

Article



# Identifying Brain Network Structure for an fMRI Effective Connectivity Study Using the Least Absolute Shrinkage and Selection Operator (LASSO) Method

Xingfeng Li<sup>1,\*</sup> and Yuan Zhang<sup>2</sup>

- <sup>1</sup> Department of Surgery & Cancer, Hammersmith Campus, Imperial College London, Du Cane Road, London W12 0HS, UK
- <sup>2</sup> Key Laboratory of Language, Cognition and Computation of Ministry of Industry and Information Technology, School of Foreign Languages, Beijing Institute of Technology, 5 Zhongguancun South Street, Beijing 100081, China; yz@bit.edu.cn
- \* Correspondence: xingfengli@gmail.com

Abstract: Background: Studying causality relationships between different brain regions using the fMRI method has attracted great attention. To investigate causality relationships between different brain regions, we need to identify both the brain network structure and the influence magnitude. Most current methods concentrate on magnitude estimation, but not on identifying the connection or structure of the network. To address this problem, we proposed a nonlinear system identification method, in which a polynomial kernel was adopted to approximate the relation between the system inputs and outputs. However, this method has an overfitting problem for modelling the inputoutput relation if we apply the method to model the brain network directly. Methods: To overcome this limitation, this study applied the least absolute shrinkage and selection operator (LASSO) model selection method to identify both brain region networks and the connection strength (system coefficients). From these coefficients, the causality influence is derived from the identified structure. The method was verified based on the human visual cortex with phase-encoded designs. The functional data were pre-processed with motion correction. The visual cortex brain regions were defined based on a retinotopic mapping method. An eight-connection visual system network was adopted to validate the method. The proposed method was able to identify both the connected visual networks and associated coefficients from the LASSO model selection. Results: The result showed that this method can be applied to identify both network structures and associated causalities between different brain regions. Conclusions: System identification with LASSO model selection algorithm is a powerful approach for fMRI effective connectivity study.

**Keywords:** fMRI; model selection; system identification; brain imaging; visual cortex; effective connectivity; least absolute shrinkage and selection operator (LASSO)

## 1. Introduction

Since the discovery of blood oxygenation level-dependent (BOLD) functional MRI (fMRI) in 1990 [1], it has been extensively applied for studying brain functions. Both linear and nonlinear dynamic system modelling methods have been implemented for fMRI data analysis [2,3]. Dealing with the nonlinearity of the response of fMRI using dynamic system theory is a challenge due to the current limited availability of mathematical tools. In addition, because many biological systems and systems involving interactions between brain regions are available only as a black box, the relation between the system input (e.g., stimuli) and output (e.g., fMRI response) is not so derivable [4]. Wiener's theory of nonlinear systems has been developed to address the brain system with fMRI methods [5]. Other methods such as the Granger causality (GC) [6] and the multivariate vector autoregression (MVAR) model [7] have also been designed to study the causal relation between



Citation: Li, X.; Zhang, Y. Identifying Brain Network Structure for an fMRI Effective Connectivity Study Using the Least Absolute Shrinkage and Selection Operator (LASSO) Method. *Tomography* 2024, 10, 1564–1576. https://doi.org/10.3390/ tomography10100115

Academic Editors: Ho-Ling Anthony Liu and Nirbhay Narayan Yadav

Received: 18 June 2024 Revised: 6 September 2024 Accepted: 24 September 2024 Published: 30 September 2024



**Copyright:** © 2024 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/). different brain regions from neuroimage, and there is increasing interest in applying MVAR for fMRI effective connectivity studies [8–10]. MVAR can be regarded as a special case of a nonlinear system identification method as the basis of a function using polynomials to model the system inputs and outputs. But this method does not consider the nonlinear dynamics in the fMRI response.

The brain neural population response is nonlinear; therefore, nonlinear modelling methods for modelling the response are better as it is close to reality. Different methods and models have been proposed to address this problem. For example, a phase portrait and fuzzy recurrence plot method [11] has been explored to model the fMRI response. Additionally, the Hopf model has been invented to study the link between the brain structure and dynamics in the whole brain [12]. A dynamical flexible inference for nonlinear embeddings has been created to model nonlinear brain dynamics [13]. However, it is difficult to infer the causality relationship between different brain regions using these methods.

System identification is one of the methods used to build mathematical models of dynamic systems using measurements of the input and output signals of the system [14]. This method includes finding the structure of the network and the associated parameters, although most current methods concentrate on identifying system parameters only, not on the structure of the network for fMRI effective connectivity studies. The main purpose of this paper was to improve the system identification method by using a statistical model selection algorithm to determine the structure of the network for brain region causality studies. Particularly, the least absolute shrinkage and selection operator (LASSO) model selection method [15,16] was applied to decide the effective connectivity for fMRI studies. Using a nonlinear system identification method with non-linear basis functions, such as polynomial or radial basis functions, a regularization step is required to reduce the number of predictors and smooth the estimation results. For fMRI studies, the current model selection is based on either L1 (sparsity) [17–20] or L2 norm (ridge regression) [21,22] and these methods have been mainly applied to functional connectivity studies [23-25]. Although the LASSO method has been adopted for causality studies using the electroencephalogram (EEG) method [26], the applicability of this method for nonlinear fMRI responses has not been studied.

Moreover, for fMRI effective connectivity studies with the MVAR method [27], this method does not include nonlinear terms and/or system inputs [28]. To overcome this limitation, the purpose of this study was to suggest a LASSO model selection algorithm [15,16] for effective connectivity based on the nonlinear system identification method. Specifically, a nonlinear system identification method with a model selection algorithm was applied to investigate the cause–effect relationship from different brain regions.

#### 2. Materials and Methods

#### 2.1. Subjects, Experimental Design, MRI Collection, and fMRI Image Pre-Processing

This is a retrospective study of a previous study [29], in which brain responses to external stimuli between healthy controls and amblyopic patients were compared. Subjects, experimental design, data collection, and fMRI image pre-processing can be found in the previous study [29]. Briefly, visual eccentricity and polar angle stimuli were adopted for the experimental design, and subjects viewed the stimuli monocularly. A 1.5-T (Magnetom; Siemens, Madison, WI, USA) scanner was adopted to collect structural and functional image data. The repetition time (TR) was 3 s for the fMRI data collection. The fMRI image resolution was 4 mm  $\times$  4 mm, and the data were obtained from the visual cortex. The fMRI was pre-processed with motion correction and the visual cortex boundaries were defined based on the responses. A Z-score method was employed to normalise the time series temporally. For each visual cortex region, a single voxel response was selected randomly to represent the fMRI response from the region. There were 128 image frames for the fMRI time series. The first 8 images were excluded from the analysis due to the instability of the magnetic fields at the beginning of the fMRI scan; therefore, only 120 fMRI time series were used. fMRI scans from one subject viewing the eccentricity and polar

angle stimuli with the right eye were included, and only first-level (within each run) fMRI analysis was involved in this study.

#### 2.2. Theory

Briefly, the discrete nonlinear dynamic brain can be described using a differential equation as follows [30]:

$$\begin{cases} x(t+1) = f(x(t), u(t), \theta) \\ y(t) = g(x(t), \theta) \end{cases}$$
(1)

where x(t) is the state variable,  $\theta$  is the parameter, y(t) is the system output (fMRI response in this study), f and g are nonlinear functions, and u(t) is the system input (stimuli for task fMRI). For resting-state fMRI, u(t) = 0. For the effective connectivity study, we need to identify the system connection structure and parameters in Equation (1).

Taking a 3-connection network (Figure 1A) as an example, it has a 3-connection network structure with responses from 3 brain regions. In the network, each region is connected to other regions mutually with an additional input. The inputs are used to represent the fMRI response due to the stimuli from the experimental design. Figure 1B shows another example with a larger network, which has an eight-connection network. For this network,  $Y_1(t)$ ,  $Y_2(t)$ ,  $Y_3(t)$ ,  $Y_4(t)$ ,  $Y_5(t)$ ,  $Y_6(t)$ ,  $Y_7(t)$ , and  $Y_8(t)$  denote the fMRI time series from visual cortex regions of V1, V2, V3, Vp, V3a, V4v, V3b, and V5, respectively. For illustration purposes, we will only use these two network structures as shown in Figure 1, but this method can be applied to other network structures in the same way for brain effective connectivity studies.

To model the input–output relation of this nonlinear brain system as shown in Figure 1A, a polynomial kernel is adopted to approximate the nonlinear brain system. Using the first order auto-regression (AR) model, without taking into account nonlinear terms and neglecting the system inputs, for this 3-connection network (Figure 1A), we can write the first level fMRI causality analysis model as:

$$Y_1(t) = a_1 Y_1(t-1) + a_2 Y_2(t-1) + a_3 Y_3(t-1) + \varepsilon_1(t)$$
(2)

$$Y_2(t) = b_1 Y_1(t-1) + b_2 Y_2(t-1) + b_3 Y_3(t-1) + \varepsilon_2(t)$$
(3)

$$Y_3(t) = c_1 Y_1(t-1) + c_2 Y_2(t-1) + c_3 Y_3(t-1) + \varepsilon_3(t)$$
(4)

In Equations (2)–(4),  $Y_1(t)$ ,  $Y_2(t)$ , and  $Y_3(t)$  are the fMRI response from the brain regions at time point *t*. *a*, *b*, and *c* are the model coefficients, and  $\varepsilon$  is the model residual. For instance, from Equation (2), we can test the coefficient  $a_2$  to study the influence from  $Y_2(t-1)$  on  $Y_1(t)$ . This can be achieved by the F/T test. We adopted the F test to test whether  $a_2$  is significantly larger than 0 or not. To achieve this, from Equation (2), we have a null hypothesis:

$$H_0: a_2 = 0$$
 (5)

The reduced model of Equation (2) can be written as:

$$Y_1(t) = a_1 Y_1(t-1) + a_3 Y_3(t-1) + \varepsilon_4(t)$$
(6)

The Granger causality from  $Y_2(t)$  to  $Y_1(t)$  is quantified by the F test as:

$$F_{1-\alpha}(q, N-k) = \frac{(RRSS - URSS)/q}{(URSS)/(N-k)}$$
(7)

where the RRSS denotes the Restricted Residual Sum of Squares (RRSS) as the residual sum of squares obtained from estimating the restricted model (i.e., from Equation (6),  $RRSS = \varepsilon_4(t)\varepsilon_4(t)^T$ ). URSS represents the unrestricted residual sum of squares (URSS) as the residual sum of squares obtained from estimating the unrestricted model (i.e., from Equation (2),  $RRSS = \varepsilon_1(t)\varepsilon_1(t)^T$ ).  $\alpha$  is the level of significance. *q* is the degrees of freedom

of the numerator, as we only test one parameter, i.e., q = 1 in this example. N - k is the denominator degree of freedom. In our case, the T test square is the F test, and the corresponding T test is:

$$T_{N-k}^2 = F(q, N-k)$$
 (8)

with the denominator degree of freedom N - k. N is the sample size (length of fMRI time series), and k is the number of predictors in the reduced model, i.e., k = 2 in this example.



**Figure 1.** (**A**) An example of a three-connection network structure. (**B**) An example of an eight-connection network structure. Line arrow indicates the direction of information flow (causality).

## 2.3. Visual System

In this study, an eight-connection human visual system was used, and the connection network is displayed in Figure 1B. In the system, we adopted the polynomial basis function to approximate the nonlinear dynamic system. If we neglect these nonlinear interacting terms and high-order autoregression terms (using a 3-order autoregression term or below), the system output  $Y_1(t)$  (fMRI response from V1 region) can be modelled as:

$$Y_{1}(t) = a_{0} + a_{1} \quad Y_{1}(t-1) + a_{2}Y_{2}(t-1) + \dots + a_{8}Y_{8}(t-1) + b_{1}Y_{1}(t-2) + b_{2}Y_{2}(t-2) + \dots + b_{8}Y_{8}(t-2) + c_{1}Y_{1}(t-3) + c_{2}Y_{2}(t-3) + \dots + c_{8}Y_{8}(t-3) + d_{1}u_{1}(t-1) + d_{2}u_{2}(t-1) + \dots + d_{8}u_{8}(t-1) + e_{1}u_{1}(t-2) + e_{2}u_{2}(t-2) + \dots + e_{8}u_{8}(t-2) + f_{1}u_{1}(t-3) + f_{2}u_{2}(t-3) + \dots + f_{8}u_{8}(t-3) + g_{1}Y_{1}^{2}(t-1) + g_{2}Y_{2}^{2}(t-1) + \dots + g_{8}Y_{8}^{2}(t-1) + h_{1}Y_{1}^{2}(t-2)$$
(9)  
$$+ h_{2}Y_{2}^{2}(t-2) + \dots + h_{8}Y_{8}^{2}(t-2) + i_{1}Y_{1}^{2}(t-3) + i_{2}Y_{2}^{2}(t-3) + \dots + i_{8}Y_{8}^{2}(t-3) + j_{1}u_{1}^{2}(t-1) + j_{2}u_{2}^{2}(t-1) + \dots + j_{8}u_{8}^{2}(t-1) + k_{1}u_{1}^{2}(t-2) + k_{2}u_{2}^{2}(t-2) + \dots + k_{8}u_{8}^{2}(t-2) + l_{1}u_{1}^{2}(t-3) + l_{2}u_{2}^{2}(t-3) + \dots + l_{8}u_{8}^{2}(t-3) + \varepsilon(t)$$

There are 97 parameters that need to be estimated in Equation (9), i.e., 24 items from the AR of the fMRI response, 24 items from the system inputs, 24 nonlinear (square) responses, 24 nonlinear system input items, and  $a_0$  is the intercept. The reason for excluding nonlinear interaction terms such as  $Y_2(t-1)Y_3(t-2)$  is because the effect is the combination of two different regions, which is not easy to interpret from the brain science viewpoint. Equation (9) can be written in matrix form as:

$$Y = \beta_0 + x_i^T \beta + \varepsilon \tag{10}$$

where *Y* is the  $(n - AR) \times 1$  vector to represent the fMRI response  $Y_1(t)$ , *n* is the number of fMRI time frames, and *AR* is the maximum order of AR.  $x_i$  is the  $(n - AR) \times 1$  vector value at observation *i*, i.e., the data vector.  $x_1 = Y_1(t - 1), x_2 = Y_2(t - 1), \dots, x_{96} = u_8^2(t - 3)$ .  $\beta$  is the coefficient vector of each term in Equation (9),  $\beta_0 = a_0$  is the intercept, and  $\varepsilon$  is the residuals.

For the causality within the seed region  $Y_1$  (for the other regions, the same method can be applied), the least square solution of the equation is to minimise the mean square error objective function:

$$\min_{\beta_{0,\beta}} \left( \frac{1}{2N} \sum_{i=1}^{N} \left( y_i - \beta_0 - x_i^T \beta \right)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right)$$
(11)

where N = n - AR and p is the number of coefficients, i.e., p = 97 in this study (including the intercept).  $\lambda$  is a positive regularisation parameter corresponding to one value of Lambda.

MATLAB (version 2022b) was employed for modelling the visual system using Equation (10) while minimising Equation (11). Function lassoglm.m from MATLAB was adopted to implement the LASSO model selection.

#### 3. Results

The brain response to the eccentricity stimulus was adopted to validate the method in this example. Based on the boundaries from the region of the visual cortex, one fMRI time series from each visual cortex region was selected randomly. For the eight-connection brain visual system (Figure 1B), the fMRI response or system output time series is displayed as the dotted asterisk curves in Figure 2. The red curve is the input for the time series, i.e., it assumes that the system input contributes to generate an fMRI response. The red curve is the fundamental frequency of the fMRI response, which was obtained from the fast Fourier transform (FFT) to model the system input.



**Figure 2.** The fMRI time series from the eight-connection visual system in response to an eccentricity stimulus for the effective connectivity study. (**A**) (V1), (**B**) (V2), (**C**) (V3), (**D**) (Vp), (**E**) (V3a), (**F**) (V4v), (**G**) (V3b), and (**H**) (V5/MT) were included in the brain network. The dotted curve shows the fMRI response, while the red curve indicates the input from the experimental design. The *x*-axis is the normalised fMRI signal, while the *y*-axis is time (in terms of TR = 3 s).



**Figure 3.** Predictors matrix, e.g.,  $x_i^T$  in Equation (10).

From Figure 2, we can generate a predictor matrix as shown in Figure 3. In the Figure, the first 24 columns of the matrix (denoted as AR on the top) are derived from an auto-regression term from different regions, e.g.,  $Y_1(t-1), Y_2(t-1), \dots, Y_8(t-3)$ . The second 24 columns ("Input" on the top) are the inputs for different regions, e.g.,  $u_1(t-1), u_2(t-1), \dots, u_8(t-3)$ . The third 24 columns ("ArNon" on the top) are nonlinear terms of the responses, e.g.,  $u_1^2(t-1), u_2^2(t-1), \dots, u_8^2(t-3)$ . The last 24 columns (denoted by "iNonlin") are the nonlinear terms of inputs from eight-connected regions, e.g.,  $Y_1^2(t-1), Y_2^2(t-1), \dots, Y_8^2(t-3)$ .

Resolving Equation (10) by minimizing Equation (11), we obtained the L1 regularization parameter Lambda for model selection. The selection process with the 10-fold cross-validation method is illustrated in Figure 4.

In Figure 4, there are two vertical dashed lines. The vertical dashed line on the right (green) corresponds to the Lambda for the minimum mean cross-validated error. The vertical dotted dashed line on the left (blue) locates the point with a minimum cross-validation error plus one standard deviation (Lambda = 0.1643). Lambda with one standard error (1SE) is the value that gives the most regularised model such that the cross-validated error is within one standard error of the minimum. In this study, the Lambda value corresponding to the minimum deviance was adopted for LASSO model selection.



**Figure 4.** LASSO selection results. The blue circle and (**left**) dotted line locate the point with the minimum cross-validation error plus one standard deviation (Lambda 1SE = 0.1643). The green circle and (**right**) dotted line locate the Lambda with minimum cross-validation error (LambdaMinDeviance = 0.0648).

For the model as described in Equation (9), and the responses shown in Figure 2, applying the LASSO method, the selected model  $y_1(t)$  for the response is:

$$y_1(t) = -0.0735 - 0.0121y_3(t-3) + 6.4472u_3(t-2) + 3.9456u_3(t-3) + 0.019y_4^2(t-1) + 0.3069u_3^2(t-2) + \varepsilon$$
(12)

Five items have been selected in Equation (12), and the associated coefficients are displayed. Applying Equations (7) and (8), we obtained the *t*-test value (the degree of freedom is N - k = 117 - 6 = 111) as 1.6383 (intercept), 0.9175 ( $y_3(t - 3)$ ), 12.7305 ( $u_3(t - 2)$ ), 8.2911 ( $u_3(t - 3)$ ), 1.4245  $y_4^2(t - 1)$ , and 2.1506  $u_3^2(t - 2)$ . From the t table, the influence from  $u_3(t - 2)$ ,  $u_3(t - 3)$ , and  $u_3^2(t - 2)$  to  $y_1(t)$  was significant (p < 0.05).

Based on Equation (12), we can plot the final prediction result as shown in Figure 5. The dotted blue curve is the response from V1, while the red curve is the fitted results with the LASSO model response from Equation (12). Another validation study example of an effective connectivity study in the visual cortex with polar angle stimuli is presented in the Appendix A.



**Figure 5.** A visual cortex (V1 region) response to the stimuli and final prediction results from the identified model structure and its parameters. The red curve shows the prediction result from the V1 region.

## 4. Discussion

Nonlinear system identification methods [5] and MAR methods [7] have been extensively suggested for modelling dynamic brain systems. Technically, the main difference between these two methods is that for the MAR method, system inputs and nonlinear terms are not considered for modelling the dynamic system. For the nonlinear system identification method, both nonlinear terms and system inputs can be included in the model. For the nonlinear system identification method, other kernels such as radial basis functions can also be applied to model the system [31]. While the MAR method is a special case for nonlinear system identification methods, if the identification kernel is polynomial and without system input, and there are nonlinear terms in the model for selection, the MAR method is the same as the identification method [2].

Stimuli from experimental design can be regarded as a system input as it is something put into the brain system or expended in its operation to achieve output. Statistical model selection was employed to decide whether the response was caused by system inputs (stimuli) or the connected regions. In a block design with periodic stimuli, including the phase-encoding block design as shown in this study, the input is a travelling square wave, which has delays, and the shape has been distorted. Therefore, the fundamental frequency of the response was adopted for a representative system input [3]. For resting-state fMRI studies, it can be regarded as a special case, i.e., a system without input or the input is 0.

In addition to the LASSO method, the last-angle regression (LARS) method for model selection has been applied for selecting models to study effective connectivity [20]. The LARS algorithm is a greedy method that does not yield a provably consistent estimator. The LASSO method has been applied to functional and effective connectivity studies [10,27], while these methods are based on MAR models. Stepwise regression, backwards selection such as recursive feature elimination (RFE), and elastic net regularisation methods [16] can be applied to identify the structure of the brain connection networks.

Individual regional effective connectivity variability is bigger when applying the model selection method for causality inference (comparing the selected model in the Section 3 and Appendix A). This is because the model selection is sensitive to the fMRI response. The fMRI image spatial resolution in this study is  $4 \text{ mm} \times 4 \text{ mm} \times 4 \text{ mm}$ . A single voxel in fMRI measures the response of a neural population, which includes the response from millions of single neurons. The response at each voxel is different, so the causality inference based on each voxel is different accordingly. A small change in an fMRI response will lead to a bigger change in effective connectivity. Furthermore, when applying the LASSO method for model selection, a ten-fold cross-validation method was used, which also has randomness.

The limitation of this study is that the method has not been validated on an eventrelated fMRI response. In this study, phase-encoded experiments were adopted. This experimental design can be regarded as a travelling wave, and it is similar to a block design. The advantage of this type of experimental design is that the Fourier transform can be applied to approximate the system input. This makes the analysis simpler but more powerful. For the other types of experimental design, system inputs are much more difficult to model, and the system input needs to be adjusted to fit the study.

Compared with other studies [10,27], the advantage of this study is that this method takes nonlinear and system inputs into account for modelling the effective connectivities. In addition, this study applied LASSO methods for identifying the system structure and associated parameters automatically. This study combined the results from model selection with causality inference and was tested in a real human visual system.

One of the future directions for applying the results of this study is to overcome the limitations of the randomness for the causality study. Statistical methods need to be developed to address this problem. At the subject level, there are individual variabilities for measuring causality. Methods need to be developed to deal with individual variabilities and combine different causality results with statistical measurements for group analysis. To reduce the model randomness, robust statistical methods need to be applied to detect the outliers from the estimation.

Furthermore, a deep learning method has been proposed to model the nonlinear system [32]. In the framework for system identification, an autoencoder structure of the neuron network was introduced, in which an encoder was employed to estimate the deep features (latent variables) of the model. Then a decoder was applied to restore the system output estimation. Using deep features, the causality relationship between different deep features could be inferred.

Finally, this study used the fMRI from the visual cortex. Other cortexes and other stimuli have not been verified. For example, it is possible to apply this method to study the motor cortex system causality relationship. Additionally, it may be of interest to apply this method to study the human brain under disease conditions. Lastly, similar to a dynamic functional connectivity study that adopted sliding windows [33], this method can be extended to use the sliding windows method to study dynamic casualty.

**Author Contributions:** Conceptualization, X.L.; methodology, X.L.; software, X.L.; validation, X.L.; formal analysis, X.L.; writing—original draft preparation, X.L. and Y.Z.; writing—review and editing, X.L. and Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

**Institutional Review Board Statement:** Ethical review and approval were waived for this study because of the retrospective data analysis and method development nature of the study. All original studies were performed with the informed consent of the subjects and approved by the Montreal Neurologic Institute Research Ethics Committee and adhered to the tenets of the Declaration of Helsinki.

**Informed Consent Statement:** Informed consent was waived because of the retrospective nature of the study. The analysis also used anonymous data.

**Data Availability Statement:** The original contributions presented in the study are included in the article (https://doi.org/10.1167/iovs.06-1021) Further inquiries can be directed to the corresponding authors.

Conflicts of Interest: The authors declare no conflicts of interest.

#### Appendix A

In this appendix, effective connectivity at the same voxel in response to polar angle stimuli viewing with the right eye from the same subject was employed in the analysis. The fMRI response in V1 (A), V2 (B), V3 (C), Vp (D), V3a (E)), V4v (F), V3b (G), and V5/MT (H) is displayed in Figure A1. Compared with Figure 2, the fMRI response is very different, particularly in the V5/MT region.

Similar to Figure 3, Figure A2 shows the predictor matrix obtained from fMRI responses (Figure A1). In the figure, the first 24 columns of the matrix (denoted as AR on the top) are derived from auto-regression terms from different visual cortex regions. The second 24 columns ("Input" on the top) are the inputs for different regions. The third 24 columns ("ArNon" on the top) are nonlinear terms of the responses, and the last 24 columns (denoted by "iNonlin") are the nonlinear terms of inputs from the eight-connected visual cortical regions.

Resolving Equation (10) by minimising Equation (11), we obtained Lambda from the LASSO method. The selection process is illustrated as shown in Figure A3.

In Figure A3, there are two vertical dashed lines. The vertical dashed line on the left (blue) corresponds to the lambda with minimum cross-validation error plus one standard deviation (Lambda 1SE = 0.1045). The vertical dashed line on the right (green) locates the Lambda with minimum cross-validation error (Lambda = 0.0497). The Lambda with minimum deviance value was adopted for model selection in this study.



**Figure A1.** Time series from the eight-connection visual system in response to polar angle stimulus for effective connectivity study. The fMRI response was obtained from (**A**) (V1), (**B**) (V2), (**C**) (V3), (**D**) (Vp), (**E**) (V3a), (**F**) (V4v), (**G**) (V3b), and (**H**) (V5/MT). The dotted curve shows the fMRI response, while the red curve indicates the input.



**Figure A2.** Predictor matrix, i.e.,  $x_i^T$  in Equation (10).

For the model as described in Equation (9), and the responses as shown in Figure A1, applying the LASSO method, the selected model  $y_1(t)$  for response is:

 $\begin{array}{l} y_1(t)=0.0376+0.1819y_1(t-1)+0.0904y_2(t-1)+0.0618y_3(t-1)+\\ 0.2021y_3(t-2)+0.0219y_5(t-3)+0.0581y_8(t-2)+0.1512u_3(t-1)+\\ 0.4077u_5(t-2)+0.0560y_3^2(t-1)+0.0343y_3^2(t-2)+0.0237y_3^2(t-3)-\\ 0.0069y_4^2(t-2)+0.0816y_7^2(t-2)+0.0962u_3^2(t-1)-1.6881u_4^2(t-3)+\varepsilon. \end{array}$ 

Sixteen items have been selected in the above equation, and the associated coefficients were also included. Applying Equations (7) and (8), we obtained the T test value (degree of freedom is N - k = 117 - 6 = 111) for each coefficient as 0.7835 (intercept), 4.7356, 2.7119, 2.0795, 5.1189, 1.0681,1.9739, 2.0008, 3.4942, 2.5979, 1.8091, 1.4014, 0.6195, 3.2746, 0.7606, and 6.6170, respectively. From the t statistic table, the influence from  $y_2(t - 1)$  and  $y_3(t - 1)$  to y1 is significant (p < 0.05). Based on the selected model, the prediction of the response (red curve) is displayed in Figure A4.



**Figure A3.** LASSO selection results and Lambda values. Lambda 1SE = 0.1045 is the regularisation corresponding to the minimum cross-validation error plus one standard deviation. The vertical dotted line on the right indicates the minimum cross-validation error value (Lambda MinDeviance = 0.0497).



**Figure A4.** The final prediction result based on the identified structure and its parameters. The dotted curve shows the fMRI response, while the red curve indicates the predicted response for the LASSO selected model. The *x*-axis is the magnitude of the normalized fMRI response, while the *y*-axis is time in terms of TR.

## References

- 1. Ogawa, S.; Lee, T.M.; Kay, A.R.; Tank, D.W. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc. Natl. Acad Sci. USA* **1990**, *87*, 9868–9872. [CrossRef]
- Rugh, W.J. Nonlinear System Theory—The Volterra-Wiener Approach, Web version prepared in 2002; Johns Hopkins University Press: Baltimore, MD, USA, 1981; Chapter 1; pp. 1–53. ISBN O-8018-2549-0.
- 3. Boynton, G.M.; Engel, S.A.; Glover, G.H.; Heeger, D.J. Linear Systems Analysis of Functional Magnetic Resonance Imaging in Human V1. *J. Neurosci.* 1996, *16*, 4207–4221. [CrossRef]
- 4. Schetzen, M. Nonlinear system modeling based on the Wiener theory. Proc. IEEE 1981, 69, 1557–1573. [CrossRef]
- Li, X.; Marrelec, G.; Hess, R.F.; Benali, H. A nonlinear identification method to study effective connectivity in functional MRI. Med. Image Anal. 2010, 14, 30–38. [CrossRef] [PubMed]
- Goebel, R.; Roebroeck, A.F.; Kim, D.-S.; Formisano, E. Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. *Magn. Reason. Imaging* 2003, 21, 1251–1261. [CrossRef] [PubMed]
- 7. Harrison, L.; Penny, W.; Friston, K.; Harrison, L.; Penny, W.; Friston, K.; Harrison, L.; Penny, W.; Friston, K. Multivariate autoregressive modeling of fMRI time series. *NeuroImage* **2003**, *19*, 1477–1491. [CrossRef] [PubMed]
- 8. Cao, J.; Yang, L.; Sarrigiannis, P.G.; Blackburn, D.; Zhao, Y. Dementia classification using a graph neural network on imaging of effective brain connectivity. *Comput. Biol. Med.* **2024**, *168*, 107701. [CrossRef]
- 9. Bressler, S.L.; Kumar, A.; Singer, I. Brain Synchronization and Multivariate Autoregressive (MVAR) Modeling in Cognitive Neurodynamics. *Front. Syst. Neurosci.* 2021, 15, 638269. [CrossRef]
- Nagle, A.; Gerrelts, J.P.; Krause, B.M.; Boes, A.D.; Bruss, J.E.; Nourski, K.V.; Banks, M.I.; Van Veen, B. High-dimensional multivariate autoregressive model estimation of human electrophysiological data using fMRI priors. *NeuroImage* 2023, 277, 120211. [CrossRef] [PubMed]
- 11. Li, Q.; Calhoun, V.D.; Pham, T.D.; Iraji, A. Exploring Nonlinear Dynamics in Brain Functionality through Phase Portraits and Fuzzy Recurrence Plots. *bioRxiv* 2024. [CrossRef]
- 12. Ponce-Alvarez, A.; Deco, G. The Hopf whole-brain model and its linear approximation. *Sci. Rep.* **2024**, *14*, 2615. [CrossRef] [PubMed]
- 13. Abbaspourazad, H.; Erturk, E.; Pesaran, B.; Shanechi, M.M. Dynamical flexible inference of nonlinear latent factors and structures in neural population activity. *Nat. Biomed. Eng.* **2024**, *8*, 85–108. [CrossRef] [PubMed]
- 14. Åström, K.J.; Eykhoff, P. System identification—A survey. Automatica 1971, 7, 123–162. [CrossRef]
- 15. Hastie, T.; Tibshirani, R.; Friedman, J.H.; Friedman, J.H. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction;* Springer: New York, NY, USA, 2009; Volume 2.
- 16. Zou, H.; Hastie, T. Regularization and variable selection via the elastic net. J. R. Stat. Soc. Ser. B (Stat. Methodol.) 2005, 67, 301–320. [CrossRef]
- Valdés-Sosa, P.A.; Sánchez-Bornot, J.M.; Lage-Castellanos, A.; Vega-Hernández, M.; Bosch-Bayard, J.; Melie-García, L.; Canales-Rodríguez, E. Estimating brain functional connectivity with sparse multivariate autoregression. *Philos. Trans. R. Soc. B Biol. Sci.* 2005, 360, 969–981. [CrossRef] [PubMed]
- Liu, A.; Chen, X.; McKeown, M.J.; Wang, Z.J. A Sticky Weighted Regression Model for Time-Varying Resting-State Brain Connectivity Estimation. *IEEE Trans. Biomed. Eng.* 2015, 62, 501–510. [CrossRef] [PubMed]
- 19. Shojaie, A.; Michailidis, G. Discovering graphical Granger causality using the truncating lasso penalty. *Bioinformatics* **2010**, *26*, i517–i523. [CrossRef] [PubMed]
- Li, X.; Coyle, D.; Maguire, L.; McGinnity, T.M.; Benali, H. A Model Selection Method for Nonlinear System Identification Based fMRI Effective Connectivity Analysis. *IEEE Trans. Med. Imaging* 2011, 30, 1365–1380. [CrossRef] [PubMed]
- Pongrattanakul, A.; Lertkultanon, P.; Songsiri, J. Sparse system identification for discovering brain connectivity from fMRI time series. In *The SICE Annual Conference*; IEEE: Piscataway, NJ, USA, 2013.
- Haufe, S.; Müller, K.R.; Nolte, G.; Krämer, N. Sparse causal discovery in multivariate time series. In Proceedings of the 2008th International Conference on Causality: Objectives and Assessment, Vancouver, BC, Canada, 12 December 2008; JMLR.org: Whistler, BC, Canada, 2008; Volume 6, pp. 97–106.
- Ryali, S.; Chen, T.; Supekar, K.; Menon, V. Estimation of functional connectivity in fMRI data using stability selection-based sparse partial correlation with elastic net penalty. *NeuroImage* 2012, *59*, 3852–3861. [CrossRef]
- Shojaie, A.; Fox, E.B. Granger Causality: A Review and Recent Advances. Annu. Rev. Stat. Its Appl. 2022, 9, 289–319. [CrossRef] [PubMed]
- Teipel, S.J.; Grothe, M.J.; Metzger, C.D.; Grimmer, T.; Sorg, C.; Ewers, M.; Franzmeier, N.; Meisenzahl, E.; Klöppel, S.; Borchardt, V.; et al. Robust Detection of Impaired Resting State Functional Connectivity Networks in Alzheimer's Disease Using Elastic Net Regularized Regression. *Front. Aging Neurosci.* 2017, *8*, 318. [CrossRef] [PubMed]
- Siggiridou, E.; Kugiumtzis, D. Granger Causality in Multivariate Time Series Using a Time-Ordered Restricted Vector Autoregressive Model. *IEEE Trans. Signal Process.* 2016, 64, 1759–1773. [CrossRef]
- 27. Tang, W.; Bressler, S.L.; Sylvester, C.M.; Shulman, G.L.; Corbetta, M. Measuring Granger Causality between Cortical Regions from Voxelwise fMRI BOLD Signals with LASSO. *PLoS Comput. Biol.* **2012**, *8*, e1002513. [CrossRef] [PubMed]
- 28. Das, P.; Babadi, B. Non-Asymptotic Guarantees for Reliable Identification of Granger Causality via the LASSO. *IEEE Trans. Inf. Theory* **2023**, *69*, 7439–7460. [CrossRef] [PubMed]

- 29. Li, X.; Dumoulin, S.O.; Mansouri, B.; Hess, R.F. Cortical Deficits in Human Amblyopia: Their Regional Distribution and Their Relationship to the Contrast Detection Deficit. *Investig. Ophthalmol. Vis. Sci.* **2007**, *48*, 1575–1591. [CrossRef] [PubMed]
- 30. Nozari, E.; Bertolero, M.A.; Stiso, J.; Caciagli, L.; Cornblath, E.J.; He, X.; Mahadevan, A.S.; Pappas, G.J.; Bassett, D.S. Macroscopic resting-state brain dynamics are best described by linear models. *Nat. Biomed. Eng.* **2023**, *8*, 68–84. [CrossRef]
- Pattanaik, R.K.; Mohanty, M.N. Nonlinear System Identification Using Robust Fusion Kernel-Based Radial basis function Neural Network. In Proceedings of the 2022 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 9–11 March 2022.
- Gedon, D.; Wahlström, N.; Schön, T.B.; Ljung, L. Deep state space models for nonlinear system identification. *IFAC-PapersOnLine* 2021, 54, 481–486. [CrossRef]
- Hutchison, R.M.; Womelsdorf, T.; Allen, E.A.; Bandettini, P.A.; Calhoun, V.D.; Corbetta, M.; Della Penna, S.; Duyn, J.H.; Glover, G.H.; Gonzalez-Castillo, J.; et al. Dynamic functional connectivity: Promise, issues, and interpretations. *NeuroImage* 2013, *80*, 360–378. [CrossRef]

**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.