*. 2000;19(2):80-93. doi:10.1109/42.836368
56. Schilling KG, Blaber J, Huo Y, et al. Synthesized b0 for diffusion distortion correction (Synb0-DisCo). *Magnetic Resonance Imaging*. 2019;64:62-70. doi:10.1016/j.mri.2019.05.008
57. Rohde G k., Barnett A s., Basser P j., Marengo S, Pierpaoli C. Comprehensive approach for correction of motion and distortion in diffusion-weighted MRI. *Magnetic Resonance in Medicine*. 2004;51(1):103-114. doi:10.1002/mrm.10677
58. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. *NeuroImage*. 2012;62(2):782-790. doi:10.1016/j.neuroimage.2011.09.015
59. Avants, B. B., Tustison, N., & Song, G. (2009). Advanced normalization tools (ANTs). *Insight j*, 2(365), 1-35.
60. Zöllei L, Iglesias JE, Ou Y, Grant PE, Fischl B. Infant FreeSurfer: An automated segmentation and surface extraction pipeline for T1-weighted neuroimaging data of infants 0–2 years. *NeuroImage*. 2020;218:116946. doi:10.1016/j.neuroimage.2020.116946
61. Tournier JD, Calamante F, Connelly A. Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage*. 2007;35(4):1459-1472. doi:10.1016/j.neuroimage.2007.02.016

62. Tournier JD, Smith R, Raffelt D, et al. *MRtrix3*: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage*. 2019;202:116137. doi:10.1016/j.neuroimage.2019.116137
63. Le Bihan D, Mangin JF, Poupon C, et al. Diffusion tensor imaging: Concepts and applications. *Journal of Magnetic Resonance Imaging*. 2001;13(4):534-546. doi:10.1002/jmri.1076
64. Behrens TEJ, Berg HJ, Jbabdi S, Rushworth MFS, Woolrich MW. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage*. 2007;34(1):144-155. doi:10.1016/j.neuroimage.2006.09.018
65. Smith RE, Tournier JD, Calamante F, Connelly A. SIFT: Spherical-deconvolution informed filtering of tractograms. *NeuroImage*. 2013;67:298-312. doi:10.1016/j.neuroimage.2012.11.049
66. Schirner M, Domide L, Perdakis D, et al. Brain simulation as a cloud service: The Virtual Brain on EBRAINS. *NeuroImage*. 2022;251:118973. doi:10.1016/j.neuroimage.2022.118973
67. Mancini M, Karakuzu A, Cohen-Adad J, Cercignani M, Nichols TE, Stikov N. An interactive meta-analysis of MRI biomarkers of myelin. Jbabdi S, Baker CI, Jbabdi S, Does M, eds. *eLife*. 2020;9:e61523. doi:10.7554/eLife.61523
68. Mohammadi S, Callaghan MF. Towards *in vivo* g-ratio mapping using MRI: Unifying myelin and diffusion imaging. *Journal of Neuroscience Methods*. 2021;348:108990. doi:10.1016/j.jneumeth.2020.108990
69. Homan RW, Herman J, Purdy P. Cerebral location of international 10–20 system electrode placement. *Electroencephalography and Clinical Neurophysiology*. 1987;66(4):376-382. doi:10.1016/0013-4694(87)90206-9
70. Cardinale F, Chinnici G, Bramero M, et al. Validation of FreeSurfer-Estimated Brain Cortical Thickness: Comparison with Histologic Measurements. *Neuroinform*. 2014;12(4):535-542. doi:10.1007/s12021-014-9229-2
71. Zhang Y, van Drongelen W, He B. Estimation of *in vivo* brain-to-skull conductivity ratio in humans. *Applied Physics Letters*. 2006;89(22):223903. doi:10.1063/1.2398883
72. Kana RK, Uddin LQ, Kenet T, Chugani D, Müller RA. Brain connectivity in autism. *Front Hum Neurosci*. 2014;8. doi:10.3389/fnhum.2014.00349
73. Keown CL, Datko MC, Chen CP, Maximo JO, Jahedi A, Müller RA. Network Organization Is Globally Atypical in Autism: A Graph Theory Study of Intrinsic Functional Connectivity. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*. 2017;2(1):66-75. doi:10.1016/j.bpsc.2016.07.008
74. Sarvas J. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Physics in Medicine & Biology*. 1987;32(1):11.
75. Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*. 2004;134(1):9-21.
76. Chang CY, Hsu SH, Pion-Tonachini L, Jung TP. Evaluation of Artifact Subspace Reconstruction for Automatic EEG Artifact Removal. In: *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. ; 2018:1242-1245. doi:10.1109/EMBC.2018.8512547
77. Yao D. A method to standardize a reference of scalp EEG recordings to a point at infinity. *Physiol Meas*. 2001;22(4):693. doi:10.1088/0967-3334/22/4/305

78. Pion-Tonachini L, Kreutz-Delgado K, Makeig S. ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*. 2019;198:181-197. doi:10.1016/j.neuroimage.2019.05.026
79. Gerster M, Waterstraat G, Litvak V, et al. Separating Neural Oscillations from Aperiodic 1/f Activity: Challenges and Recommendations. *Neuroinform*. 2022;20(4):991-1012. doi:10.1007/s12021-022-09581-8
80. van den Broek SP, Reinders F, Donderwinkel M, Peters M. Volume conduction effects in EEG and MEG. *Electroencephalography and clinical neurophysiology*. 1998;106(6):522-534.