brain sciences

MDPI

Article

Amortization Transformer for Brain Effective Connectivity Estimation from fMRI Data

Zuozhen Zhang, Ziqi Zhang, Junzhong Ji and Jinduo Liu *

The Beijing Municipal Key Laboratory of Multimedia and Intelligent Software Technology, Beijing Institute of Artificial Intelligence, Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China; zzz3582@emails.bjut.edu.cn (Z.Z.); ziqiz447@gmail.com (Z.Z.); jjz01@bjut.edu.cn (J.J.)
* Correspondence: jinduo@bjut.edu.cn

Abstract: Using machine learning methods to estimate brain effective connectivity networks from functional magnetic resonance imaging (fMRI) data has garnered significant attention in the fields of neuroinformatics and bioinformatics. However, existing methods usually require retraining the model for each subject, which ignores the knowledge shared across subjects. In this paper, we propose a novel framework for estimating effective connectivity based on an amortization transformer, named AT-EC. In detail, AT-EC first employs an amortization transformer to model the dynamics of fMRI time series and infer brain effective connectivity across different subjects, which can train an amortized model that leverages the shared knowledge from different subjects. Then, an assisted learning mechanism based on functional connectivity is designed to assist the estimation of the brain effective connectivity network. Experimental results on both simulated and real-world data demonstrate the efficacy of our method.

Keywords: amortization learning; transformer; brain effective connectivity; functional magnetic resonance imaging

check for updates

Citation: Zhang, Z.; Zhang, Z.; Ji, J.; Liu, J. Amortization Transformer for Brain Effective Connectivity Estimation from fMRI Data. *Brain Sci.* **2023**, *13*, 995. https://doi.org/10.3390/brainsci13070995

Academic Editor: Chantal Delon-Martin

Received: 2 June 2023
Revised: 18 June 2023
Accepted: 22 June 2023
Published: 25 June 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Recently, there has been an increasing interest in brain functional network analysis. A brain functional network can be represented as a graph structure naturally, which consists of brain regions as nodes and connections among brain regions as edges [1]. According to the characteristics of functional integration and separation in the brain, brain connectivity is typically divided into two categories: functional connectivity (FC), and effective connectivity (EC).

With the development of functional magnetic resonance imaging (fMRI) technology, it has become a valuable tool for studying the workings of the human brain and has been used to investigate a wide range of cognitive processes. fMRI is a non-invasive neuroimaging technique to measure changes in blood flow and oxygenation levels in the brain that allows researchers to visualize brain activity in response to various stimuli or tasks. Both functional and effective connectivity can be generated from functional magnetic resonance imaging (fMRI) data, FC describes statistical dependencies between brain regions, while EC is the causal influence that one brain region exerts over another [2,3], which characterizes the causality between brain regions. As different individuals exhibit different EC networks, the discrepancies among EC networks offer an effective way to evaluate normal brain functions and the brain injuries related to some neurodegenerative diseases [1], such as Alzheimer's disease [4], Parkinson's disease [5] and autism spectrum disorder [6]. Therefore, the estimation of brain effective connectivity from fMRI data is a critical scientific problem in the investigation of the human brain connectome.

Over the past few years, there has been a discernible surge in research and exploration focused on the utilization of causal learning methods to estimate brain effective connectivity. Generally, these methods can be categorized into two types: traditional machine learning

methods and deep learning methods. Traditional machine learning methods [7–13] have the advantage of fewer parameters of model and are easy to interpret, which mainly include dynamic causal model (DCM) [14,15], structural equation model (SEM) approaches [16,17], Bayesian network (BN) methods [11,12], linear non-Gaussian acyclic model (LiNGAM) methods [8,18], Granger causality (GC) methods [19,20] and so on. However, such methods rely on assumptions about the model or data. With the rapid development of deep learning techniques, several deep learning methods have been applied exploratively to identify brain effect connectivity networks from fMRI data, such as multilayer perceptron (MLP) [21,22], recurrent neural networks (RNN) [23,24], and generative adversarial networks (GAN) [23,25]. These methods are more flexible and scalable, enabling the design of neural networks based on the characteristics of fMRI data. However, these methods are usually necessary to train models for each subject to estimate the brain effective connectivity network, which may lead to difficulties in fully utilizing the shared effective connectivity information between subjects and extremely high computational costs.

In this paper, we propose a novel framework for estimating effective connectivity based on an amortization transformer, named AT-EC. In detail, AT-EC first employs an amortization transformer encoder to estimate the brain EC across different subjects and a decoder to model the dynamics of fMRI time series under the learned brain EC network. This enables us to train an amortized model that leverages the shared knowledge from different subjects. Then, an assisted learning mechanism is designed that uses the statistical dependency information of FC networks to guide the estimation of EC networks. We have tested our model on both simulated and real-world fMRI data, and the experimental results show that the proposed method has certain advantages in performance compared with existing state-of-the-art methods.

The main contributions of this paper can be summarized as follows:

- We propose an end-to-end single amortization model that can model fMRI time series dynamics for different subjects and learn brain EC network. This model allows us to estimate EC with unseen subjects without refitting the model.
- We develop an assisted learning mechanism that uses the brain functional connectivity network as an additional embedding to guide the estimating of the brain EC network.
- Systematic experiments on both simulated and real-world data show that the proposed method achieves better performance compared to some state-of-the-art methods.

2. Related Work and Preliminary

This section presents a related work of brain effective connectivity estimation and provides the preliminary of amortization learning paradigm.

2.1. Brain Effective Connectivity Estimation

In this section, we introduce some state-of-the-art methods for estimating brain effective connectivity from fMRI data and summarize their advantages and disadvantages. We first introduce the traditional machine learning methods, i.e., Bayesian network (BN) methods and Granger causality (GC) methods, and then show the deep learning methods.

The Bayesian network (BN) method [26] employs probabilistic graphical models that can represent causal relationships between brain regions and has been widely used to infer EC from fMRI data. Bayesian networks provide a flexible and expandable framework for searching brain EC networks, but they are computationally demanding and cannot model cyclic or bidirectional EC.

The Granger causality (GC) method [27] calculates how past values of one brain region time series improve the prediction of another brain region time series, and estimate causal relationship (effective connectivity). These methods can learn brain EC networks from fMRI time series data by using regression models. However, the traditional GC method only considers the causal relations between two variables.

With the wide success of deep learning technology, several novel brain EC learning methods based on deep learning technology have been constantly presented. Compared

to traditional machine learning methods, deep learning methods are better at handling highly noisy and non-linear data. In detail, Liu et al. [25] present a brain EC learning algorithm based on generative adversarial networks (EC-GAN). The method uses the generative adversarial process to obtain brain EC networks, where the generator uses structural equation models (SEM) to quantify causal relationships among brain regions and the discriminator distinguishes between the joint distribution of real and generated fMRI time series. Li et al. [28] design a CR-VAE method, which learns EC by an encoder and multi-head decoders under the general concept of Granger causality. Zou et al. [29] offer an EC estimation method based on a controllable variational autoencoder (CVAEEC). This method first employs an encoder to encode the fMRI time series input as latent variables and then uses SEM to learn brain EC network during the generation of the fMRI time series data.

Although these methods achieve favorable results, they fail to fully utilize the shared effective connectivity information between subjects. To more accurately estimate brain EC, it is necessary to extract and utilize the knowledge shared between different subjects.

2.2. Amortization Learning

For traditional optimization problems such that:

$$y^{\star}(x) \in \arg\min_{y} f(y; x), \tag{1}$$

where x denotes the input context, y is the output we control, $y^{\star}(x)$ denotes the optimal solution to the optimization problem, which reflects the similarities between problems across various contexts. It is not explicitly defined, but rather emerges from the optimization process and is usually considered unique. However, since the optimization problem (1) usually does not have a closed-form solution, it is usually necessary to use approximate numerical methods during iterative optimization to obtain an approximate optimal solution with extremely high computational cost, such as the directed acyclic graph searching problem [30].

Amortized learning can amortize the cost of solving the optimization problems across many contexts to approximate the solution mapping [31], and has become a widely used modeling technique, particularly in scenarios that need to solve multiple instances of the same problem repeatedly. This approach enables the prediction of solutions to such problems, making use of the structural similarities that exist between them [32].

The solution for the amortization learning can be expressed as follows:

$$\hat{\Phi} = \arg\min_{\Phi} \mathbb{E}[\mathcal{L}(f, \mathcal{Y}, \mathcal{X}, p(x), \hat{y}_{\Phi})], \tag{2}$$

where $\mathcal{X}$ describes the context space, $\mathcal{Y}$ describes the control space, f denotes optimization objectives, $p(x)$ represents the distribution of data, and $\hat{y}_{\Phi}$ denotes the approximate optimal solution, i.e., the mapping from $\mathcal{X}$ to $\mathcal{Y}$, Φ denotes the solution-related parameters, in the context of deep learning, Φ can represent the parameters of a neural network. $\mathcal{L}$ needs to be optimized under all the contexts involved in amortization.

3. Methodology

In this section, we present our proposed novel model, i.e., AT-EC, which can estimate brain effective connectivity from fMRI time series data. Specifically, we first give an overview of the proposed method, and then describe the details of the main components. Finally, we show the description of AT-EC.

3.1. Main Idea

The AT-EC consists of two components: (1) Amortization transformer for fMRI time series dynamics modeling; and (2) estimating EC with FC guidance. First, the amortization transformer takes advantage of shared dynamics among different fMRI subjects to improve

the prediction of brain EC. In detail, the amortization transformer is an encoder–decoder model, the encoder extracts Granger causal relations of fMRI time series, and the decoder predicts the next time step of the fMRI time series data under the estimated Granger causal relationship (brain EC). The concept of amortized learning is utilized to train an amortized model on multiple fMRI subjects, enabling it to learn the brain EC of new unseen subjects. In this way, the smaller the difference between the generated fMRI time series and the original fMRI time series, the more accurate brain EC we can obtain. However, the brain EC network of each subject is different, and even a generalized amortization model has difficulty accurately learning all the brain effective connectivities. Then, to obtain an optimal brain EC network, the FC network of each subject is as embedded weight matrices to guide the estimating of brain EC networks. The structure of the proposed AT-EC is shown in Figure 1.

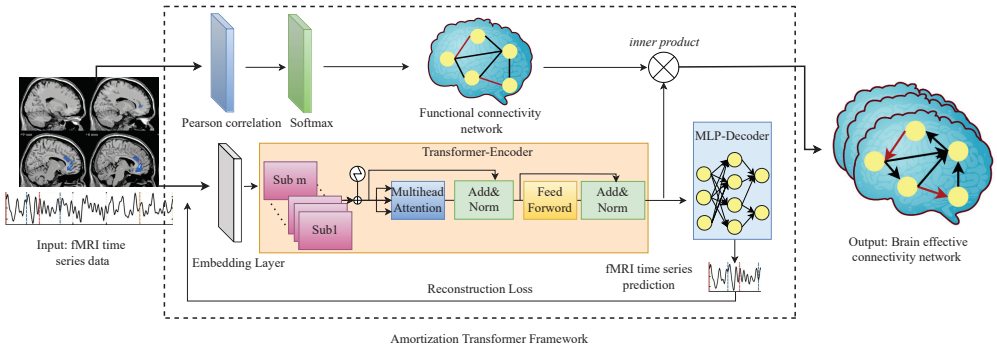


Figure 1. The structure of AT-EC.

3.2. Amortization Transformer for fMRI Time Series Dynamics Modeling

3.2.1. Amortization Transformer Architecture

In this section, we introduce the amortization transformer architecture. The amortization transformer component is an encoder–decoder model, which takes the real fMRI time series data as input and generates brain EC network and fMRI time series data for loss function. We use the idea of amortization learning to train an amortized encoder–decoder model. As the optimization problem described in (2), $\mathcal{X}$ denotes fMRI time series data of different subjects and $\mathcal{Y}$ denotes the generated fMRI time series data, $p(x)$ represents the distribution of fMRI time series data. $\hat{y}_\Phi$ represents the brain EC network and Φ is the parameter of encoder–decoder model. In order to obtain the optimal brain EC network, the optimized loss function $\mathcal{L}$ needs to be minimized among all the subjects involved in the training.

Given the fMRI time series of m subjects with n brain regions x_i $(i = 1, \ldots, n)$ and t length, the input training data X can be represented as:

$$X = (x_1, x_2, \ldots, x_n)^\top \in \mathbb{R}^{m \times n \times t}, \tag{3}$$

where the m subjects may have different brain EC networks, but most ECs are the same. Amortization transformer can learn the knowledge shared between these subjects and does not need to retrain the model when faced with new unseen subjects. Next, we introduce the encoder–decoder components in detail.

3.2.2. Transformer-Based Encoder

The amortization encoder utilized in this paper is based on the transformer model. When dealing with short-length fMRI time series, the transformer model can quickly capture important features of brain region series through multi-head attention mechanisms, and integrate these features into a global representation, thereby better extracting the causal relationships between brain regions. The transformer-based encoder involves first embedding the inputs via a linear layer, followed by processing them through multiple

identical encoder blocks comprising a multi-head self-attention layer and a feed-forward layer; we posit that multi-head self-attention is well-suited for extracting temporal features from fMRI time series data, as it reduces reliance on external information and better captures internal correlations within the fMRI data. The encoder block operations are as follows:

$$\begin{aligned} X' &= \text{Linear}(X) \in \mathbb{R}^{m\times n\times h}, \\ Q_l &= W^Q X'^l + \epsilon^Q, \\ K_l &= W^K X'^l + \epsilon^K, \\ V_l &= W^V X'^l + \epsilon^V, \end{aligned} \tag{4}$$

where X' denotes the embedded input, Linear is a fully connected linear layer that provides a linear transformation of the input X', h denotes the number of hidden layer nodes of the fully connected linear layer, and X'^l expresses the l-th input after dividing the embedding X' into L head self-attention layer. The attention mechanism can be described as a function that takes both a value and a set of key-query pairs as input and produces an output by calculating a weighted sum of the values. Q, K and V describe the query, key, and value, Q, K is used to calculate the similarity of the input features and acquire the attention features with value V. W^Q, W^K and W^V denote the network parameters for the self-attention layer, respectively, ϵ^Q, ϵ^K and ϵ^V are the bias vector. Then, we can obtain Q_l, K_l and V_l, which denote the output of the linear transformation of the input X'^l. So the self-attention can be calculated as follows:

$$Attention_l = \text{softmax}\left(\frac{Q_l K_l^T}{\sqrt{D_{K_l}}}\right) V_l, \tag{5}$$

where Q_l, K_l and V_l denote the query, key, and value vector of the self-attentive calculation of the l-th head, D denotes the number of elements in the last dimension of the query, key, and value vector Q_l, K_l, and V_l, $Attention_l$ describes the l-th head attention vector. Here, the softmax function is used for stable training which has a lower computational cost than the other commonly used sigmoid function. Then, we can collect all L heads of the self-attention vectors as follows:

$$X_{MultiHead} = \text{Concat}(Attention_1, Attention_2, \ldots, Attention_L). \tag{6}$$

Finally, we can obtain the brain EC adjacency matrices $\mathcal{G}$ by putting $X_{MultiHead}$ into a feed-forward layer, which consists of two liner layers and a ReLU activation.

$$\mathcal{G} \in \mathbb{R}^{m\times n\times n} = \text{softmax}(\text{Feedforward}(X_{MultiHead})), \tag{7}$$

where n denotes the number of nodes (brain regions) and m is the number of fMRI subjects for amortization learning.

3.2.3. MLP-Based Decoder

Based on the brain EC network $\mathcal{G}$ estimated by the amortization transformer encoder, we can define the decoder as follows:

$$x_n^{t+1} = g\left(x_n^{\leq t}, \mathcal{G}\right) + \varepsilon_n^{t+1}, t \in [1, T], \tag{8}$$

where ε_n^{t+1} is the exogenous noise variables, t describe the time step of the fMRI time series, T denotes the length of the time series. The decoder function g describes the dynamics of all time series data $x_n \in X$. In practice, the function g of the decoder can be defined as any model, such as a neural network model or predefined nonlinear formulation. Due to the high cost of data collection and pre-processing, the length of fMRI time series is usually

short, so we consider a multilayer perceptron (MLP) with 3 hidden layers and a single activation function $\sigma : \mathbb{R} \to \mathbb{R}$ for the decoder:

$$x_n^{t+1} = \text{Linear}(x_n^{\leq t} \times \mathcal{G}). \tag{9}$$

MLP is a type of feed-forward neural network that performs well in short-term time series prediction tasks. Compared to the commonly used long short-term memory model (LSTM) and recurrent neural network (RNN), MLP has less parameters and is less prone to over-fitting. Therefore, using MLP may be a better choice.

The input of the decoder model is $x_n^{\leq t}$ and their brain EC network $\mathcal{G}$. The purpose is to use the underlying causal relationships between brain regions to generate the time series of each brain region, which can be described as x_n^{t+1}. During the generation of fMRI time series data, the decoder strictly obeys the fundamental principles of Granger causality (GC), which results in a clear understanding of the data-generating process. When the generated fMRI time series data highly match the original fMRI time series data, we can obtain an optimal brain effective connectivity network. Therefore, the amortization transformer encoder and decoder parameters are trained with the reconstruction loss function.

3.2.4. Loss Function

In this section, we present the loss function for the optimization task of AT-EC. The objective loss function is given such that:

$$\begin{aligned} f_{\phi_\star}, g_{\theta_\star} &= \text{argmin}_{f_\phi, f_\theta} \mathcal{L}\big(X, f_\phi, g_\theta\big), \\ \text{where } \mathcal{L}(X, \phi, \theta) &= \sum_{m=1}^{M} \sum_{t=1}^{T-1} \ell\big(X, g_\theta\big(X, f_\phi(X)\big)\big) + r\big(f_\phi(X)\big), \end{aligned} \tag{10}$$

where f_ϕ and g_θ denote the amortization transformer encoder and decoder, respectively, ϕ and θ are the parameters of them, t describes the time step of the fMRI time series, m is the number of the fMRI subjects for amortization learning, X denote the original fMRI time series inputs which collect all dimension in $\mathbb{R}^{m \times n}$. ℓ denotes the metrics of data regression (i.e., evidence lower bound and least squares loss). In this paper, we use the evidence lower bound (ELBO) as the reconstruction loss function described as follows:

$$\ell_{ELBO} = \mathbb{E}_{f_\phi(\mathcal{G}|X)}[\log g_\theta(X \mid \mathcal{G})] - \text{KL}\big[f_\phi(\mathcal{G} \mid X) \| p(\mathcal{G})\big]. \tag{11}$$

The ℓ_{ELBO} loss is a negative log-likelihood value with a KL-Divergence to a prior distribution over $\mathcal{G}$. When training the model, we want to maximize the log-likelihood and keep the KL-Divergence as small as possible. To overcome overfitting and infer sparse effective connectivity networks (causal graphs), we add an L1-regularization sparsity penalty $r\big(f_\phi(X)\big) = \|f_\phi(X)\|_1$ in our loss function.

3.3. Estimating EC with FC Guidance

Existing research suggests that FC can explain the existence of connections between different brain regions, and EC can reveal the influence of these connections on behavioral and cognitive processes [2]. Typically, the study of brain functional connectivity provides the basis for the study of EC. To improve the accuracy of the brain EC network, we propose an assisted learning mechanism to estimate EC with FC guidance.

We first employ the Pearson correlation (PC), which is widely used to calculate the statistical dependencies of other variables, to obtain the skeleton of the causal diagram (FC network) of each subject [33]. Then, we use the FC network to guide the brain EC estimation. Next, we introduce the processes of PC and EC estimation in turn.

Assuming the brain FC network learned by PC is W and $w_{i,j} \in W$, where $w_{i,j}$ describes the statistical correlation of variables (brain region) between x_j and x_j. $w_{i,j}$ is calculated as follows:

$$w_{x_i,x_j} = \frac{\mathrm{cov}(x_i, x_j)}{\sigma_{x_i}\sigma_{x_j}}, \tag{12}$$

where $\mathrm{cov}(.)$ represents the covariance between two variables, and $\sigma_{x_i}, \sigma_{x_j}$ is the standard deviation of the two variable x_i, x_j. The value of $w_{i,j}$ and $w_{j,i}$ should be the same since the FC network has no direction and is not affected by changes in the position or size of the two variables x_i and x_j.

Based on the learned FC network, we can modify the EC network similar to the attention mechanism, and we consider the FC network W as the query-key pair and the EC network learned by the amortization transformer as the value vector. Therefore, the estimating EC with FC Guidance can be obtained by:

$$G' = G \times \mathrm{Softmax}(W), \tag{13}$$

where $G \in \mathcal{G}$ denotes the EC network of a single subject, which has the same dimension with FC network W in $\mathbb{R}^{n\times n}$. Here, the softmax function is used to normalize the FC network W, and the properties of the softmax function can be used to highlight the weights of important brain regions to guide the learning of the EC network. G' is the final brain EC network of the subject.

3.4. Algorithm Description

Above all, the AT-EC algorithm mainly consists of two phases: fMRI time series dynamics modeling phase, and estimating EC with FC guidance phase. The algorithm is formally stated in Algorithm 1.

Algorithm 1 AT-EC

Input: Original fMRI time-series data.
Output: Brain effective connectivity network
Initialization:
m for the number of subjects for amortization learning,
T for the length of time slice to be predicted,
Parameters of the amortization transformer encoder and decoder: f_ϕ and g_θ.
Brain EC Estimation with Amortization Transformer:
for $R - iterations$ **do**
  Learn Granger causal (GC) relations by the amortization transformer encoder as (7);
  Predict the T time slice of the fMRI time series data under the estimated GC by the MLP decoder as (9);
  Compute the reconstruction loss function as (10);
  Update the transformer encoder f_ϕ and MLP decoder g_θ;
end
Learn the FC network by Pearson correlation as (12);
Estimate EC network with the guidance of FC network as (13);
Return: Brain effective connectivity network $\mathcal{G}$.

(1) In the fMRI time series dynamics modeling phase, AT-EC uses m fMRI subjects as an amortization learning task with simultaneous input models for training, then performs model initialization and sets some basic parameters. The fMRI time series of m subjects first enter the encoder simultaneously to obtain m brain EC networks. Then, based on these EC networks, T time slices are predicted by the MLP decoder. Finally, the predicted T time slices are used to calculate the reconstruction loss and guide the update of parameters. In

each iteration, the learned EC network can be constantly optimized with the reduction of loss value. (2) In the estimating EC with FC guidance phase, the algorithm first use PC to learn the FC network for each subject, and then use the FC network to guide the brain EC estimation. Finally, we can obtain the optimal brain EC network for m subjects.

4. Experiment Setting

This section presents our experimental setting, which includes data description, baseline methods, evaluation metrics, and model configuration. To evaluate the performance of AT-EC, we compare AT-EC with other state-of-the-art methods using simulated fMRI data generated from known ground-truth networks to demonstrate its efficacy. Moreover, we also apply AT-EC to public real fMRI data to showcase its potential for practical applications.

4.1. Data Description

4.1.1. Benchmark Simulation Data

The benchmark simulation data we used are supported by Smith et al. [7] and Sanchez-Romero et al. [34], which are generated by dynamic causal models [35] (DCM). Smith dataset (https://www.fmrib.ox.ac.uk/datasets/netsim/index.html) has been widely used in the literature owing to its rich, realistic simulated fMRI data for a wide range of underlying networks. In our experiment, we utilize three types of typical simulation cases to verify the performance of the AT-EC algorithm. In detail, sim1 is set to non-linearity. Sim2 shares the external inputs of a network. Sim3 adds the global mean confound to all nodes. All three datasets have 5 nodes, 50 subjects, 5 arcs, and 200 data points, which are instructive to simulate a real scene because the real scanning session time is usually shorter. The detailed information of the Smith benchmark simulation data is shown in Table 1.

Table 1. Description of the Smith benchmark simulation data.

Dataset	No. Arc	No. Node	Samples	Noise (%)	Other Factors	Session (min)	TR(s)
Sim1	5	5	200	1.0		10	3.00
Sim2	5	5	200	1.0	shared inputs	10	3.00
Sim3	5	5	200	1.0	global mean confound	10	3.00

The simple_network simulated fMRI data (https://github.com/cabal-cmu/feedback-discovery) is an expansion of the Smith dataset which increase in data points of the fMRI time series from 200 to 500 and reduces the original non-Gaussian of the BOLD data [36]. The ground-truth networks of simple_network simulated fMRI data contain different bidirectional structures, which are similar to the actual brain effective connectivity network. In detail, sim4 has the same ground-truth network of sim1, sim5 has more additional bidirectional structures, and sim6 has more nodes and edges. The detailed information of the simple_network benchmark simulation data is shown in Table 2.

Table 2. Description of the simple_network benchmark simulation data.

Dataset	No. Arc	No. Node	Samples	No. Subject	Session (min)	TR(s)
sim4	5	5	500	60	10	1.20
sim5	5	7	500	60	10	1.20
sim6	10	18	500	60	10	1.20

4.1.2. Real fMRI Dataset

The real-resting-state fMRI data of 23 human subjects are obtained from the MTlnet database (https://github.com/shahpreya/MTlnet). These real fMRI data are acquired at TR (Repetition Time) = 1 s, 7 min fMRI sessions for each subject, resulting in 421 data points of the fMRI time series. We consider the following seven regions of interest (ROIs) from the medial temporal lobe, which is referred to [37]. The detailed information of ROIs is shown in Table 3.

Table 3. The ROIs of the real fMRI dataset.

NO.	ROI	Detailed Description
1	$CA1$	Cornu Ammonis1
2	$CA23DG$	Cornu Ammonis 2, 3 and Dentate Gyrus
3	SUB	Subiculum
4	ERC	Entorhinal Cortex
5	$BA35$	Brodmann Areas 35
6	$BA36$	Brodmann Areas 36
7	PHC	Parahippocampal Cortex

4.2. Evaluation Metrics

To evaluate the effectiveness of methods for learning effective connectivity, we employ the following three widely-used graph metrics to analyze the learned results: Precision, Recall, F1-measure ($F1$), and Structural Hamming distance (SHD).

$$Precision = \frac{CA}{TA}, \tag{14}$$

$$Recall = \frac{CA}{|\mathbb{G}|}, \tag{15}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}, \tag{16}$$

$$SHD = EA + MA + RA, \tag{17}$$

where extra arcs (EA), missing arcs (MA), reverse arcs (RA), correct arcs (CA), and total arcs (TA) are obtained from comparison of the learned results and ground-truth networks. And $\mathbb{G}$ is the arc set of ground-truth networks. $|\mathbb{G}|$ is the cardinal number of the set $\mathbb{G}$.

4.3. Baseline Methods

To illustrate the competitiveness of the AT-EC algorithm, we compare AT-EC with the other seven algorithms. These comparison algorithms contain classical machine learning methods and some novel deep learning methods, which are isGC (2017), ACOCTE (2022), EC-GAN (2020), CR-VAE (2023), and CVAEEC (2021) algorithms. The parameter configurations of the corresponding methods are shown in Table 4.

Table 4. Parameter settings of these baseline methods.

Methods	Parameter Settings
isGC [27]	cmp = 4, ARorder = 2, normalize = 1
ACOCTE [26]	$\alpha = 1.0, \beta = 2.0, q_0 = 0.98, \rho = 0.2$
EC-GAN [25]	lr = 0.01, dlr = 0.01, l1 = 5, II = 3, nh = 100, dnh = 100
CR-VAE [28]	context = 20, $\lambda = 0.1$, lr = 0.05, nh = 64
CVAEEC [29]	$\tau_A = 0, \lambda_A = 0, c_A = 1$, lr = 0.003, nh = 64

4.4. Model Configuration

In our framework, our encoder construction first adopts a Transformer encoder module with 6 layers and uses the ReLU activation function. The number of hidden nodes is 2048 and the number of multi-head attention groups is 4. Two feed forward layers are used at the end of the encoder module to identify Granger causal relationships between brain regions. The decoder simply adopts the MLP layer decoder, which is divided into two layers, and each layer contains three fully connected linear layers. We use the Adam optimizer to optimize and fix the learning rate of 0.0025, and the learning rate of test time adaptation logic is 0.05. Sparse punishment can help us more effectively observe

the correct and sparse EC network and our sparse punishment threshold is set at 0.3. For model training, we use five subjects for amortized learning for 200 epochs. When facing new subjects, we only iterate training 10 epochs to obtain an optimal brain EC network. For baseline methods, we run or train each subject using the parameters shown in Table 4 with the same sparse punishment threshold of 0.3.

5. Experimental Results

5.1. Results on Benchmark Simulated fMRI Dataset

To evaluate the performance of the learning methods, we apply five baseline methods and AT-EC to the benchmark simulation datasets, the Smith simulated datasets have 50 subjects and the simple_network simulated datasets have 60 subjects. we run the six methods on every subject to simulate a real scene. We report the mean μ and standard deviation σ of results across all subjects and assess the performance of these six methods based on Precision, Recall, $F1$, and SHD. A method is considered to perform well if it achieves higher precision, recall, and F1 scores, with lower SHD values.

Results on simulation datasets are shown in Table 5. The bold values indicate that the method achieved the best results. Sim2 has the same ground truth as Sim1, but the external neuronal noises are mixed into the nodes. From the results of Sim2, we can notice that traditional machine learning algorithms are more affected by external noise inputs, the precision, recall, and $F1$ values of the IsGC approach have declined somewhat. The ACOCTE approach is a score-based search algorithm that has stronger noise resistance. The deep learning methods are less affected, the EC-GAN, CR-VAE, CVAEEC methods achieve better performance despite the influence of noise. Moreover, compared with other algorithms, AT-EC still maintains its relative advantages over other algorithms.

Table 5. The mean and the standard deviation results of 6 methods on benchmark simulated dataset using single subject data.

Data	Metrics	Methods					
		isGC	**ACOCTE**	**EC-GAN**	**CR-VAE**	**CVAEEC**	**AT-EC**
sim1	Precision	0.36 ± 0.20	0.44 ± 0.34	0.29 ± 0.04	0.22 ± 0.09	**0.60 ± 0.35**	0.52+0.01
	Recall	0.37 ± 0.21	0.27 ± 0.23	**0.93 ± 0.13**	0.41 ± 0.22	0.24 ± 0.15	0.80+0.02
	SHD	4.94 ± 1.65	3.80 ± 1.22	4.90 ± 0.85	6.32 ± 1.26	3.86 ± 0.74	**1.90+0.49**
	$F1$	0.35 ± 0.17	0.32 ± 0.26	0.44 ± 0.05	0.28 ± 0.12	0.32 ± 0.17	**0.62+0.01**
sim2	Precision	0.32 ± 0.20	0.45 ± 0.31	0.28 ± 0.04	0.26 ± 0.14	0.48 ± 0.24	**0.51+0.01**
	Recall	0.34 ± 0.22	0.33 ± 0.22	**0.91 ± 0.14**	0.48 ± 0.22	0.30 ± 0.16	0.82+0.02
	SHD	5.14 ± 1.45	4.00 ± 1.56	5.22 ± 0.84	6.10 ± 1.40	4.00 ± 1.07	**1.90+0.89**
	$F1$	0.31 ± 0.17	0.37 ± 0.25	0.43 ± 0.06	0.32 ± 0.14	0.35 ± 0.15	**0.61+0.00**
sim3	Precision	0.28 ± 0.15	0.52 ± 0.30	**0.66 ± 0.15**	0.24 ± 0.10	0.64 ± 0.30	0.47+0.01
	Recall	0.39 ± 0.21	0.42 ± 0.24	0.63 ± 0.25	0.41 ± 0.21	0.30 ± 0.13	**0.92+0.15**
	SHD	5.66 ± 1.61	3.20 ± 1.39	5.70 ± 1.43	6.14 ± 1.46	3.64 ± 0.73	**2.30+0.61**
	$F1$	0.32 ± 0.15	0.46 ± 0.26	**0.61 ± 0.11**	0.29 ± 0.13	0.39 ± 0.15	0.58+0.00
sim4	Precision	0.25 ± 0.00	**0.69 ± 0.21**	0.25 ± 0.00	0.25 ± 0.00	0.53 ± 0.19	0.64 ± 0.01
	Recall	**1.00 ± 0.00**	0.56 ± 0.18	**1.00 ± 0.00**	**1.00 ± 0.00**	0.30 ± 0.17	0.85 ± 0.01
	SHD	5.00 ± 0.00	2.62 ± 1.28	5.00 ± 0.00	4.98 ± 0.13	3.52 ± 0.89	**1.48 ± 0.78**
	$F1$	0.40 ± 0.00	0.62 ± 0.19	0.40 ± 0.00	0.40 ± 0.00	0.38 ± 0.15	**0.73 ± 0.01**
sim5	Precision	0.35 ± 0.00	**0.76 ± 0.17**	0.52 ± 0.00	0.35 ± 0.00	0.92 ± 0.12	0.66 ± 0.01
	Recall	**1.00 ± 0.00**	0.55 ± 0.12	**1.00 ± 0.00**	**1.00 ± 0.00**	0.54 ± 0.13	0.84 ± 0.01
	SHD	5.00 ± 0.00	3.12 ± 1.53	5.00 ± 0.00	5.00 ± 0.00	2.65 ± 0.84	**1.87 ± 0.72**
	$F1$	0.52 ± 0.00	0.64 ± 0.14	0.52 ± 0.00	0.52 ± 0.00	0.67 ± 0.11	**0.73 ± 0.01**
sim6	Precision	0.20 ± 0.00	0.53 ± 0.11	0.20 ± 0.00	0.33 ± 0.00	**0.86 ± 0.11**	0.59 ± 0.01
	Recall	**1.00 ± 0.00**	0.36 ± 0.07	**1.00 ± 0.00**	**1.00 ± 0.00**	0.42 ± 0.05	0.64 ± 0.01
	SHD	31.00 ± 0.00	11.00 ± 2.32	31.00 ± 0.00	31.00 ± 0.00	**9.97 ± 0.69**	10.23 ± 6.75
	$F1$	0.33 ± 0.00	0.42 ± 0.09	0.33 ± 0.00	0.33 ± 0.00	0.56 ± 0.05	**0.60 ± 0.01**

Sim3 includes global mean confounding factors into the fMRI time series of brain regions, which are variables that are not measured or controlled in a causal relationship. Many causal discovery methods will identify spurious effective connectivity under the

influence of confounding factors. The results of Sim3 indicate that most algorithms are robust to global confounding factors. All other methods outperform or equal the results of Sim1, except the isGC approach. AT-EC achieves the best recall and was second only to $F1$ with EC-GAN, which means AT-EC estimates almost all sides but identifies too many two-way connections.

Sim4 reduces the original non-Gaussian of the BOLD data. From the results of sim4, we can notice that all methods traditional machine learning methods show improved performance, however, the deep learning methods perform worse. Theoretically speaking, deep learning methods have stronger expressive power and adaptability, and they should achieve better performance. We believe that the threshold for the set sparse penalty is too small (0.3), making it difficult to filter out invalid edges. This is confirmed by the small value of precision and the recall of one in the results. Compared to other methods, the AT-EC method still achieves the best results in terms of precision, recall, SHD, and $F1$.

Sim5 has two additional bidirectional structures, which are similar to the real scene since the bidirectional connection structures are common in the brain effective connectivity network. From the results of sim5, we find that ACOCTE and CVAEEC have high precision and a low recall, which means both methods learn brain EC with high accuracy and rarely learn the wrong EC, but miss some ECs. Conversely, EC-GAN and CR-VAE achieve high recall and low precision, indicating that the ECs were basically learned but many are learned incorrectly, probably because these methods identify too many two-way connections. Our method has a high accuracy and recall values, indicating that the learned ECs have few errors and missing.

Sim6 has more nodes and edges, which also are similar to the real scene, since the complex human brain can be divided into many brain regions, usually multiple brain regions collaborate to complete an action. From the results of sim6, we notice that the performance of most methods is declining, while CVAEEC and AT-EC methods are performing well. This may be because, for most methods, it is difficult to learn causal relationships for short-term time series. Thanks to our amortization transformer architecture and FC-guided EC learning mechanism, our approach still achieves the best performance ($F1$).

Overall, deep learning methods are capable of extracting deep features from fMRI data, allowing for more accurate and precise results. AT-EC achieves better performance than state-of-the-art deep learning algorithms since it exploits correlation information across subjects.

5.2. Results on Real Resting-State fMRI Dataset

Unlike simulated data, it is not possible to evaluate the performance of causal search algorithms on real fMRI data using fully defined ground truth. Instead, our evaluation relies on partial knowledge about the structural connections between brain regions in the medial temporal lobe based on existing studies [34,38].

For the real fMRI data, we run AT-EC on every individual subject (each subject has 421 time points) for the seven medial temporal lobe regions of interest of the left and right hemispheres separately. We also do the same on each baseline algorithm for each hemisphere and for individual cases the list of directed edges and their frequency of appearance across the 23 individual subjects. The EC between two brain regions is estimated when we consider edges that appear in 40% of the 23 individual subjects. Figure 2 illustrates the EC networks estimated by AT-EC and the baseline methods from the left hemisphere medial temporal lobe and right hemisphere medial temporal lobe.

In Figure 2f. The effective connectivity networks of the left and right hemisphere medial temporal lobes are similar, but exhibit some differences. These differences are mainly caused by the connections of $CA1$, $CA23DG$, SUB, ERC, $BA35,36$, and PHC. Effective connectivities $CA1 \rightarrow BA36$, $CA23DG \rightarrow ERC$, $SUB \rightarrow CA23DG$, $PHC \rightarrow CA1$, and $SUB \rightarrow BA36$ are in the left hemisphere while $BA36 \rightarrow ERC$ and $PHC \rightarrow ERC$ are in the right hemisphere.

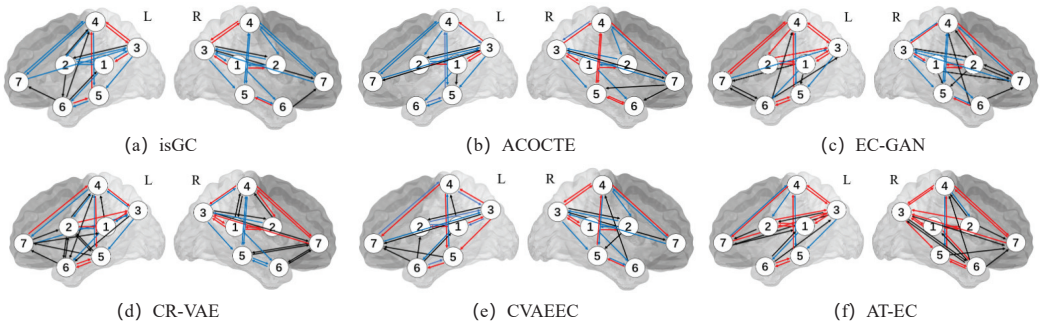


Figure 2. The EC network inferred by 6 methods on real fMRI data. The red lines represent the important connections between brain regions. In contrast, the black lines may indicate incorrect or spurious connections, while the blue lines may represent missing connections that were not captured by the algorithm. The ROIs include: (1) $CA1$; (2) $CA23DG$; (3) SUB; (4) ERC; (5) $BA35$; (6) $BA36$; and (7) PHC.

Compared with the other methods, AT-EC learns the most correct edges (red line) with few missed edges (blue line) in both the left and right hemispheres. As is suggested by Lavenex and Amaral [38], the flow of information from the medial temporal lobe cortices ($BA35, BA36, PHC$) directly into the entorhinal cortex (ERC) and travel to $CA23DG$ to $CA1$, this is the main pathway connecting the medial temporal lobe cortices with the hippocampus. We can find that those important effective connectivities ($CA23DG \leftrightarrow CA1$, $CA1 \leftrightarrow SUB$, $BA35 \leftrightarrow BA36$, $PHC \rightarrow SUB$ and $CA23DG \rightarrow SUB$) are both learned in the left and right hemispheres, and other methods cannot identify them accurately. We also missed some important brain effective connectivities, such as the one-way connection between ERC and $CA23DG$ ($ERC \rightarrow CA23DG$), which is an important connection in the main pathway connecting the medial temporal lobe cortices with the hippocampus. Overall, the new method AT-EC performs better than the state-of-the-art methods and could provide a reliable perspective for the analysis of brain effective connectivity networks.

6. Conclusions

Estimating brain EC from fMRI time series data is still a challenging problem in the study of the brain connectome. In this paper, we propose a novel EC estimation method based on an amortization transformer, named AT-EC. AT-EC first employs an amortization transformer to model the dynamics of fMRI time series and infers brain EC across different subjects. Then, an assisted learning mechanism is designed to assist the estimation of the brain EC network. Experimental results on both synthetic and real datasets show that AT-EC performs well compared to the state-of-the-art methods, which shows that the amortization causal discovery approach has great development potential in brain EC estimation. In the future, we will extend this work to more large-scale brain EC networks.

Author Contributions: J.L. designed the study; Z.Z. (Zuozhen Zhang) and Z.Z. (Ziqi Zhang) produced the results and wrote the manuscript. J.J. and J.L. were the guarantors of this study. J.L. provided fund grants. All authors listed have approved it for publication. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partly supported by National Natural Science Foundation of China Research Program (62106009, 62276010), in part by R&D Program of Beijing Municipal Education Commission (KM202210005030, KZ202210005009).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used for this study are available in the original studies cited.

Conflicts of Interest: The authors have declared no conflict of interest.

References

1. Ji, J.; Zou, A.; Liu, J.; Yang, C.; Zhang, X.; Song, Y. A Survey on Brain Effective Connectivity Network Learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *34*, 1879–1899. [CrossRef]
2. Friston, K.J. Functional and effective connectivity: A review. *Brain Connect.* **2011**, *1*, 13–36. [CrossRef]
3. Lv, H.; Liu, J.; Chen, Q.; Zhang, Z.; Wang, Z.; Gong, S.; Ji, J.; Wang, Z. Brain effective connectivity analysis facilitates the treatment outcome expectation of sound therapy in patients with tinnitus. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2023**, *31*, 1158–1166. [CrossRef]
4. Zhang, J.; Xia, J.; Liu, X.; Olichney, J. Machine Learning on Visibility Graph Features Discriminates the Cognitive Event-Related Potentials of Patients with Early Alzheimer's Disease from Healthy Aging. *Brain Sci.* **2023**, *13*, 770. [CrossRef]
5. Latella, D.; Maresca, G.; Formica, C.; Sorbera, C.; Bringandì, A.; Di Lorenzo, G.; Quartarone, A.; Marino, S. The Role of Telemedicine in the Treatment of Cognitive and Psychological Disorders in Parkinson's Disease: An Overview. *Brain Sci.* **2023**, *13*, 499. [CrossRef]
6. Cremone, I.M.; Carpita, B.; Nardi, B.; Casagrande, D.; Stagnari, R.; Amatori, G.; Dell'Osso, L. Measuring Social Camouflaging in Individuals with High Functioning Autism: A Literature Review. *Brain Sci.* **2023**, *13*, 469. [CrossRef]
7. Smith, S.M.; Miller, K.L.; Salimi-Khorshidi, G.; Webster, M.; Beckmann, C.F.; Nichols, T.E.; Ramsey, J.D.; Woolrich, M.W. Network modelling methods for FMRI. *Neuroimage* **2011**, *54*, 875–891. [CrossRef]
8. Shimizu, S.; Hoyer, P.O.; Hyvärinen, A.; Kerminen, A.; Jordan, M. A linear non-Gaussian acyclic model for causal discovery. *J. Mach. Learn. Res.* **2006**, *7*, 2003–2030.
9. Seth, A.K.; Barrett, A.B.; Barnett, L. Granger causality analysis in neuroscience and neuroimaging. *J. Neurosci.* **2015**, *35*, 3293–3297. [CrossRef]
10. Wang, Y.; David, O.; Hu, X.; Deshpande, G. Can Patel's τ accurately estimate directionality of connections in brain networks from fMRI? *Magn. Reson. Med.* **2017**, *78*, 2003–2010. [CrossRef]
11. Ramsey, J.D.; Hanson, S.J.; Hanson, C.; Halchenko, Y.O.; Poldrack, R.A.; Glymour, C. Six problems for causal inference from fMRI. *Neuroimage* **2010**, *49*, 1545–1558. [CrossRef]
12. Liu, J.; Ji, J.; Jia, X.; Zhang, A. Learning brain effective connectivity network structure using ant colony optimization combining with voxel activation information. *IEEE J. Biomed. Health Inform.* **2019**, *24*, 2028–2040. [CrossRef] [PubMed]
13. Ambrogioni, L.; Hinne, M.; Van Gerven, M.; Maris, E. GP CaKe: Effective brain connectivity with causal kernels. *Adv. Neural Inf. Process. Syst.* **2017**, *30, 951-60* .
14. Park, H.J.; Friston, K.J.; Pae, C.; Park, B.; Razi, A. Dynamic effective connectivity in resting state fMRI. *NeuroImage* **2018**, *180*, 594–608. [CrossRef] [PubMed]
15. Gilson, M.; Tauste Campo, A.; Chen, X.; Thiele, A.; Deco, G. Nonparametric test for connectivity detection in multivariate autoregressive networks and application to multiunit activity data. *Netw. Neurosci.* **2017**, *1*, 357–380. [CrossRef]
16. Kim, J.; Zhu, W.; Chang, L.; Bentler, P.M.; Ernst, T. Unified structural equation modeling approach for the analysis of multisubject, multivariate functional MRI data. *Hum. Brain Mapp.* **2007**, *28*, 85–93. [CrossRef]
17. Henry, T.R.; Feczko, E.; Cordova, M.; Earl, E.; Williams, S.; Nigg, J.T.; Fair, D.A.; Gates, K.M. Comparing directed functional connectivity between groups with confirmatory subgrouping GIMME. *NeuroImage* **2019**, *188*, 642–653. [CrossRef]
18. Xu, L.; Fan, T.; Wu, X.; Chen, K.; Guo, X.; Zhang, J.; Yao, L. A pooling-LiNGAM algorithm for effective connectivity analysis of fMRI data. *Front. Comput. Neurosci.* **2014**, *8*, 125. [CrossRef]
19. Friston, K.J.; Bastos, A.M.; Oswal, A.; van Wijk, B.; Richter, C.; Litvak, V. Granger causality revisited. *Neuroimage* **2014**, *101*, 796–808. [CrossRef]
20. Farokhzadi, M.; Hossein-Zadeh, G.A.; Soltanian-Zadeh, H. Nonlinear effective connectivity measure based on adaptive Neuro Fuzzy Inference System and Granger Causality. *NeuroImage* **2018**, *181*, 382–394. [CrossRef]
21. Talebi, N.; Nasrabadi, A.M.; Mohammad-Rezazadeh, I.; Coben, R. NCREANN: Nonlinear causal relationship estimation by artificial neural network; applied for autism connectivity study. *IEEE Trans. Med. Imaging* **2019**, *38*, 2883–2890. [CrossRef]
22. Khadem, A.; Hossein-Zadeh, G.A. Estimation of direct nonlinear effective connectivity using information theory and multilayer perceptron. *J. Neurosci. Methods* **2014**, *229*, 53–67. [CrossRef]
23. Ji, J.; Liu, J.; Han, L.; Wang, F. Estimating Effective Connectivity by Recurrent Generative Adversarial Networks. *IEEE Trans. Med. Imaging* **2021**, *40*, 3326–3336. [CrossRef] [PubMed]
24. Wang, Y.; Wang, Y.; Lui, Y.W. Generalized recurrent neural network accommodating dynamic causal modeling for functional MRI analysis. *NeuroImage* **2018**, *178*, 385–402. [CrossRef]
25. Liu, J.; Ji, J.; Xun, G.; Yao, L.; Huai, M.; Zhang, A. EC-GAN: Inferring brain effective connectivity via generative adversarial networks. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 4852–4859.
26. Liu, J.; Ji, J.; Xun, G.; Zhang, A. Inferring effective connectivity networks from fMRI time series with a temporal entropy-score. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *33*, 5993–6006. [CrossRef]
27. DSouza, A.M.; Abidin, A.Z.; Leistritz, L.; Wismüller, A. Exploring connectivity with large-scale Granger causality on resting-state functional MRI. *J. Neurosci. Methods* **2017**, *287*, 68–79. [CrossRef]
28. Li, H.; Yu, S.; Principe, J. Causal Recurrent Variational Autoencoder for Medical Time Series Generation. *arXiv* **2023**, arXiv:2301.06574.

29. Zou, A.; Ji, J. Learning brain effective connectivity networks via controllable variational autoencoder. In Proceedings of the 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Houston, TX, USA, 9–12 December 2021; pp. 284–287.
30. Chickering, D.M.; Meek, C.; Heckerman, D. Large-sample learning of bayesian networks is np-hard. *arXiv* **2012**, arXiv:1212.2468.
31. Amos, B. Tutorial on amortized optimization for learning to optimize over continuous domains. *arXiv* **2022**, arXiv:2202.00665.
32. Löwe, S.; Madras, D.; Zemel, R.; Welling, M. Amortized causal discovery: Learning to infer causal graphs from time-series data. In Proceedings of the Conference on Causal Learning and Reasoning. PMLR, Eureka, CA, USA, 11–13 April 2022; pp. 509–525.
33. Linke, A.C.; Mash, L.E.; Fong, C.H.; Kinnear, M.K.; Kohli, J.; Wilkinson, M.; Tung, R.; Keehn, R.J.; Carper, R.A.; Fishman, I.; et al. Dynamic time warping outperforms Pearson correlation in detecting atypical functional connectivity in autism spectrum disorders. *Neuroimage* **2020**, *223*, 117383. [CrossRef]
34. Sanchez-Romero, R.; Ramsey, J.D.; Zhang, K.; Glymour, M.R.; Huang, B.; Glymour, C. Estimating feedforward and feedback effective connections from fMRI time series: Assessments of statistical methods. *Netw. Neurosci.* **2019**, *3*, 274–306. [CrossRef]
35. Friston, K.J.; Harrison, L.; Penny, W. Dynamic causal modelling. *Neuroimage* **2003**, *19*, 1273–1302. [CrossRef]
36. Ramsey, J.D.; Sanchez-Romero, R.; Glymour, C. Non-Gaussian methods and high-pass filters in the estimation of effective connections. *Neuroimage* **2014**, *84*, 986–1006. [CrossRef] [PubMed]
37. Shah, P.; Bassett, D.S.; Wisse, L.E.; Detre, J.A.; Stein, J.M.; Yushkevich, P.A.; Shinohara, R.T.; Pluta, J.B.; Valenciano, E.; Daffner, M.; et al. Mapping the structural and functional network architecture of the medial temporal lobe using 7T MRI. *Hum. Brain Mapp.* **2018**, *39*, 851–865. [CrossRef]
38. Lavenex, P.; Amaral, D.G. Hippocampal-neocortical interaction: A hierarchy of associativity. *Hippocampus* **2000**, *10*, 420–430. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.