Nonlinear System Identification of Neural Systems from Neurophysiological Signals

Fei He[1,*] and Yuan Yang[2,3,4,†]

[1] *Centre for Data Science, Coventry University, Coventry CV1 2JH, UK.*
[2] *Stephenson School of Biomedical Engineering, The University of Oklahoma, Tulsa, Oklahoma-74135, USA.*
[3] *Department of Physical Therapy and Human Movement Sciences, Feinberg School of Medicine, Northwestern University, Chicago, Illinois-60611, USA.*
[4] *Laureate Institute for Brain Research, Tulsa, Oklahoma-74136, USA.*

The human nervous system is one of the most complicated systems in nature. The complex nonlinear behaviours have been shown from the single neuron level to the system level. For decades, linear connectivity analysis methods, such as correlation, coherence and Granger causality, have been extensively used to assess the neural connectivities and input-output interconnections in neural systems. Recent studies indicate that these linear methods can only capture a small amount of neural activities and functional relationships, and therefore cannot describe neural behaviours in a precise or complete way. In this review, we highlight recent advances on nonlinear system identification of neural systems, corresponding time and frequency domain analysis, and novel neural connectivity measures based on nonlinear system identification techniques. We argue that nonlinear modelling and analysis are necessary to study neuronal processing and signal transfer in neural systems quantitatively. These approaches can hopefully provide new insights to advance our understanding of neurophysiological mechanisms underlying neural functions. They also have the potential to produce sensitive biomarkers to facilitate the development of precision diagnostic tools for evaluating neurological disorders and the effects of targeted intervention.

I. INTRODUCTION

The human nervous system is a complicated network comprised of more than 10 billion neurons, with trillions of synapses connecting them. The behaviour of a single neuron is highly nonlinear, showing a step-like 'none-or-all' firing response [1], while the behaviour of neurons in a population could be relatively similar. Therefore, the nonlinear response of each individual neuron may be smoothed out by the distribution of membrane thresholds across the population, known as the pool effect [2]. This typically occurs in a mono-synaptic neural system such as the cortico-spinal tract where the supraspinal motor command is linearly transferred to the motor output due to the pool effect of motor units [3]. However, multi-synaptic neural systems, such as the somatosensory system, have been reported highly nonlinear, showing harmonic responses to periodic stimuli [4–7]. The cross-frequency coupling in the corticothalamic interactions has also been reported [8]. The nonlinear behaviour in neural systems is thought to be associated with various neural functions, including neuronal encoding, neural processing of synaptic inputs, communication between different neuronal populations and functional integration [9–12].

Various functional and effective connectivity measures have been developed [13] to characterise such (linear and nonlinear) functional integration, from neurophysiological signals – such as electroencephalogram (EEG) and magnetoencephalogram (MEG) from the brain, Electromyogram (EMG) from muscles. While functional connectivity (e.g. correlation, coherence, mutual information) only quantifies the statistical dependencies among signals from different areas, effective connectivity attempts to quantify the causal influences of one neural system exerts over another, either at a synaptic or population level [14]. Mostly, effective connectivity depends on models of interactions or coupling (although there exist model-free measures like transfer entropy [15]) and is often time-dependent (dynamic). Therefore, effective connectivity has a strong link with dynamic modelling (from input-output neurophysiological signals), also known as system identification [16, 17], and corresponding model-based causality analysis.

Many nonlinear system identification methods have been previously proposed and developed to investigate nonlinear behaviours in neural systems. These approaches provide us with useful tools to explore the nonlinear nature of neural systems [9]. In this review article, we highlight recent advances in nonlinear system identification of neural systems, as well as novel neural connectivity methods based on nonlinear system identification techniques. A diagram that summarises the linear and nonlinear functional and effective connectivity measures and their links with system identification is provided in Figure 1.

II. NONLINEARITY IN THE NEURONAL LEVEL AND NEURAL SYSTEMS

In a single neuron level, the action potential spike is the principal basis of information encoding, which allows signal transmission across different neuronal populations [12]. The spike timing is thought to be associated with the coding scheme in neural systems [18]. The

* fei.he@coventry.ac.uk
† yuan.yang@northwestern.edu

nonlinear nature of the neuronal process of synaptic input influences the temporal firing behaviour of individual neurons. Different types of neurons have their own repertoire of ion channels that are responsible for their characteristic nonlinear firing patterns and associated neural functions. For example, persistent inward currents mediated by their voltage-gated sodium and calcium channels are an important source of the nonlinear behaviour of spinal motoneurons and are instrumental for generating the sustained force outputs required for postural control [19]. Activation of the L-type calcium channels in nigral dopaminergic neurons results in intrinsic bursting behaviour, exhibiting low-dimensional determinism and likely encodes meaningful information in the awake state of the brain [20]. The nonlinearity of the neuronal transfer function mediated by its component ion channels can generate various types of nonlinear output patterns such as harmonic, subharmonic and/or intermodulation of input patterns.

Despite plenty of knowledge of the nonlinear behaviour of a single neuron, the input-output relation at the neural system level is yet to be understood entirely. The system-level neural response is a composite output of collective neuronal activity from a large number of neurons. In a neuronal population, the pool effect can reduce the nonlinear generated by each individual neuron. This has been previously demonstrated by both a computational model and an in vivo study in the human motor system, where the motor command can be transmitted linearly via the mono-synaptic corticospinal tract when there are more than five motoneurons are activated [3]. However, a small amount of nonlinearity may still be present [21]. The nonlinearity in each neuronal population can cumulatively increase if the system involves multiple synaptic connections [22]. A recent study in hemiparetic stroke shows that the nonlinearity in the motor system increases due to an increased usage of multi-synaptic indirect motor pathways (e.g. corticoreticulospinal tract [23]) following damage to the monosynaptic corticospinal tract [24].

Assessing the input-output relation in neural systems (e.g. sensory, motor and cognitive processes) is essential to a better understanding of the nervous system. It could gain a better insight into the normal and pathological neural functions. A linear system generates only iso-frequency interactions between an input and the output, e.g. the coupling of neuronal oscillations at a specific frequency (band) [25]. For decades, correlation and coherence measures have been used to identify the linear interaction in neural systems. More recently, various studies indicate the input-output neural interactions can cross different frequency components or bands (named cross-frequency coupling) [26–28], which is a distinctive feature of a nonlinear system. In the following sections, we review both the linear and nonlinear approaches for identifying neural systems and associate neural connectivity, especially from a system identification perspective.

III. LINEAR CONNECTIVITY AND SYSTEM IDENTIFICATION

The nervous system is a highly cooperative network comprised of different groups of neurons. Neural connectivity, i.e., the synchronization of neural activity across these groups, is crucial to the coordination among distant, but functionally related, neuronal groups [29]. Linear neural connectivity can be assessed by determining the signal correlation or causality between the recorded neural signals. This section reviews commonly used linear connectivity, system identification methods and their interconnections in studying neural systems.

A. Correlation and coherence

The most widely used measure of interdependence between two time series in the time domain is the cross-correlation function [13], which measures the linear correlation between two signals (or stochastic processes) X and Y with discrete observations $x(t)$ and $y(t)$, at $t = 1, 2, ..., N$, as a function of their delay time:

$$C_{xy}(\tau) = \frac{1}{N-\tau} \sum_{k=1}^{N-\tau} x(k+\tau)y(k) \tag{1}$$

where N is the number of samples and τ the time lag between two signals. The function ranges from -1 (complete linear inverse correlation) to 1 (complete linear direct correlation). The value of τ that maximizes this function can be used to estimate the (linearly related) delay between signals. The well-known Pearson correlation coefficient is equal to $C_{xy}(\tau)$ when $\tau = 0$. The linear dependence between two signals in the frequency domain is usually measured by the spectral coherence. The coherence between two signals at frequency f is defined as:

$$C_{xy}(f) = \frac{|S_{xy}(f)|^2}{S_{xx}(f)S_{yy}(f)} \tag{2}$$

where $S_{xy}(f)$ is the cross-spectral density between x and y, and $S_{xx}(f)$ and $S_{yy}(f)$ the auto-spectral density of x and y respectively. The cross-spectral and auto-spectral densities are normally the Fourier transforms of the cross-correlation and auto-correlation functions of the two signals. Values of coherence are always between 0 and 1. The correlation and coherence measures have been widely applied to EEG, MEG or EMG signals to characterise the neuronal interactions, from the firing of cortical neuron spike trains to complicated neural systems (for reviews see [13, 30]).

B. System identification and causality

Unlike functional connectivity, effective connectivity emphasises on the directional causal influences between

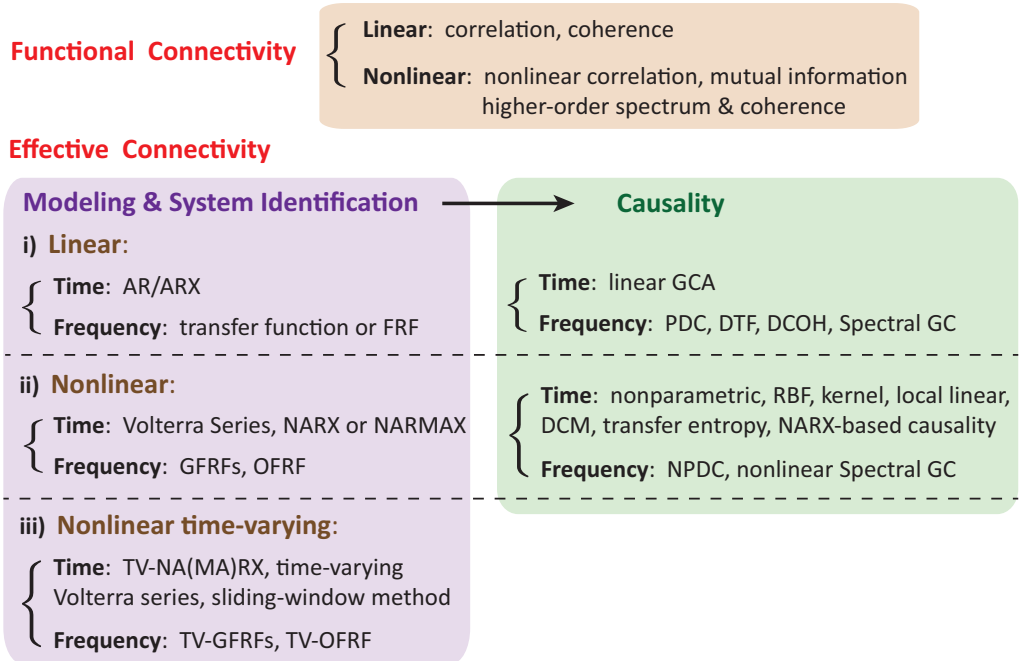

FIG. 1: Overview of the linear and nonlinear functional and effective connectivity (causality) measures and their links with system identification methods. The linear functional connectivity, linear system identification and linear causality measures are first reviewed in Section III. The nonlinear and nonlinear time-varying system identification approaches (in both time and frequency domains) are then investigated in Section IV. The recently proposed nonlinear function connectivity measures and nonlinear causality measures (based on nonlinear system identification) are introduced in Section V.

neural areas or signals. Here, we first introduce the classical Granger causality and its link with the time-domain linear system identification (i.e. regression models of bivariate time series). The frequency-domain causality measures can then be linked with the frequency response function of linear systems.

Consider two signals/variables X and Y, the interactions of the signals can be described by bivariate linear autoregressive with exogenous input (ARX) models jointly,

$$\begin{aligned} x(t) &= \sum_{k=1}^{q} a_{11,k} x(t-k) + \sum_{k=1}^{p} a_{12,k} y(t-k) + e_x(t) \\ y(t) &= \sum_{k=1}^{q} a_{21,k} x(t-k) + \sum_{k=1}^{p} a_{22,k} y(t-k) + e_y(t) \end{aligned} \tag{3}$$

where p and q are the model order of y and x regressors; $e_x(t)$ and $e_y(t)$ are the model (uncorrelated) prediction errors over time. A linear causal influence from X to Y defined by Granger can be expressed as the log ratio of the prediction error variances of the corresponding restricted (AR) and unrestricted (ARX) models:

$$F_{X\to Y} = \ln \frac{var(y|y^-)}{var(y|y^-, x^-)} = \ln \frac{\Sigma_{22}^{y^-}}{\Sigma_{22}} \tag{4}$$

where x^- and y^- denotes contributions from lagged input and output terms respectively; ${y^-}_{22}$ denotes the variance of e_y when there are only regression terms of Y. The linear ARX models (3) can be rewritten in matrix form and mapped to the frequency domain by Fourier transformation:

$$\begin{pmatrix} A_{11}(f) & A_{12}(f) \\ A_{21}(f) & A_{22}(f) \end{pmatrix} \begin{pmatrix} X(f) \\ Y(f) \end{pmatrix} = \begin{pmatrix} E_x(f) \\ E_y(f) \end{pmatrix} \tag{5}$$

where the components of the coefficient matrix $\mathbf{A}(f)$ are $A_{lm}(f) = \delta_{lm} - \sum_{k=1}^{p(or\ q)} a_{lm,k} e^{-j2\pi kf/f_s}$ with f_s the sampling frequency and δ_{lm} the Kronecker delta function. We can re-write the above equation by inverting the coefficient matrix $\mathbf{G}(f) = \mathbf{A}^{-1}(f)$ and moving the so-called transfer function matrix $\mathbf{G}(f)$ to the right-hand-side the equation. Different frequency-domain Granger causality measures, such as partial directed coherence (PDC), directed transfer function (DTF), spectral Granger causality and directed coherence (DCOH) [31–33], can then be expressed as a function of the elements of either the coefficient matrix $\mathbf{A}(f)$ or the transfer function matrix $\mathbf{G}(f)$ (Baccala and Sameshima, 2001; Chicharro, 2011). By dividing both sides of (5) with the corresponding diagonal elements in the coefficient matrix $\mathbf{A}$, the off-diagonal elements in the transformed coefficient matrix are actually related to the negative frequency response functions (FRFs) of the linear (ARX) systems, if one signal is treated as the input while the other is treated as the output. For instance,

$$\frac{A_{21}(f)}{A_{21}(f)} = -H_{X\to Y}(f) \tag{6}$$

the FRF, $H_{X\to Y}(f)$, describes the input-output relationship (i.e., with input X and output Y) of the (noise-free) 'system' in the frequency domain. It is also known as the 'transfer function' in the linear system theory. Frequency-domain Granger causality measure, e.g. PDC, can be expressed as a function of the FRFs of the corresponding linear ARX and AR models:

$$PDC_{X\to Y}(f) = \frac{-H_{X\to Y}(f)}{\sqrt{|A_{11}(f)/A_{21}(f)|^2 + |H_{X\to Y}(f)|^2}} \tag{7}$$

Establishing such a link between the causality measures and linear system identification (in both time and frequency domains) is crucial to the further development of nonlinear model-based causality measures via nonlinear system identification, which will be investigated in Section IV and V.

C. Limitation of linear approaches on identifying neural system

Linear connectivity and system identification allow the assessment of communication between neuronal populations at the same oscillatory frequency band or similar neuronal firing patterns. The applications of linear approaches have been thoroughly reviewed previously [34, 35]. However, it is not clear how much information is missing when using the linear approach, since the behaviour of various neural systems can be highly nonlinear [9, 10]. When one uses a linear measure to investigate a neural system, the nonlinear neural interaction is ignored, especially between the neuronal populations which have very different mean firing rates such as the central nervous system and the periphery. A recent study in human somatosensory system reported that over 80% of the cortical response to wrist joint sensory input comes from nonlinear interactions, where a linear model explains only 10% of the cortical response [36]. Therefore, nonlinear connectivity and modelling approaches are needed to investigate neural systems in a more complete way.

IV. NONLINEAR SYSTEM IDENTIFICATION OF NEURAL SYSTEMS

It is often impossible to derive a mechanistic model of a neural system, due to the complexity of the underlying biological process and many unknown state variables are unobservable. In this section, we focus on the generic nonlinear model representation of a single input single output (SISO) neural dynamic system, its identification process in the time domain, and corresponding frequency-domain analysis. We first investigate the identification of nonlinear time-invariant systems, and then time-varying nonlinear systems.

A. Time-domain nonlinear system identification

1. *Volterra series*

The Volterra series model is a direct generalisation of the linear convolution integral and provides an intuitive representation for a nonlinear input-output system. The output $y(t)$ of a SISO nonlinear system can be expressed as a Volterra functional of the input signal $u(t)$:

$$\begin{aligned} y(t) &= \sum_{n=1}^{M} y^{(n)}(t) + e_y(t) \\ y^{(n)}(t) &= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} h_n(\tau_1, \cdots, \tau_n) \prod_{i=1}^{n} u(t-\tau_i) d\tau_i \end{aligned} \tag{8}$$

where $y^{(n)}(t)$ is the nth-order output and M is the maximum order of the system's nonlinearity; $h_n(\tau_1, \cdots, \tau_n)$ is the nth-order impulse response function or Volterra kernel, which describes nonlinear interactions among n copies of input and generalises the linear convolution integral to deal with nonlinear systems. Neurobiologically, Volterra series can be directly interpreted as the effective connectivity - 'the influence that one neural system exerts over another, either at a synaptic or population level' [37]. The first-order kernel describes the linear 'driving' efficacy or linear synchronous interactions, and the second- or higher-order kernels represent the 'modulatory' influence or asynchronous interactions [9]. The Fourier transform of the first-order kernel is the FRF (or transfer function) and describe the interactions in the same frequencies, while the frequency-domain counterparts of the higher-order kernels are the GFRFs (to be discussed in Section IV B) which quantifies the cross-frequency interactions.

Volterra model has been widely used in physiological systems (including neural systems) modelling, including the study of nonlinear interactions in hippocampal-cortical neurons [38], in the spectrotemporal receptive fields of the primary auditory cortex [39], in the sensory mechanoreceptor system [36], in multiple-input and multiple-output (MIMO) spiking neural circuits [40] and hippocampal memory prostheses [41]. To deal with large number of coefficients, a group regularised (LASSO) estimation approach is often employed. The Volterra model also has a strong theoretical link with the NARMAX model [42] and the dynamic causal modeling [43] to be discussed next.

2. *NARMAX model*

Although Volterra series can provide an intuitive representation for nonlinear systems, there are several critical limitations including i) it cannot represent severely nonlinear systems; ii) the order of the Volterra series expansion can be very high in order to achieve a good approximation accuracy; however iii) the estimation of high

order Volterra kernel requires a large number of data and can be computationally very expensive. Nonlinear Autoregressive Moving Average Model with Exogenous Inputs (NARMAX) [42, 44] has therefore been developed as an alternative to the Volterra series. NARMAX model normally contains a much less number of terms (due to the inclusion of output delay terms), and its identification process is computationally more efficient. Similar to the Volterra series, a polynomial Nonlinear Autoregressive Model with Exogenous Inputs, NARX model (the simplest NARMAX), can be expressed as a summation of a series of output terms with different orders of nonlinearity:

$$y^{(n)}(t) = \sum_{p=0}^{n} \sum_{k_1,k_{p+q}=1}^{K} c_{p,q}(k_1, ..., k_{p+q}) \times \prod_{i=1}^{p} y(t-k_i) \prod_{i=p+1}^{p+q} x(t-k_i) \quad (9)$$

where $p+q = n, k_i = 1, ..., K$, and $\sum_{k_1,k_{p+q}=1}^{K} \equiv \sum_{k_1}^{K} \cdots \sum_{k_{p+q}=1}^{K}$. The number of model terms depends on the order of input and output (q and p) and the maximum lags (K). The NARX model structure and parameters are typically identified based on the forward regression with orthogonal least squares (FROLS) method [45]. In cases where the system under study is stochastic with unknown coloured noise, noise moving average (MA) models should be employed to form a full NARMAX model. The identified model can be statistically validated using the nonlinear correlation tests [46, 47].

A wide range of nonlinear (including severely nonlinear) systems can be represented by NARMAX method, including exotic nonlinear behaviours such as harmonics, bifurcations, and chaos, as observed in the human nervous system [48]. Until now, NARMAX methodology has been employed to develop dynamic models for nonlinear sensory processing circuit from spiking neuron data [49] (as an improvement to the previous Volterra model-based studies [40]) to investigate the somatosensory afferent pathways (from muscles to the brain) [50, 51]; as well as to study the corticothalamic nonlinear interactions during tremor active and resting states [8]. Apart from efficient time-domain predictive modelling, NARMAX is also an important base for the nonlinear frequency-domain analysis, nonlinear time-varying modelling, and nonlinear causality analysis to be discussed in the following sections.

3. *Dynamic causal modelling*

Most of the effective connectivity models (e.g. linear and nonlinear autoregressive models) are directly identified from functional neurophysiological signals. However, sometimes it would be more accurate and meaningful to identify the causal interactions (of the underlying neuronal activity) at the level of neuronal dynamics [52]. The aim of dynamic causal modelling (DCM) [43, 53] is to infer such connectivity among brain regions (or sources) under different experimental factors or inputs. A DCM normally comprises two parts: a neuronal part that describes dynamics among neuronal sources and a measurement part that describes how the source dynamics generate measurable observations (e.g. EEG or MEG) [54, 55]. Therefore, DCM can be naturally expressed as a nonlinear state-space model with hidden states denoting unobserved neuronal dynamics and the observation equation (e.g. the lead-field) assumed linearly in the states. The effective connectivity among those sources can be identified via Bayesian model selection and Bayesian inference of the neuronal model parameters. One strength of DCM is its biophysical and neuronal interpretation of how the neurophysiological signals are generated from the underlying neuronal system, through the generative or forward (state-space) modelling. Due to the complexity and computational cost of Bayesian model selection, DCM is more suitable to investigate the connectivity among pre-defined regions of interest, rather than exploratory analysis of relatively large brain or neuronal networks [56]. Compared to the hypothesis-driven DCM, the NARMAX or Volterra models are more flexible in terms of model structure identification and their direct frequency-domain mapping (to be discussed) is a powerful tool to study the nonlinear cross-frequency interactions between neurological regions.

B. Frequency-domain nonlinear system analysis: nonlinear frequency response functions

Many nonlinear effects, such as harmonics, intermodulations, and energy transfer, can only be accurately and intuitively characterised in the frequency domain. Thus, it is important to 'map' the identified time-domain nonlinear models to the frequency domain for further analysis. A multidimensional Fourier transform of the nth-order Volterra kernel in (8) yields the so-called nth-order generalised frequency response function (GFRF): $H_n(f_1, \cdots, f_n)$, which is a nature extension of the linear frequency response function (or transfer function) to the nonlinear case [42]. The output spectrum $Y(f)$ of a nonlinear system can then be expressed as a function of the input spectrum $X(f)$ and GFRF, known as the output frequency response function (OFRF) [42, 57]:

$$Y(f) = \sum_{n=1}^{M} \Big(\frac{1}{\sqrt{n}} \int_{f_1+\cdots+f_n=f} H_n(f_1, \cdots, f_n) \prod_{i=1}^{n} X(f_i) df \Big) \quad (10)$$

Compared with the Volterra series, the GFRFs can be more efficiently computed from the identified time-domain NARMAX model (9) and corresponding model parameters [58]. As shown in Figure 2, the peaks in 1st-order GFRF indicate the well-known 'resonance fre-

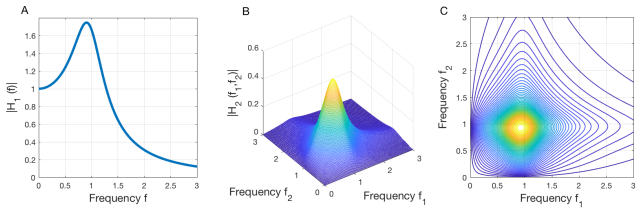

FIG. 2: The GFRFs of an exemplar nonlinear system. (A) The linear 1st-order GFRF, $H_1(f)$, shows a 'resonance' peak at $f = 0.9$Hz; (B) and (C) the 3-D and contour plots of the 2nd-order GFRF, $H_2(f_1, f_2)$. It shows a peak at $f_1 = 0.9Hz, f_2 = 0.9Hz$, which indicates harmonics at $2f = f_1 + f_2 = 1.8Hz$ can be introduced in the output spectrum if input contains a $0.9Hz$ component.

quencies' of the linear part of the system; and the peaks (or ridges) in the 2nd-order GFRF would indicate (nonlinear) harmonics ($f_1 + f_2$ when $f_1 = f_2$) or intermodulation effects ($f_1 \pm f_2$ when $f_1 \neq f_2$) in the output spectrum, and so on. Since $Y(f) = \sum_{n=1}^{m} Y_n(f)$, the nth OFRF $Y_n(f)$ indicates the nth-order (linear or nonlinear) contribution from the input to the output spectrum. Practically, by comparing the OFRF with the spectrum of output signal (obtained from a classical nonparametric estimation such as fast Fourier transform), one can also 'validate' the accuracy of the time-domain modelling in addition to the aforementioned NARMAX model validation [8].

NARMAX-based frequency-domain analysis method has been applied to quantify the dynamic characteristics of nonlinear sensory processing circuit model from spiking neuron data [49], the cross-frequency interactions in the corticothalamic loops with respect to tremor [8], and the characterisation of epileptic brain states [47]. More details will be discussed in Section VI.

C. Time-varying nonlinear system identification

Many neurological subsystems are inherently nonstationary, since the brain is a dissipative and adaptive dynamical system [59, 60]. Modelling and identification approaches of nonstationary processes have been well developed for linear systems, i.e. linear time-varying (LTV) systems. It is primarily based on adaptive recursive methods, such as recursive least squares, least mean squares, and the Kalman filter [61], or based on a finite basis function expansion of the time-varying coefficients [62–65]. The identification of a nonlinear time-varying system is more sophisticated. The primary difficulty is how to effectively combine the nonlinear model structure selection and the time-varying coefficient estimation processes. Approaches based on time-varying Volterra series combining artificial neural networks [66], principal dynamic modes [67], or sliding-window based strategy [68] have been proposed. However, the model structure selection is still an unsolved issue, and the identification and frequency-domain analysis are computationally costly.

A better strategy is to extend the basis function expansion approach (proposed for LTV identification) to nonlinear time-varying cases [69]. The time-varying (TV) parameters in TV-NARX models are first expanded using multi-wavelet basis functions, and TV nonlinear model is transformed into an expanded time-invariant model; the challenging TV model selection and parameter estimation problem can then be solved using the computational efficient FROLS algorithm. To accommodate the stochastic perturbations or additive coloured noise, this procedure can also be extended to more general TV-NARMAX models by introducing a modified extended least squares (ELS) algorithm [70]. Several modifications to the TV-NARX model has recently been proposed using different basis functions or model selection procedure [71]. The corresponding frequency-domain analysis for nonlinear time-varying systems has been recently developed based on the identified time-domain (TV-NARX or TV-NARMAX) model and the TV-GFRF concepts [70, 72]. The TV-NARX model has been applied to two fragments of intracranial EEG recordings measured from epileptic patients. The corresponding frequency-domain analysis (i.e. TV-GFRF and TV-OFRF) shows the nonlinear energy transfer effect – the underlying neural system transfers the energy from low frequencies to higher frequencies as seizure spreading (over time) from the left to the right brain areas [72, 73].

The implementation of NARMAX model-based system identification framework, including both time-invariant (NARMAX) and time-varying (TV-NARX) modelling as well as corresponding frequency-domain analysis to neurophysiological signal analysis, is summarised in Figure 3.

V. NONLINEAR NEURAL CONNECTIVITY ANALYSIS

The communication between different neuronal populations which have very different firing behaviours can result in nonlinear neural connectivity, showing neural coupling across two or more different frequencies. To quantitatively study such a 'cross-frequency coupling', this section reviews recent advances in nonlinear neural functional and effective connectivity analysis.

A. High-order spectrum and nonlinear coherence

The power spectra and coherence discussed in Section III A are Fourier transforms of the auto- and cross-correlations of signals, hence they are only linear frequency-domain measures. Practically these measures cannot detect certain nonlinear effects such as quadratic moments in or between signals that have a zero mean [42]. Higher-order spectral analysis has been developed to detect nonlinear coupling between spectral compo-

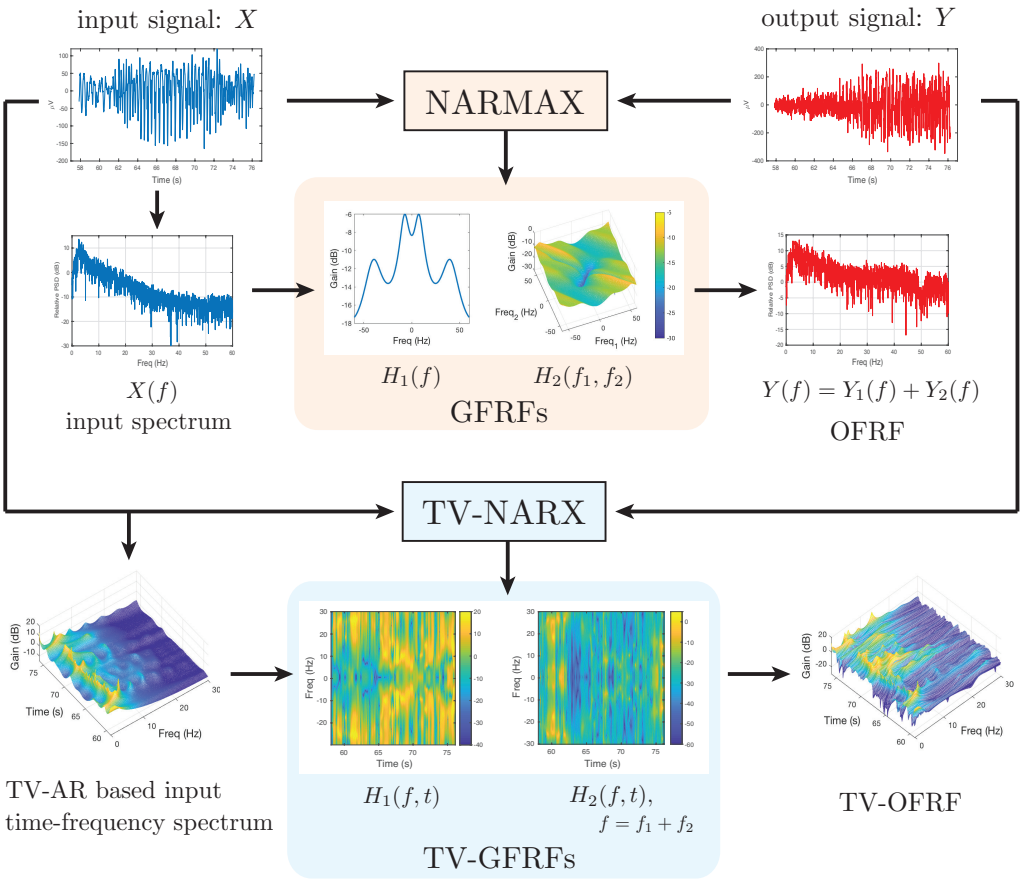

FIG. 3: Analysing neurophysiological signals using nonlinear time-invariant and time-varying system identification and corresponding frequency-domain analysis methods. The upper part of the diagram illustrates the nonlinear time-invariant modelling: first a NARMAX model is identified from the input and output neurophysiological signals (e.g. EEG, EMG, MEG, LFP); this time-domain model is then mapped to the frequency-domain with GFRFs (i.e. $H_1(f)$, $H_2(f)$, ...), and the OFRF ($Y(f) = Y_1(f) + Y_2(f), ...$) can be computed from the input spectrum and GFRFs. The lower part of the diagram shows the nonlinear time-varying system identification using a TV-NARX model, and the identified time-varying model can then be mapped to the time-frequency domain with (averaged) TV-GFRFs (i.e. $H_1(f,t)$, $H_2(f,t)$, ...). The TV-OFRF can therefore be computed from a combination of the input time-frequency spectrum and the TV-GFRFs.

nents [74]. For example, the most widely used bispectral analysis is the frequency-domain mapping of the third-order statistics, and can be used to quantify the quadratic nonlinear interactions, i.e. quadratic phase coupling. The bispectrum or bicoherence (or the bivariate cross-biespectrum/-bicoherence) analysis is well-known in engineering signal processing, but it has only relatively recently been applied to study the nonlinear interactions in neurophysiological signals [75–77]. For example, bispectral measures were used to detect long-range nonlinear coupling/synchronization in healthy subjects from human EEG [76, 78], to characterise and predict epileptic seizures [79], and to study the nonlinear interactions between different frequency components related to Parkinson's disease and tremor [8, 80, 81].

However, bispectrum or bicoherence cannot detect nonlinearity beyond second order, such as the higher-order harmonics and intermodulation effects, or the subharmonic coupling. A generalised nonlinear coherence analysis framework has therefore been proposed, based on two different nonlinear mappings from the input to the output (of an open-loop system) in the frequency domain [5]:

1) n:m Mapping: to measure harmonic or subharmonic coupling related to individual input frequencies. $Y^m(f_Y) = H(n:m)X^n(f_X)$, $X^n \equiv \underbrace{XX\cdots X}_{n}$, where the output frequencies (f_Y) are related to the input frequencies (f_X) by the ratio n/m (n and m are co-prime positive integers), and $H(n:m)$ is the n:m mapping function. The n:m mapping can generate cross-frequency (e.g. harmonic $m = 1$ or subharmonic $m > 1$) coupling between the input and the output.

2) Integer Multiplication Mapping: to quantify intermodulation coupling among multiple (≥ 2) input fre-

quencies.

$$Y(f_Y) = H(f_1, f_2, \cdots, f_N; a_1, a_2, \ldots, a_N; f_Y) M \prod_{i=1}^{N} X^{a_i}(f_i) \tag{11}$$

where $f_Y = a_1 f_1 + a_2 f_2 + \cdots + a_N f_N$. The M is the corresponding multinomial coefficient, and $H(f_1, f_2, \cdots, f_N; a_1, a_2, \ldots, a_N; f_Y)$ indicates amplitude scaling and phase shift from the input to the output.

According to these two different types of nonlinear mapping, Yang and colleagues proposed two basic metrics for quantifying nonlinear coherence: (i) n:m coherence and (ii) multi-spectral coherence [5].

1. n:m coherence

The n:m coherence is a generalized coherence measure for quantifying nonlinear coherence between two frequency components of the input $X(f)$ and the output $Y(f)$ [5]:

$$C_{nm}(f_X, f_Y) = \frac{|S_{XY}(f_X, f_Y)|}{\sqrt{S^n_{XX}(f_X) S^m_{YY}(f_Y)}} \tag{12}$$

where $n : m = f_Y : f_X$. $S_{XY}(f_X, f_Y) =< X^n(f_X)(Y^m(f_Y))^* >$ is n:m cross-spectrum, with $< \cdot >$ represents the averaging over repetitions. $S^n_{XX}(f_X) =< X^n(f_X)(X^n(f_X))^* >=< |\underbrace{X(f_X)X(f_X)\cdots X(f_X)}_{n}|^2 >$ is the nth order auto-spectra. According to Cauchy-Schwarz-inequality, we have:

$$\begin{aligned} &| < X^n(f_X)(Y^m(f_Y))^* > | \\ &\quad \leq \left(< |X^n(f_X)|^2 >\right)^{1/2} \cdot \left(< |Y^n(f_Y)|^2 >\right)^{1/2} \end{aligned} \tag{13}$$

Thus, n:m coherence is bounded by 0 and 1, where 1 indicates that two signals are perfectly coupled for the given frequency pair (f_X, f_Y).

A simplified version of n:m coherence that considers only the phase relation between the input and the output is known as n:m phase synchronization index [82]. The n:m coherence and n:m phase synchronization index has been widely applied to neuroscience and brain research to investigate nonlinear functional connectivity in the brain [83, 84], as well as the nonlinear connectivity between the brain and muscles [21].

2. Multi-spectral coherence

Multi-spectral coherence measures the multi-frequency nonlinear coupling generated by the integer multiplication mapping [5]. It is defined as:

$$\begin{aligned} &C_{MS}(f_1, \cdots, f_N; a_1, \cdots, a_N; f_Y) \\ &\quad = \frac{|S_{XY}(f_1, \cdots, f_N; a_1, \cdots, a_N; f_Y)|}{\sqrt{(\prod_{i=1}^{n} S^{|a_i|}_{XX}(f_i)) S_{YY}(f_Y)}} \end{aligned} \tag{14}$$

where $f_Y = a_1 f_1 + \cdots + a_N f_N$; $S_{XY}(f_1, \cdots, f_N; a_1, \cdots, a_N; f_Y)$ is the high-order cross-spectrum between X and Y, and equal to $< \prod_{i=1}^{n} X^{|a_i|}(f_i)) Y^*(f_Y) >$. Here '*' indicates the complex conjugate. When there are only two input frequencies, the multi-spectral coherence is degraded to the bicoherence [85]. The multi-spectral coherence or bicoherence has been applied to study the nonlinear behaviours in visual [86], auditory [87] and somatosensory systems [5], which are thought to be associated with neural coding and functional integration of various sensory inputs [88].

A simplified version of multi-spectral coherence that considers only the phase relation between the input and the output is known as multi-spectral phase coherence [89]. Similarly, there is a degraded measure, named bi-phase locking value [90], for the case involving only two input frequencies. The advantage of multi-spectral phase coherence or bi-phase locking value is that it allows precise estimation of time delay in the nervous system based on the relative phase relationship between the input and output [89, 91]. The multi-spectral phase coherence or bi-phase locking value has been previously used to determine neural transmission delays in the human visual system [86] and the stretch reflex loops [21]

B. Nonlinear causality analysis: system identification based approaches

1. Time-domain analysis

In terms of effective connectivity, classical linear Granger causality analysis (GCA) (as discussed in Section III B) may provide misleading results when used to analyse EEG/MEG or EMG signals, as the possible nonlinear interactions within a neural system are not modelled explicitly by simply using linear regression models. The Granger causality definition has been extended to nonlinear cases in the time domain, based on nonparametric methods [92, 93], radial basis functions [94], kernel methods [95], local linear models [96]. Dynamic causal modelling (DCM) [43] (see Section IV A 3) was developed to accommodate both linear and nonlinear causal effects using a dynamic state-space model, and the effective connectivity among hidden states (unobserved neuronal dynamics) can be identified via Bayesian inference. Information-theoretical effective connectivity measures have also been proposed, such as the bivariate transfer entropy (TE) [97, 98] and phase transfer entropy (PTE), which use conditional probabilities (via the Kullback-Leibler divergence) to quantify the amount of uncertainty in the future target signal conditioned on only the target's past compared to both its own past and the source's past (in a conceptually similar way as the GCA). A more recent work [99] generalised the TE method by using multivariate transfer entropy, which can overcome the problem of inferring spurious or redundant causality and missing synergistic interactions between

multiple sources and target.

Another strategy to implement nonlinear granger causality under a system identification framework is to use the NARX models [100, 101], by calculating the relative predictability improvement obtained from the NARX model compared to the corresponding NAR model. More importantly, compared to other nonlinear causality measures (e.g. nonparametric or information-theoretic measures), the advantage of using NARX-based causal inference [101] is that one can easily separate the linear and nonlinear causal influence, for example from an input X to an output Y. After fitting a polynomial NARX model with the form (9), the linear causality can still be calculated from (4) (from the linear parts of the NARX model), while the nonlinear causal influence from X to Y can be defined as:

$$F^n_{X\to Y} = \ln \frac{var(Y|Y_l^-, Y_n^-, X_l^-)}{var(Y|Y_l^-, Y_n^-, X_l^-, X_n^-, (XY)_n^-)} \quad or \quad \ln \frac{var(Y|X_l^-)}{var(X_l^-, X_n^-, (XY)_n^-)} \tag{15}$$

Here Y_n^-, X_n^- and $(XY)_n^-$ denote the sets of all nonlinear delayed terms of Y, X and nonlinear product terms XY. This nonlinear causality measure can also be generalised to nonlinear time-varying systems, by computing similar linear and nonlinear causality indices based on the identified TV-NARX models (as described in Section IV C), as proposed in [101, 102].

2. Frequency-domain analysis

In the frequency domain, linear Granger causality measures, such as PDC, DTF and spectral Granger causality, can all be expressed as a function of the elements of either the coefficient matrix or its inverse the transfer function matrix of the corresponding linear ARX models (3). By identifying the link between the PDC and the FRFs of the corresponding linear ARX models (as described in Section III B), a new nonlinear PDC (NPDC) measure has been proposed [47] by generalising the spectrum decomposition with respect to a nonlinear NARX model in a similar way as the linear case. The NPDC from X to Y can then be expressed as a direct generalization of the linear PDC:

$$NPDC_{X\to Y}(f) = \frac{-H_{X\to Y}(f)}{\sqrt{|H_y^e(f)/H_x^e(f)|^2 + |H_{X\to Y}(f)|^2}} \tag{16}$$

Here, the $H_{X\to Y}(f)$ is the 'nonlinear FRF' which replaces the FRF in the linear PDC (7), and $H_y^e(f)$ and $H_x^e(f)$ are the error-driven GFRFs correspond to the restricted NAR models with respect to Y and X as discussed in [47]. The NPDC measures both linear and nonlinear causal influences from X to Y. The linear causal effects can be expressed as a special case of (16) by only considering the 1st-order nonlinear FRFs of NARX (i.e. $H_{1,X\to Y}(f)$) and NAR (i.e. $H^e_{1,x}(f)$ and $H^e_{1,y}(f)$) models.

This new NPDC measure has recently been applied to predict epileptic seizures [103] by advancing the construction of functional brain networks, nonlinear feature selection and classification procedure, which provides better prediction accuracy compared to other standard graph theory or nonlinear classification based methods. A nonlinear generalization of Geweke's spectral Granger causality has also been proposed [73] using the NARX methodology.

VI. NEUROLOGICAL AND CLINICAL APPLICATIONS

A. Nonlinear cortical response to somatosensory inputs

The human somatosensory system is highly nonlinear [10]. Previous studies applied periodic sinusoid tactile stimulations to the index finger and measured the cortical response, where they found harmonic and subharmonic patterns in the response [4, 104]. Several recent studies used sum-of-sinusoid stimulations to the wrist joint and found not only harmonics/subharmonics but also intermodulation patterns [5, 36]. The majority of intermodulation responses presented the second-order nonlinearity, which is the sum or the difference between input frequencies [7]. These findings indicate that the nonlinearity in the somatosensory system allows the functional integration of input signals at different frequencies and transmitted in different somatosensory ascending pathways.

Yang and colleagues recently built a hierarchical neural network based on known neuroanatomical connections and corresponding transmission delays in neural pathways to model the cortical response to somatosensory input [50]. The proposed computational model contains a neural layer at the thalamus that integrates the inputs from different ascending pathways, including Ia and II afferents. The computational model well captured the majority of the cortical response to the given somatosensory input, indicating that the functional integration of different somatosensory input signals may occur at the thalamus and is then transmitted to the cortex via the thalamocortical radiation.

B. Tremor: nonlinearity in the thalamocortical loop

Essential tremor is a common neurological movement disorder widely considered to have a centrally-driven origin. There is both neurophysiological and clinical evidence of thalamic involvement in the central oscillatory network generating essential tremor [105–107], and local field potential (LFP) recordings of thalamic ventralis in-

termedius (Vim) nucleus show a strong linear correlation with the contralateral electromyography (EMG) during tremor [106]. Some studies using EEG and magnetoencephalography (MEG) suggest the sensorimotor cortex is also part of the central tremor-related oscillatory network with significant coupling in some cases with the contralateral tremorgenic EMG [108–110]. Despite a well-established reciprocal anatomical connection between the thalamus and cortex, the functional association between the two structures during 'tremor-on' periods had not been extensively defined.

He and co-authors [8] investigated the functional connectivity among cortical EEG, thalamic (Vim) LFPs and contralateral EMG signals over both 'tremor-on' and 'tremor-off' states, using linear (i.e. coherence) and nonlinear (i.e. bispectral analysis) methods. In addition to expected strong coherence between EMG and thalamic LFP, nonlinear interactions (i.e. quadratic phase coupling) at different frequencies (i.e. low frequency during tremor on and higher frequency during tremor off) in LFPs have been reported. More importantly, by using the NARX-based system identification and frequency-domain analysis (as described in Section IV B), two distinct (and not overlapping) frequency 'channels' of communication between thalamic Vim and the ipsilateral motor cortex were identified, which robustly defines the 'tremor-on' versus 'tremor-off' states. Longer corticothalamic (nonlinear) lags in the tremor active state is also uncovered, suggesting the possible involvement of an indirect multi-synaptic loop. This work demonstrates, for the first time, the importance of cross-frequency nonlinear interactions between the cortex and the thalamus in characterising the essential tremor.

C. Nonlinear analysis for determining motor impairment in stroke

After a stroke, damage to the brain increases reliance on indirect motor pathways resulting in motor impairments and changes in neural connectivity between the brain and muscles. A hallmark of impairments post-stroke is a loss of independent joint control that leads to abnormal co-activation between shoulder, arm and hand muscles, known as the upper limb synergy [111]. The upper limb synergy is thought to be caused by progressive recruitment of indirect motor pathways via the brainstem following a stroke-induced loss of corticospinal projections [112]. Thus, a neural connectivity measure that quantifies the recruitment of these indirect motor pathways would be crucial to evaluate post-stroke motor impairments. Recent model simulation and clinical studies indicate that the increased usage of indirect motor pathways enhances nonlinear distortion of motor command transmissions, which leads to stronger nonlinear interaction between the brain and muscles [22, 24]. The ratio of nonlinear interaction over linear interaction, known as the nonlinear-over-linear index (N-L Index), has been reported to be associated with the relative ratio of the recruitment of indirect versus direct motor pathways [24]. This new measure may facilitate the future determination of the effect of new therapeutic interventions that aim to optimise the usage of motor pathways, and thus facilitate the stroke recovery.

D. Epilepsy

It has been widely recognised that epileptic seizures are highly nonlinear phenomena, due to low dimensional chaos during epileptic seizure or transitions between ordered and disordered stats [113]. Currently, the treatment mainly relies on long-term medication with antiepileptic drugs or neurosurgery, which can cause cognitive or other neurological deficits. New treatment strategies such as on-demand therapies during the pre-seizure (preictal) state or electrical stimulation are therefore needed. A vital part of this new on-demand strategy is the accurate and timely detection of the preictal state, even seconds before seizure onset [114]. A range of univariate, bivariate and multivariate linear and nonlinear measures have been developed for the characterisation and detection/prediction of epileptic brain states and gain a better understanding of the spatial and temporal dynamics of the epileptic process [114, 115], including a comprehensive review of using different parametric and nonparametric nonlinear features (in time, frequency and time-frequency domains) for the automated epilepsy stage detection and classification [116].

Given the current challenges in epilepsy detection and diagnostics [114, 117] (e.g. to improve the understanding of brain dynamics/mechanisms during the seizure-free interval and seizure initiation and termination), there is a great need to develop more accurate nonlinear methods to improve the detectability of directional interactions in underlying functional and anatomical networks - developing new nonlinear system identification and nonlinear causality measures are therefore crucial. A nonlinear causality measures, partial transfer entropy [118], has been applied to analyse the EEG of epileptic patients during preictal, ictal and postictal states, and can provide better detection of causality strength variations compared to linear PDC. An adaptive nonlinear Granger causality measure [119] was also proposed and applied to LFP data (intracranial EEG in cortex and thalamus) in rats. It was reported to provide more sensitive detection of changes in the dynamics of network properties compared to linear Granger causality. The recently proposed nonlinear frequency-domain causality measure NPDC [47] (reviewed in Section V B 2) has been applied to analyse EEG recordings of two bipolar channels from a patient with childhood absence epilepsy. It shows this nonlinear measure can detect extra frequency-domain causal interactions compared to standard linear PDC.

VII. CONCLUSIONS AND PERSPECTIVES

The nonlinearity and complexity of neural systems require advanced system identification approaches to understand its properties and mechanisms. This review investigated the links between connectivity analysis and system identification, as well as recent progress of nonlinear system identification of neural systems. With the state-of-the-art examples of clinical applications, we argued that nonlinear dynamic modelling and corresponding connectivity analysis allow new insights into the underlying neural functions and neuropathological mechanisms of the abnormality caused by various neurological disorders. These novel approaches may facilitate the development of precision diagnostic tools and, therefore, improve the diagnosis and treatment of neurological disorders.

Compared to the linear system identification and iso-frequency connectivity analysis, nonlinear dynamic modelling and cross-frequency analysis are much more complicated. Such complexity brings challenges but also research opportunities. Potential future work includes: i) further develop multivariate system identification techniques and corresponding multivariate nonlinear frequency-domain analysis and causality analysis measures. Most existing nonlinear system identification based (time or frequency-domain) analysis or causality analysis are primarily bivariate, which limits nonlinear analysis to only pre-defined local (brain or neuronal) regions. New multivariate system identification (e.g. multivariate NARMAX modelling) or inference approaches would generalise the existing nonlinear connectivity analysis to larger neuronal networks, although developing efficient model selection and reduce the computational cost would be challenging tasks; ii) many neuronal systems or interactions are in nature nonstationary and nonlinear, how to automatically distinguish the nonlinearity and time-varying effects (nonstationarity) via novel system identification technique is still an open and important research topic, although significant progress has been made (as reviewed in Section IV C); iii) machine learning and deep learning techniques have recently been applied to Granger causality analysis [120–122], an interesting future work is to further explore and combine the advantages of deep learning (e.g. accurate quantification of complex and long-range nonlinear interactions) and nonlinear system identification approaches to study the nonlinear causal interactions in complex neuronal networks; iv) apart from those clinical applications described in Section VI, the importance of nonlinearity in understanding and characterising other important neurological disorders, e.g. Parkinson's disease and Alzheimer's disease [123–125], has been reported recently. Therefore nonlinear system identification approaches will have great potential in developing new diagnostic tools for those critical neurological disorders that affect a large population worldwide.

[1] Christof Koch and Idan Segev. The role of single neurons in information processing. *Nature neuroscience*, 3(11):1171–1177, 2000.

[2] Robert Rosenbaum, James Trousdale, and Kresimir Josic. Pooling and correlated neural activity. *Frontiers in computational neuroscience*, 4:9, 2010.

[3] Francesco Negro and Dario Farina. Linear transmission of cortical oscillations to the neural drive to muscles is mediated by common projections to populations of motoneurons in humans. *The Journal of physiology*, 589(3):629–637, 2011.

[4] Angela J Langdon, Tjeerd W Boonstra, and Michael Breakspear. Multi-frequency phase locking in human somatosensory cortex. *Progress in biophysics and molecular biology*, 105(1-2):58–66, 2011.

[5] Yuan Yang, Teodoro Solis-Escalante, Frans CT van der Helm, and Alfred C Schouten. A generalized coherence framework for detecting and characterizing nonlinear interactions in the nervous system. *IEEE transactions on biomedical engineering*, 63(12):2629–2637, 2016.

[6] Jing Ren, Jing Xiang, Yueqiu Chen, Feng Li, Ting Wu, and Jingping Shi. Abnormal functional connectivity under somatosensory stimulation in migraine: a multi-frequency magnetoencephalography study. *The journal of headache and pain*, 20(1):1–10, 2019.

[7] Yuan Yang, Jun Yao, Julius Dewald, Frans CT Van der Helm, and Alfred C Schouten. Quantifying the nonlinear interaction in the nervous system based on phase-locked amplitude relationship. *IEEE Transactions on Biomedical Engineering*, page in press.

[8] Fei He, Ptolemaios G Sarrigiannis, Stephen A Billings, Hualiang Wei, Jeremy Rowe, Charles Romanowski, Nigel Hoggard, Marios Hadjivassilliou, Dasappaiah Ganesh Rao, Richard Grünewald, et al. Nonlinear interactions in the thalamocortical loop in essential tremor: a model-based frequency domain analysis. *Neuroscience*, 324:377–389, 2016.

[9] Karl J Friston. Book review: brain function, nonlinear coupling, and neuronal transients. *The Neuroscientist*, 7(5):406–418, 2001.

[10] Yuan Yang, Julius PA Dewald, Frans CT van der Helm, and Alfred C Schouten. Unveiling neural coupling within the sensorimotor system: directionality and nonlinearity. *European journal of neuroscience*, 48(7):2407–2415, 2018.

[11] Juergen Fell and Nikolai Axmacher. The role of phase synchronization in memory processes. *Nature reviews neuroscience*, 12(2):105–118, 2011.

[12] Andrew S French and MJ Korenberg. A nonlinear cascade model for action potential encoding in an insect sensory neuron. *Biophysical journal*, 55(4):655–661, 1989.

[13] Ernesto Pereda, Rodrigo Quian Quiroga, and Joydeep Bhattacharya. Nonlinear multivariate analysis of neurophysiological signals. *Progress in neurobiology*, 77(1-2):1–37, 2005.

[14] Karl J Friston. Functional and effective connectivity: a review. *Brain connectivity*, 1(1):13–36, 2011.

[15] Alexander G Dimitrov, Aurel A Lazar, and Jonathan D Victor. Information theory in neuroscience. *Journal of computational neuroscience*, 30(1):1–5, 2011.

[16] Lennart Ljung. System identification. *Wiley encyclopedia of electrical and electronics engineering*, pages 1–19, 1999.

[17] Rik Pintelon and Johan Schoukens. *System identification: a frequency domain approach*. John Wiley & Sons, 2012.

[18] Eberhard E Fetz. Temporal coding in neural populations? *Science*, 278(5345):1901–1902, 1997.

[19] JORN Hounsgaard, HANS Hultborn, B Jespersen, and O Kiehn. Bistability of alpha-motoneurones in the decerebrate cat and in the acute spinal cat after intravenous 5-hydroxytryptophan. *The Journal of physiology*, 405(1):345–367, 1988.

[20] Jaeseung Jeong, Wei-Xing Shi, Ralph Hoffman, Jihoon Oh, John C Gore, Benjamin S Bunney, and Bradley S Peterson. Bursting as a source of non-linear determinism in the firing patterns of nigral dopamine neurons. *European Journal of Neuroscience*, 36(9):3214–3223, 2012.

[21] Yuan Yang, Teodoro Solis-Escalante, Mark van de Ruit, Frans CT van der Helm, and Alfred C Schouten. Nonlinear coupling between cortical oscillations and muscle activity during isotonic wrist flexion. *Frontiers in computational neuroscience*, 10:126, 2016.

[22] Nirvik Sinha, Julius Dewald, Charles J Heckman, and Yuan Yang. Cross-frequency coupling in descending motor pathways: Theory and simulation. *Frontiers in Systems Neuroscience*, 13:86, 2020.

[23] Sheng Li, Yen-Ting Chen, Gerard E Francisco, Ping Zhou, and William Zev Rymer. A unifying pathophysiological account for post-stroke spasticity and disordered motor control. *Frontiers in neurology*, 10, 2019.

[24] Y Yang, N Sinha, R Tian, N Gurari, JM Drogos, and JPA Dewald. Quantifying altered neural connectivity of the stretch reflex in chronic hemiparetic stroke. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(6):1436 – 1441, 2020.

[25] Pascal Fries. Rhythms for cognition: communication through coherence. *Neuron*, 88(1):220–235, 2015.

[26] Ole Jensen and Laura L Colgin. Cross-frequency coupling between neuronal oscillations. *Trends in cognitive sciences*, 11(7):267–269, 2007.

[27] Ryan T Canolty and Robert T Knight. The functional role of cross-frequency coupling. *Trends in cognitive sciences*, 14(11):506–515, 2010.

[28] Alexandre Hyafil, Anne-Lise Giraud, Lorenzo Fontolan, and Boris Gutkin. Neural cross-frequency coupling: connecting architectures, mechanisms, and functions. *Trends in neurosciences*, 38(11):725–740, 2015.

[29] Francisco Varela, Jean-Philippe Lachaux, Eugenio Rodriguez, and Jacques Martinerie. The brainweb: phase synchronization and large-scale integration. *Nature reviews neuroscience*, 2(4):229–239, 2001.

[30] André M Bastos and Jan-Mathijs Schoffelen. A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. *Frontiers in systems neuroscience*, 9:175, 2016.

[31] Luiz A Baccalá and Koichi Sameshima. Partial directed coherence: a new concept in neural structure determination. *Biological cybernetics*, 84(6):463–474, 2001.

[32] D Chicharro. On the spectral formulation of granger causality. *Biological cybernetics*, 105(5-6):331–347, 2011.

[33] Boris Gourévitch, Régine Le Bouquin-Jeannès, and Gérard Faucon. Linear and nonlinear causality between signals: methods, examples and neurophysiological applications. *Biological cybernetics*, 95(4):349–369, 2006.

[34] Vangelis Sakkalis. Review of advanced techniques for the estimation of brain connectivity measured with eeg/meg. *Computers in biology and medicine*, 41(12):1110–1117, 2011.

[35] Katarzyna J Blinowska. Review of the methods of determination of directed connectivity from multichannel data. *Medical & biological engineering & computing*, 49(5):521–529, 2011.

[36] Martijn P Vlaar, Teodoro Solis-Escalante, Alistair N Vardy, Frans CT Van Der Helm, and Alfred C Schouten. Quantifying nonlinear contributions to cortical responses evoked by continuous wrist manipulation. *IEEE transactions on neural systems and rehabilitation engineering*, 25(5):481–491, 2016.

[37] Karl J Friston. Functional and effective connectivity in neuroimaging: a synthesis. *Human brain mapping*, 2(1-2):56–78, 1994.

[38] Dong Song, Rosa HM Chan, Vasilis Z Marmarelis, Robert E Hampson, Sam A Deadwyler, and Theodore W Berger. Nonlinear dynamic modeling of spike train transformations for hippocampal-cortical prostheses. *IEEE Transactions on Biomedical Engineering*, 54(6):1053–1066, 2007.

[39] Martin Pienkowski and Jos J Eggermont. Nonlinear cross-frequency interactions in primary auditory cortex spectrotemporal receptive fields: a wiener–volterra analysis. *Journal of computational neuroscience*, 28(2):285–303, 2010.

[40] Aurel A Lazar and Yevgeniy B Slutskiy. Spiking neural circuits with dendritic stimulus processors. *Journal of computational neuroscience*, 38(1):1–24, 2015.

[41] Dong Song, Brian S Robinson, Robert E Hampson, Vasilis Z Marmarelis, Sam A Deadwyler, and Theodore W Berger. Sparse large-scale nonlinear dynamical modeling of human hippocampus for memory prostheses. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 26(2):272–280, 2016.

[42] Stephen A Billings. *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains*. John Wiley & Sons, 2013.

[43] Karl J Friston, Lee Harrison, and Will Penny. Dynamic causal modelling. *Neuroimage*, 19(4):1273–1302, 2003.

[44] IJ Leontaritis and Stephen A Billings. Input-output parametric models for non-linear systems part i: deterministic non-linear systems. *International journal of control*, 41(2):303–328, 1985.

[45] Sheng Chen, Stephen A Billings, and Wan Luo. Orthogonal least squares methods and their application to non-linear system identification. *International Journal of control*, 50(5):1873–1896, 1989.

[46] SA Billings and WSF Voon. Structure detection and model validity tests in the identification of nonlinear systems. In *IEE Proceedings D-Control Theory and Applications*, volume 130, pages 193–199. IET, 1983.

[47] Fei He, Stephen A Billings, Hua-Liang Wei, and Ptolemaios G Sarrigiannis. A nonlinear causality measure in

the frequency domain: Nonlinear partial directed coherence with applications to eeg. *Journal of neuroscience methods*, 225:71–80, 2014.

[48] Michael Breakspear. Dynamic models of large-scale brain activity. *Nature neuroscience*, 20(3):340, 2017.

[49] Dorian Florescu and Daniel Coca. Identification of linear and nonlinear sensory processing circuits from spiking neuron data. *Neural computation*, 30(3):670–707, 2018.

[50] Runfeng Tian, Yuan Yang, Frans CT van der Helm, and Julius Dewald. A novel approach for modeling neural responses to joint perturbations using the narmax method and a hierarchical neural network. *Frontiers in computational neuroscience*, 12:96, 2018.

[51] Yuanlin Gu, Yuan Yang, Julius Dewald, Frans CT Van der Helm, Alfred C Schouten, and Liang-Hua Wei. Nonlinear modeling of cortical responses to mechanical wrist perturbations using the narmax method. *IEEE Transactions on Biomedical Engineering*, page in press.

[52] Klaas E Stephan, Lee M Harrison, Stefan J Kiebel, Olivier David, Will D Penny, and Karl J Friston. Dynamic causal models of neural system dynamics: current state and future extensions. *Journal of biosciences*, 32(1):129–144, 2007.

[53] Karl J Friston, Katrin H Preller, Chris Mathys, Hayriye Cagnan, Jakob Heinzle, Adeel Razi, and Peter Zeidman. Dynamic causal modelling revisited. *Neuroimage*, 199:730–744, 2019.

[54] Stefan J Kiebel, Marta I Garrido, Rosalyn J Moran, and Karl J Friston. Dynamic causal modelling for eeg and meg. *Cognitive neurodynamics*, 2(2):121, 2008.

[55] Olivier David, Stefan J Kiebel, Lee M Harrison, Jérémie Mattout, James M Kilner, and Karl J Friston. Dynamic causal modeling of evoked responses in eeg and meg. *NeuroImage*, 30(4):1255–1272, 2006.

[56] Klaas Enno Stephan, Will D Penny, Rosalyn J Moran, Hanneke EM den Ouden, Jean Daunizeau, and Karl J Friston. Ten simple rules for dynamic causal modeling. *Neuroimage*, 49(4):3099–3109, 2010.

[57] Zi-Qiang Lang and SA Billings. Output frequency characteristics of nonlinear systems. *International Journal of Control*, 64(6):1049–1067, 1996.

[58] SA Billings and KM Tsang. Spectral analysis for nonlinear systems, part ii: Interpretation of non-linear frequency response functions. *Mechanical Systems and Signal Processing*, 3(4):341–359, 1989.

[59] Alexander Ya Kaplan, Andrew A Fingelkurts, Alexander A Fingelkurts, Sergei V Borisov, and Boris S Darkhovsky. Nonstationary nature of the brain activity as revealed by eeg/meg: methodological, practical and conceptual challenges. *Signal processing*, 85(11):2190–2212, 2005.

[60] Klaus Lehnertz, Christian Geier, Thorsten Rings, and Kirsten Stahn. Capturing time-varying brain dynamics. *EPJ Nonlinear Biomedical Physics*, 5:2, 2017.

[61] Lennart Ljung and Svante Gunnarsson. Adaptation and tracking in system identification—a survey. *Automatica*, 26(1):7–21, 1990.

[62] Rui Zou, Hengliang Wang, and Ki H Chon. A robust time-varying identification algorithm using basis functions. *Annals of biomedical engineering*, 31(7):840–853, 2003.

[63] Michail K Tsatsanis and Georgios B Giannakis. Time-varying system identification and model validation using wavelets. *IEEE Transactions on Signal Processing*, 41(12):3512–3523, 1993.

[64] Maciej Niedzwiecki and Tomasz Klaput. Fast recursive basis function estimators for identification of time-varying processes. *IEEE Transactions on Signal Processing*, 50(8):1925–1934, 2002.

[65] Y Zheng, Z Lin, and David Ban Hock Tay. Time-varying parametric system multiresolution identification by wavelets. *International Journal of Systems Science*, 32(6):775–793, 2001.

[66] Maria Iatrou, Theodore W Berger, and Vasillis Z Marmarelis. Modeling of nonlinear nonstationary dynamic systems with a novel class of artificial neural networks. *IEEE transactions on neural networks*, 10(2):327–339, 1999.

[67] Yuru Zhong, Kung-Ming Jan, Ki H Ju, and Ki H Chon. Representation of time-varying nonlinear systems with time-varying principal dynamic modes. *IEEE transactions on biomedical engineering*, 54(11):1983–1992, 2007.

[68] Yang Li, Hua-Liang Wei, Stephen A Billings, and Ptolemaios G Sarrigiannis. Identification of nonlinear time-varying systems using an online sliding-window and common model structure selection (cmss) approach with applications to eeg. *International Journal of Systems Science*, 47(11):2671–2681, 2016.

[69] Stephen A Billings* and Hua-Liang Wei. The wavelet-narmax representation: A hybrid model structure combining polynomial models with multiresolution wavelet decompositions. *International Journal of Systems Science*, 36(3):137–152, 2005.

[70] Fei He, Hua-Liang Wei, and Stephen A Billings. Identification and frequency domain analysis of nonstationary and nonlinear systems using time-varying narmax models. *International Journal of Systems Science*, 46(11):2087–2100, 2015.

[71] Yang Li, Wei-Gang Cui, Yu-Zhu Guo, Tingwen Huang, Xiao-Feng Yang, and Hua-Liang Wei. Time-varying system identification using an ultra-orthogonal forward regression and multiwavelet basis functions with applications to eeg. *IEEE transactions on neural networks and learning systems*, 29(7):2960–2972, 2018.

[72] Fei He, Stephen A Billings, Hua-Liang Wei, Ptolemaios G Sarrigiannis, and Yifan Zhao. Spectral analysis for nonstationary and nonlinear systems: A discrete-time-model-based approach. *IEEE Transactions on Biomedical Engineering*, 60(8):2233–2241, 2013.

[73] Fei He, Hua-Liang Wei, Stephen A Billings, and Ptolemaios G Sarrigiannis. A nonlinear generalization of spectral granger causality. *IEEE Transactions on Biomedical Engineering*, 61(6):1693–1701, 2014.

[74] Chrysostomos L Nikias and Jerry M Mendel. Signal processing with higher-order spectra. *IEEE Signal processing magazine*, 10(3):10–37, 1993.

[75] Jeffrey C Sigl and Nassib G Chamoun. An introduction to bispectral analysis for the electroencephalogram. *Journal of clinical monitoring*, 10(6):392–404, 1994.

[76] Federico Chella, Laura Marzetti, Vittorio Pizzella, Filippo Zappasodi, and Guido Nolte. Third order spectral analysis robust to mixing artifacts for mapping cross-frequency interactions in eeg/meg. *Neuroimage*, 91:146–161, 2014.

[77] Federico Chella, Vittorio Pizzella, Filippo Zappasodi, Guido Nolte, and Laura Marzetti. Bispectral pair-

wise interacting source analysis for identifying systems of cross-frequency interacting brain sources from electroencephalographic or magnetoencephalographic signals. *Physical Review E*, 93(5):052420, 2016.

[78] Joseph R Isler, Philip G Grieve, D Czernochowski, Raymond I Stark, and David Friedman. Cross-frequency phase coupling of brain rhythms during the orienting response. *Brain research*, 1232:163–172, 2008.

[79] KC Chua, Vinod Chandran, U Rajendra Acharya, and CM Lim. Analysis of epileptic eeg signals using higher order spectra. *Journal of medical engineering & technology*, 33(1):42–50, 2009.

[80] SARA Marceglia, Guglielmo Foffani, AM Bianchi, Giuseppe Baselli, F Tamma, M Egidi, and Alberto Priori. Dopamine-dependent non-linear correlation between subthalamic rhythms in parkinson's disease. *The Journal of physiology*, 571(3):579–591, 2006.

[81] Zongbao Wang, Yongzhi Huang, Shouyan Wang, Alex Green, Tipu Aziz, and John Stein. Tremor dependant nonlinear interaction in deep brain local field potentials of parkinson's disease. In *2014 7th International Conference on Biomedical Engineering and Informatics*, pages 399–404. IEEE, 2014.

[82] Matthias Wacker and Herbert Witte. On the stability of the n: m phase synchronization index. *IEEE Transactions on Biomedical Engineering*, 58(2):332–338, 2010.

[83] Robson Scheffer-Teixeira and Adriano BL Tort. On cross-frequency phase-phase coupling between theta and gamma oscillations in the hippocampus. *Elife*, 5:e20515, 2016.

[84] Felix Darvas, Kai J Miller, Rajesh PN Rao, and Jeffrey G Ojemann. Nonlinear phase–phase cross-frequency coupling mediates communication between distant sites in human neocortex. *Journal of Neuroscience*, 29(2):426–435, 2009.

[85] Steve Elgar and Robert T Guza. Statistics of bicoherence. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(10):1667–1668, 1988.

[86] JL Shils, M Litt, BE Skolnick, and MM Stecker. Bispectral analysis of visual interactions in humans. *Electroencephalography and clinical Neurophysiology*, 98(2):113–125, 1996.

[87] Vinod Chandran. Time-varying bispectral analysis of auditory evoked multi-channel scalp eeg. In *2012 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA)*, pages 205–212. IEEE, 2012.

[88] Christopher Gundlach and Matthias M Müller. Perception of illusory contours forms intermodulation responses of steady state visual evoked potentials as a neural signature of spatial integration. *Biological psychology*, 94(1):55–60, 2013.

[89] Yuan Yang, Teodoro Solis-Escalante, Jun Yao, Andreas Daffertshofer, Alfred C Schouten, and Frans CT Van Der Helm. A general approach for quantifying nonlinear connectivity in the nervous system based on phase coupling. *International journal of neural systems*, 26(01):1550031, 2016.

[90] Felix Darvas, Jeffrey G Ojemann, and Larry B Sorensen. Bi-phase locking—a tool for probing non-linear interaction in the human brain. *Neuroimage*, 46(1):123–132, 2009.

[91] Runfeng Tian, Julius Dewald, and Yuan Yang. Assessing neural connectivity and associated time delays of muscle responses to continuous position perturbations. *Annals of Biomedical Engineering*, page in press, 2020.

[92] Cees Diks and Valentyn Panchenko. A new statistic and practical guidelines for nonparametric granger causality testing. *Journal of Economic Dynamics and Control*, 30(9-10):1647–1669, 2006.

[93] Craig Hiemstra and Jonathan D Jones. Testing for linear and nonlinear granger causality in the stock price-volume relation. *The Journal of Finance*, 49(5):1639–1664, 1994.

[94] Nicola Ancona, Daniele Marinazzo, and Sebastiano Stramaglia. Radial basis function approach to nonlinear granger causality of time series. *Physical Review E*, 70(5):056221, 2004.

[95] Daniele Marinazzo, Mario Pellicoro, and Sebastiano Stramaglia. Kernel-granger causality and the analysis of dynamical networks. *Physical review E*, 77(5):056215, 2008.

[96] Yonghong Chen, Govindan Rangarajan, Jianfeng Feng, and Mingzhou Ding. Analyzing multiple nonlinear time series with extended granger causality. *Physics letters A*, 324(1):26–35, 2004.

[97] Ying Liu and Selin Aviyente. Quantification of effective connectivity in the brain using a measure of directed information. *Computational and mathematical methods in medicine*, 2012, 2012.

[98] Md Hedayetul Islam Shovon, Nanda Nandagopal, Ramasamy Vijayalakshmi, Jia Tina Du, and Bernadine Cocks. Directed connectivity analysis of functional brain networks during cognitive activity using transfer entropy. *Neural Processing Letters*, 45(3):807–824, 2017.

[99] Dennis Joe Harmah, Cunbo Li, Fali Li, Yuanyuan Liao, Jiuju Wang, Walid MA Ayedh, Joyce Chelangat Bore, Dezhong Yao, Wentian Dong, and Peng Xu. Measuring the non-linear directed information flow in schizophrenia by multivariate transfer entropy. *Frontiers in Computational Neuroscience*, 13, 2019.

[100] Luca Faes, Giandomenico Nollo, and Ki H Chon. Assessment of granger causality by nonlinear model identification: application to short-term cardiovascular variability. *Annals of biomedical engineering*, 36(3):381–395, 2008.

[101] Yifan Zhao, Steve A Billings, Hualiang Wei, Fei He, and Ptolemaios G Sarrigiannis. A new narx-based granger linear and nonlinear casual influence detection method with applications to eeg data. *Journal of neuroscience methods*, 212(1):79–86, 2013.

[102] Yang Li, Hua-Liang Wei, Steve A Billings, and Xiao-Feng Liao. Time-varying linear and nonlinear parametric model for granger causality analysis. *Physical Review E*, 85(4):041906, 2012.

[103] Qizhong Zhang, Yuejing Hu, Thomas Potter, Rihui Li, Michael Quach, and Yingchun Zhang. Establishing functional brain networks using a nonlinear partial directed coherence method to predict epileptic seizures. *Journal of Neuroscience Methods*, 329:108447, 2020.

[104] Shozo Tobimatsu, You Min Zhang, and Motohiro Kato. Steady-state vibration somatosensory evoked potentials: physiological characteristics and tuning function. *Clinical neurophysiology*, 110(11):1953–1958, 1999.

[105] Sherwin E Hua and Frederick A Lenz. Posture-related oscillations in human cerebellar thalamus in essential tremor are enabled by voluntary motor circuits. *Journal of neurophysiology*, 93(1):117–127, 2005.

[106] JF Marsden, P Ashby, P Limousin-Dowsey, JC Rothwell, and P Brown. Coherence between cerebellar thalamus, cortex and muscle in man: cerebellar thalamus interactions. *Brain*, 123(7):1459–1470, 2000.

[107] Günther Deuschl, Jan Raethjen, Helge Hellriegel, and Rodger Elble. Treatment of patients with essential tremor. *The Lancet Neurology*, 10(2):148–161, 2011.

[108] B Hellwig, S Häußler, B Schelter, M Lauk, B Guschlbauer, J Timmer, and CH Lücking. Tremor-correlated cortical activity in essential tremor. *The Lancet*, 357(9255):519–523, 2001.

[109] Jan Raethjen and Günther Deuschl. The oscillating central network of essential tremor. *Clinical neurophysiology*, 123(1):61–64, 2012.

[110] Arthur WG Buijink, AM Madelein van der Stouwe, Marja Broersma, Sarvi Sharifi, Paul FC Groot, Johannes D Speelman, Natasha M Maurits, and Anne-Fleur van Rootselaar. Motor network disruption in essential tremor: a functional and effective connectivity study. *Brain*, 138(10):2934–2947, 2015.

[111] Jinsook Roh, William Z Rymer, Eric J Perreault, Seng Bum Yoo, and Randall F Beer. Alterations in upper limb muscle synergy structure in chronic stroke survivors. *Journal of neurophysiology*, 109(3):768–781, 2013.

[112] Jacob G McPherson, Albert Chen, Michael D Ellis, Jun Yao, CJ Heckman, and Julius PA Dewald. Progressive recruitment of contralesional cortico-reticulospinal pathways drives motor impairment post stroke. *The Journal of physiology*, 596(7):1211–1225, 2018.

[113] Leonidas D Iasemidis and J Chris Sackellares. Review: Chaos theory and epilepsy. *The Neuroscientist*, 2(2):118–126, 1996.

[114] Klaus Lehnertz. Epilepsy and nonlinear dynamics. *Journal of biological physics*, 34(3-4):253–266, 2008.

[115] Paul R Carney, Stephen Myers, and James D Geyer. Seizure prediction: methods. *Epilepsy & behavior*, 22:S94–S101, 2011.

[116] U Rajendra Acharya, S Vinitha Sree, G Swapna, Roshan Joy Martis, and Jasjit S Suri. Automated eeg analysis of epilepsy: a review. *Knowledge-Based Systems*, 45:147–165, 2013.

[117] Christian E Elger and Christian Hoppe. Diagnostic challenges in epilepsy: seizure under-reporting and seizure detection. *The Lancet Neurology*, 17(3):279–288, 2018.

[118] Angeliki Papana, Dimitris Kugiumtzis, and Pål G Larsson. Detection of direct causal effects and application to epileptic electroencephalogram analysis. *International Journal of Bifurcation and Chaos*, 22(09):1250222, 2012.

[119] Marina V Sysoeva, Evgenia Sitnikova, Ilya V Sysoev, Boris P Bezruchko, and Gilles van Luijtelaar. Application of adaptive nonlinear granger causality: Disclosing network changes before and after absence seizure onset in a genetic rat model. *Journal of neuroscience methods*, 226:33–41, 2014.

[120] Alex Tank, Ian Covert, Nicholas Foti, Ali Shojaie, and Emily Fox. Neural granger causality for nonlinear time series. *arXiv preprint arXiv:1802.05842*, 2018.

[121] Aneesh Sreevallabh Chivukula, Jun Li, and Wei Liu. Discovering granger-causal features from deep learning networks. In *Australasian Joint Conference on Artificial Intelligence*, pages 692–705. Springer, 2018.

[122] Wei Peng. Dli: A deep learning-based granger causality inference. *Complexity*, 2020, 2020.

[123] Cornelis J Stam. Nonlinear dynamical analysis of eeg and meg: review of an emerging field. *Clinical neurophysiology*, 116(10):2266–2301, 2005.

[124] Daniel J Blackburn, Yifan Zhao, Matteo De Marco, Simon M Bell, Fei He, Hua-Liang Wei, Sarah Lawrence, Zoe C Unwin, Michelle Blyth, Jenna Angel, et al. A pilot study investigating a novel non-linear measure of eyes open versus eyes closed eeg synchronization in people with alzheimer's disease and healthy controls. *Brain Sciences*, 8(7):134, 2018.

[125] Shenal RAS Gunawardena, Fei He, Ptolemaios Sarrigiannis, and Daniel J Blackburn. Nonlinear classification of eeg recordings from patients with alzheimer's disease using gaussian process latent variable model. *medRxiv*, 2020.