Model-based estimation of intra-cortical connectivity using electrophysiological data

P. Aram [a,b,*,1], D.R. Freestone [c,d,e,1], M.J. Cook [d], V. Kadirkamanathan [a], D.B. Grayden [c,d,f]

[a] Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, UK
[b] Insigneo Institute for in silico Medicine, University of Sheffield, Sheffield, UK
[c] NeuroEngineering Laboratory, Department of Electrical and Electronic Engineering, The University of Melbourne, Melbourne, VIC, Australia
[d] Department of Medicine, St. Vincent's Hospital Melbourne, The University of Melbourne, Melbourne, VIC, Australia
[e] Department of Statistics, Columbia University, New York, NY 10027, USA
[f] Centre for Neural Engineering, The University of Melbourne, Melbourne, VIC, Australia

ARTICLE INFO

ABSTRACT

This paper provides a new method for model-based estimation of intra-cortical connectivity from electrophysiological measurements. A novel closed-form solution for the connectivity function of the Amari neural field equations is derived as a function of electrophysiological observations. The resultant intra-cortical connectivity estimate is driven from experimental data, but constrained by the mesoscopic neurodynamics that are encoded in the computational model. A demonstration is provided to show how the method can be used to image physiological mechanisms that govern cortical dynamics, which are normally hidden in clinical data from epilepsy patients. Accurate estimation performance is demonstrated using synthetic data. Following the computational testing, results from patient data are obtained that indicate a dominant increase in surround inhibition prior to seizure onset that subsides in the cases when the seizures spread.

Introduction

The human brain is arguably nature's most complex system. The development and validation of a theory that underpins its function is one of the greatest challenges faced by scientists today. This great challenge is being addressed by both the experimental and theoretical neuroscience communities. The experimental neuroscience community is generating an ever increasing number of facts, describing the interaction between manipulations and observations. The theoretical neuroscience community is developing mathematical models that can explain generators of data and make non-trivial predictions about system behavior (to be validated by experiments). To date, neither approach has been successful in developing an understanding of high-level brain function. Continuously accumulating more facts has not brought us closer to an understanding of what appears to be emergent phenomena in the brain. Understanding such phenomena requires the development and acceptance of theory. However, theoretical developments have been limited by the inability to accurately measure model parameters and account for inter-subject variability. This has led to mathematical models that are either over-parameterized or overly-simplified. Over-parameterized models often provide ambiguous explanations of data leading to misleading theories. Overly-simple models can be useful in certain applications, but can often neglect important aspects of the underlying biology.

An opportunity to overcome the challenges mentioned above has arisen with the advent of data-driven neural modeling. Data-driven modeling is a process of a creating a subject-specific mathematical model of a particular subject or experimental preparation. The model is constrained by known relationships and general principles that are described by mathematical functions, where the parameters of the functions are considered unknown. For example, an excitatory post-synaptic potential will have a fast rise time (from synaptic dynamics) and a slower decay time (from membrane dynamics) that is described by

$$h(t) = \alpha \left(exp\left(-\frac{t}{t_s}\right) - exp\left(-\frac{t}{t_m}\right) \right), \tag{1}$$

where α is a synaptic gain parameter and $t_s < t_m$ are the synaptic and membrane time constants, respectively. The form of the synaptic response function (or kernel due to the convolution in time) is well accepted. However, the parameters are known to vary across subjects, brain regions, and neural population types and thus need to either measured or inferred from data to create accurate models.

* Corresponding author at: Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, UK.
E-mail addresses: p.aram@sheffield.ac.uk (P. Aram), deanrf@unimelb.edu.au (D.R. Freestone), markcook@unimelb.edu.au (M.J. Cook), visakan@sheffield.ac.uk (V. Kadirkamanathan), grayden@unimelb.edu.au (D.B. Grayden).
[1] Authors contributed equally to this work.

Similarly, the probability that a neuron at position **r** connects with a neighbor at position **r**′ decreases as the distance, **r** − **r**′, increases. This motivates the shape of connectivity functions (or kernels due to the convolution in space) that are used to describe neural fields, such as Amari's model (Amari, 1977) that has the form

$$w(\mathbf{r} - \mathbf{r}') = \sum_n \theta_n \, exp\left(-\frac{(\mathbf{r} - \mathbf{r}')^T(\mathbf{r} - \mathbf{r}')}{\sigma_n}\right), \tag{2}$$

where θ_n are the parameters that describe the effective connectivity strength of excitatory ($n = e$) and inhibitory ($n = i$) connections, and σ_n specifies the axonal–dendritic range. The form of the connectivity function is well accepted in the literature and gives rise to important phenomena, such as tuning of cortical regions to receptive fields. However, as for the case with the post-synaptic response function in Eq. (1), the parameters are subject-specific.

There are several data-driven frameworks recently presented in the literature (Friston et al., 2003; Schiff and Sauer, 2008; Ullah and Schiff, 2010; Sedigh-Sarvestani et al., 2012; Pinotsis et al., 2013; Gorzelic et al., 2013; Turner et al., 2013; Aram et al., 2012; Freestone et al., 2011, 2013, 2014). These frameworks utilize system identification techniques to solve the problem of inferring parameters from data. They have demonstrated great potential in furthering our understanding of the function and structure of neural circuits. Moreover, data-driven models provide new opportunities in the field of neural engineering to incorporate control and systems theory to optimize therapeutic bionic devices (Schiff, 2011).

Perhaps the most limiting factor in the widespread adoption of data-driven modeling frameworks is the high level of complexity of the estimation algorithms. It is often the case with complicated algorithms that the results are clouded by a lack of understanding of the methods involved. Furthermore, the high level of complexity has led to the inappropriate application of data-driven methods to certain problems. Therefore, the development of methods that do not rely on complicated, iterative algorithms represents a significant contribution to neuroscience.

This paper provides a method for data-driven neural field modeling that does not rely on complicated, computationally intense estimation algorithms. The output of the method is an estimate of an intra-cortical connectivity function that can be computed in closed-form from local field potential or other high-resolution electrophysiological measurements. The estimated connectivity function is based on the assumption that the mean field dynamics of the cortex is governed by Amari-style neural field equations (Amari, 1977), where the parameters are not known. The dynamics of this model are governed by the connectivity function, which physically describes the statistics of the axonal–dendritic projections. Computationally, the shape of the connectivity function strongly dictates the type of dynamics that the cortical field can exhibit.

Results from data-driven mesoscopic neural modeling frameworks must be interpreted with care. A common misconception is that the variables that are estimated have a direct one-to-one correspondence to the actual brain. This is not the case. The resultant models are far less complex than the actual brain. Accordingly, parameter estimates must be interpreted as being constrained by the models. This is not to say that the models are not related to actual neural dynamics or that valuable insights cannot be gained. In actual fact, mesoscopic models are leading to new hypotheses about many types of phenomena. Furthermore, the constraints that the models place on the data facilitate the estimation of variables that are normally hidden in experimental observations.

The rest of this paper is set out as follows. In the Methods section, the stochastic Amari neural field model is briefly reviewed. Then necessary formulations for the intra-cortical connectivity estimator are provided. This is followed by the data collection approach and the pre-processing steps. The Results section provides examples using synthetic data that demonstrate the estimation performance with known parameters. Following this, results using recorded intra-cranial electroencephalogram (iEEG) data over normal, seizure, and post-seizure periods are presented. Finally, in the Discussion section, the implications and limitations of this method as well as possible future extensions are discussed. All frequently used symbols in the following sections are given in Table 1.

Methods

Neural field model

This manuscript presents a new method for inferring the underlying connectivity structure of cortex with the assumption that the cortical dynamics of interest are governed by physical laws described by the neural field model of Amari (1977). The single layer Amari neural field model is

$$v(t,\mathbf{r}) = \int_{-\infty}^{t} h(t - t')\left(\int_\Omega w(\mathbf{r} - \mathbf{r}')f(v(t',\mathbf{r}'))\,d\mathbf{r}' + p(t',\mathbf{r})\right)dt'. \tag{3}$$

The spatial dynamics are governed by the connectivity function, $w(\mathbf{r})$, that collects all the presynaptic firing rates that drive the field of postsynaptic potentials, $v(t,\mathbf{r})$, and $\mathbf{r} \in \Omega \subset \mathbb{R}^n$ are spatial locations in n-dimensional physical space, $n \in \{1, 2, 3\}$. The temporal dynamics are governed by the post-synaptic response function, $h(t)$, acting on action potentials and the external inputs arriving from other neural populations. The term $p(t,\mathbf{r})$ denotes external inputs. The relationship between the presynaptic mean membrane potential, and the presynaptic mean firing rate is typically described by a sigmoid function in generative neural population models. The sigmoidal relationship is

$$f(v(t,\mathbf{r})) = \frac{f_{max}}{1 + exp(\varsigma(v_0 - v(t,\mathbf{r})))}, \tag{4}$$

where f_{max} is the maximum firing rate, v_0 describes the mean firing threshold relative to the resting membrane potential and the parameter ς defines the steepness of the sigmoid at v_0 (also specifies variability of firing thresholds).

Table 1
Notation. The symbols, description of the quantity, and the SI units where relevant.

Symbol	Quantity	Units
Domain		
Ω	Spatial domain	n.a.
$\mathbb{Z}^+$	Non-negative integers	n.a.
$\mathbb{R}^n$	n-Dimensional real numbers	n.a.
$\mathbf{r}$	Spatial location	[mm, mm]
t	Time	s
Model		
$\mathbf{y}_t(\mathbf{r}_{ij})$	Electrophysiological measurement	mV
$v_t(\mathbf{r})$	Mean membrane potential field	mV
$f(v(\mathbf{r}))$	Activation function	spike s^{-1}
$\hat{f}(v(\mathbf{r}))$	Linearized activation function	spike s^{-1}
$e_t(\mathbf{r})$	Field disturbance, with covariance function $\gamma(\mathbf{r})$	mV
$\varepsilon_t(\mathbf{r}_{ij})$	Observation noise, with covariance matrix Σ_ε	mV
$m(\mathbf{r}_{ij})$	Observation function, where $i = 1, ..., I$ and $j = 1, ... J$	n.a.
$w(\mathbf{r})$	Connectivity function	mV spike^{-1}
$\psi_i(\mathbf{r})$	Connectivity basis functions	n.a.
θ_i	Weights of the connectivity basis functions	mV spike^{-1}
μ_i	Centers of the connectivity basis functions	mm
σ_i	Widths of the connectivity basis functions	mm
Estimation		
τ	Spatial shift	mm
ν	Spatial frequency	cycles/mm
$\mathbf{y}_t^d$	Differential re-referenced observations along j-direction	mV
$R(\tau)$	Spatial correlation	mV^2
$S(\nu)$	Power spectral density	mV^2 mm/cycles

Simplifying assumptions

In order to relate the model to data, the neural field in Eq. (3) must be written as an integro-differential equation (IDE). To achieve this, the convolution of the post-synaptic response function with the firing rate is rewritten as a differential operator acting on $v(t, \mathbf{r})$,

$$\mathrm{D}v(t, \mathbf{r}) = \int_\Omega w(\mathbf{r} - \mathbf{r}') f(v(t, \mathbf{r}')) \, d\mathbf{r}' + p(t, \mathbf{r}). \tag{5}$$

Under the assumption that synaptic transmission is instantaneous (infinitely fast rise time on the post-synaptic response function), the differential operator is

$$\mathrm{D} = \frac{d}{dt} + \frac{1}{t_m}, \tag{6}$$

where t_m is the membrane time constant. Note that here we have set synaptic gain $\alpha = 1$, since it is linear with respect to the coupling strength and can be absorbed into the connectivity function. If we assume the input $p(t, \mathbf{r})$ is unknown, but modeled by a spatiotemporal Wiener process, then the system can be discretized using the Euler–Maruyama method (see Freestone et al., 2011 for details),

$$v_{t+1}(\mathbf{r}) = \xi v_t(\mathbf{r}) + T_s \int_\Omega w(\mathbf{r} - \mathbf{r}') f(v_t(\mathbf{r}')) d\mathbf{r}' + e_t(\mathbf{r}), \tag{7}$$

where

$$\xi = \left(1 - \frac{T_s}{t_m}\right), \tag{8}$$

$t \in \mathbb{Z}^+$ indexes discrete time, and T_s is the sampling time step. The disturbance term, $e_t(\mathbf{r}) = \sigma_d[p(t + T_s, \mathbf{r}) - p(t, \mathbf{r})]$, is the increment of a spatiotemporal Wiener process, independent and identically distributed (i. i. d.) with zero mean such that $e_t(\mathbf{r}) \sim \mathcal{GP}(\mathbf{0}, T_s \sigma_d^2 \delta(t - t')\gamma(\mathbf{r} - \mathbf{r}'))$. Here $\mathcal{GP}(\mathbf{0}, T_s \sigma_d^2 \delta(t - t')\gamma(\mathbf{r} - \mathbf{r}'))$ denotes a zero mean Gaussian process, spatially colored with the spatial covariance function, $\gamma(\mathbf{r} - \mathbf{r}')$, and temporally independent with the temporal covariance function, $T_s \sigma_d^2 \delta(t - t')$ (Rasmussen and Williams, 2005). The shape or width of the covariance function, $\gamma(\mathbf{r} - \mathbf{r}')$, dictates how correlated the inputs are at adjoining cortical areas.

The activation function is approximated by linearizing about the threshold using a first order truncated Taylor-series expansion as

$$f(v_t(\mathbf{r})) \approx \hat{f}(v_t(\mathbf{r})) = f_{\max}\left(\frac{1}{2} + \frac{1}{4}\varsigma(v_t(\mathbf{r}) - v_0)\right). \tag{9}$$

Given that this model is used for inversion purposes, provided that not all neurons are silent or firing at the maximal rate, this approximation will be reasonably accurate. Since maximum firing rate is simply multiplied with the connectivity strength parameter, we absorb the maximal firing rate into the connectivity function, $w(\cdot)$, in Eq. (7).

Observation system

The electrophysiological measurements are related to the neural field via the measurement function, which incorporates a differential montage, the size and position of the electrodes, and measurement noise. The electrodes are placed at discrete points within the neural field that are specified as $\mathbf{r}_{ij}$, where n indexes the electrodes and $n = 1, \cdots, n_y$. The measurement equation is

$$y_t(\mathbf{r}_{ij}) = \int_\Omega m(\mathbf{r}_{ij} - \mathbf{r}) v_t(\mathbf{r}) d\mathbf{r} + \mathbf{c}_t + \varepsilon_t(\mathbf{r}_{ij}), \tag{10}$$

where $m(\cdot)$ is a function that models the recording electrode pickup range, $\mathbf{c}_t$ is common-mode inference, and $\varepsilon_t(\mathbf{r}_{ij}) \sim \mathcal{N}(0, \Sigma_\varepsilon)$ denotes an additive noise component from a multivariate normal distribution

with mean zero and covariance matrix $\Sigma_\varepsilon = \sigma_\varepsilon^2 \mathbf{I}_{n_y}$, where $\mathbf{I}_{n_y}$ is the $n_y \times n_y$ identity matrix. The observation system models the tissue conductivity as being homogeneous and isotropic, where $m(\mathbf{r}_{ij} - \mathbf{r})$ is symmetric and the same at all spatial locations. To give a two-dimensional representation of the observations, the measurements at each time instant, t, are gathered in the matrix,

$$y_t(\mathbf{r}_{ij}) = \begin{bmatrix} y_t(\mathbf{r}_{11}) & y_t(\mathbf{r}_{12}) & \cdots & y_t(\mathbf{r}_{1J}) \\ y_t(\mathbf{r}_{21}) & y_t(\mathbf{r}_{22}) & \cdots & y_t(\mathbf{r}_{2J}) \\ \vdots & \vdots & \ddots & \vdots \\ y_t(\mathbf{r}_{I1}) & y_t(\mathbf{r}_{I2}) & \cdots & y_t(\mathbf{r}_{IJ}) \end{bmatrix}_{I \times J}. \tag{11}$$

Differential re-referencing along the *j*-direction is then defined as

$$y_t^d(\mathbf{r}_{ij}) = y_t^+(\mathbf{r}_{ij}) - y_t^-(\mathbf{r}_{ij}), \tag{12}$$

where

$$y_t^+(\mathbf{r}_{ij}) = \begin{bmatrix} y_t(\mathbf{r}_{11}) & y_t(\mathbf{r}_{12}) & \cdots & y_t(\mathbf{r}_{1J-1}) \\ y_t(\mathbf{r}_{21}) & y_t(\mathbf{r}_{22}) & \cdots & y_t(\mathbf{r}_{2J-1}) \\ \vdots & \vdots & \ddots & \vdots \\ y_t(\mathbf{r}_{I1}) & y_t(\mathbf{r}_{I2}) & \cdots & y_t(\mathbf{r}_{IJ-1}) \end{bmatrix}_{I \times (J-1)}, \tag{13}$$

and

$$y_t^-(\mathbf{r}_{ij}) = \begin{bmatrix} y_t(\mathbf{r}_{12}) & y_t(\mathbf{r}_{13}) & \cdots & y_t(\mathbf{r}_{1J}) \\ y_t(\mathbf{r}_{22}) & y_t(\mathbf{r}_{23}) & \cdots & y_t(\mathbf{r}_{2J}) \\ \vdots & \vdots & \ddots & \vdots \\ y_t(\mathbf{r}_{I2}) & y_t(\mathbf{r}_{I3}) & \cdots & y_t(\mathbf{r}_{IJ}) \end{bmatrix}_{I \times (J-1)}. \tag{14}$$

The re-referencing yields $n_d = I \times (J - 1)$ observation points.

Estimation method

Here, we consider a two-dimensional cortical sheet and present a closed-form estimate of the connectivity function from electrophysiological data. Importantly, the estimate accounts for a realistic differential montage of the data. This method of re-referencing provides a localized, noise-robust measurement of neural activities, where the common-mode interference and the effect of the reference electrode are removed.

The closed-form solution to the connectivity function is given by

$$w(\mathbf{r}) = \frac{4}{T_s \varsigma} \mathcal{F}^{-1}\left(\frac{\mathcal{S}_{y_t^d y_{t+1}^d}(\boldsymbol{\nu})}{\mathcal{S}_{y_t^d y_t^d}(\boldsymbol{\nu}) + \mathcal{S}_{\varepsilon_t^d \varepsilon_t^d}(\boldsymbol{\nu})} - \xi\right), \tag{15}$$

where $\boldsymbol{\nu}$ is the spatial frequency, $\mathcal{S}_{y_t^d y_{t+1}^d}(\boldsymbol{\nu})$ and $\mathcal{S}_{y_t^d y_t^d}(\boldsymbol{\nu})$ are time-averaged spatial power spectral densities between consecutive re-referenced observations at each time point. The term, $\mathcal{S}_{\varepsilon_t^d \varepsilon_t^d}(\boldsymbol{\nu})$, is the analytic spatial power spectral density of the re-referenced observation noise. The operator $\mathcal{F}^{-1}$ is the inverse Fourier transform. The derivation and detailed descriptions of each term in Eq. (15) are given in Appendix A. Note that the solution is independent of the disturbance covariance function, $\sigma_d^2 \gamma(\cdot)$, the firing threshold parameter, v_0, and the sensor model, $m(\cdot)$. However, the activation function slope parameter, ς, the observation noise variance, σ_ε^2, and the synaptic decay parameter, ξ, are required to estimate the connectivity function. The slope parameter, ς, can be absorbed into the connectivity function on the left hand side of Eq. (15), resulting in a scaled estimate of the connectivity function. This removes a parameter that cannot be recovered by our estimation algorithm. An error in the assumed value of ξ will result in a shift of the estimated connectivity function amplitude at $w(\mathbf{r} = 0)$ (since $\mathcal{F}^{-1}(\xi) = \xi\delta(\mathbf{r})$).

The Wiener–Khintchine theorem can be used to provide a bound on the estimate of the observation noise variance, σ_ε^2. The denominator of Eq. (15) is equivalent to a power spectral density,

which is a non-negative real quantity. Therefore, we can rearrange the denominator to give the bound

$$\mathcal{S}_{y^d_t y^d_t}(\boldsymbol{\nu}) + \mathcal{S}_{\varepsilon_t \varepsilon_t}(\boldsymbol{\nu}) \geq 0 \; \forall \boldsymbol{\nu}, \tag{16}$$

where (see Appendix A)

$$\mathcal{S}_{\varepsilon_t \varepsilon_t}(\boldsymbol{\nu}) = 2\sigma_\varepsilon^2 (cos(2\pi\boldsymbol{\nu}) - 1) \tag{17}$$

giving

$$\mathcal{S}_{y^d_t y^d_t}(\boldsymbol{\nu}) \geq 2\sigma_\varepsilon^2 (1 - cos(2\pi\boldsymbol{\nu})) \geq 0 \; \forall \boldsymbol{\nu}. \tag{18}$$

Data collection and pre-processing

Data were collected from three patients undergoing evaluation for surgical resection of epileptic brain tissue. The standard procedure for resective surgery involves implantation of subdural electrodes for intracranial electroencephalogram (iEEG) monitoring. Data from these electrodes provide information for mapping functional and pathological tissue to define the surgical margins. These data sets were collected with informed consent from patients with ethics approval from the St. Vincent's Hospital Melbourne Human Research Ethics Committee (HREC-A 006/08). Fig. 1 shows the arrangement of the iEEG electrodes for Patient I. The iEEG electrodes covered the temporal lobe and consisted of a square grid of 120 (8×15). The x-direction and y-direction electrode spacings were 0.5 cm and 1 cm, respectively, yielding a coverage of 7×7 cm^2 of the temporal cortex. The arrangement of the electrodes was the same for Patient II. For Patient III, a grid of 90 (6×15) electrodes was used. Data were acquired with a sampling rate of 5 kHz.

The iEEG recordings were pre-processed using several steps. The first step was to identify and exclude channels that had poor signal quality; this was performed by visual examination. Any linear trend was removed from the data by subtracting its least-squares fit. The data was then smoothed using a 5-point median filter to eliminate sharp fluctuations and discontinuities in the observations, which can lead to ringing when further filtering is applied. Next, the data was high-pass and low-pass filtered with cut-off frequencies of 1 Hz and
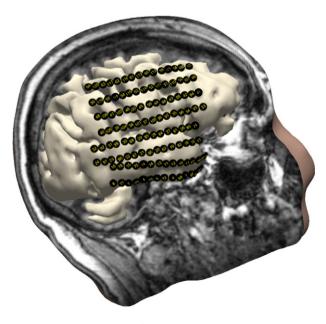

Fig. 1. X-ray (computed tomography) fused with structural magnetic resonance imaging showing the intracranial electrode positions. The electrodes were implanted to plan for epilepsy related surgery.

200 Hz, respectively, using 2nd-order Butterworth filters. To attenuate the effect of the electrical mains artifact, a notch filter with stop-band 45–55 Hz was also applied using a 2nd-order Butterworth filter. The data was then down-sampled from 5 kHz to 1 kHz and re-referenced to the differential montage. A differential montage removes the global effect of the reference electrode, which may lead to biases in the connectivity estimates. The re-referencing was accounted for in the estimation scheme.

Results

This section demonstrates the performance of the estimation scheme. The estimation results using synthetic data and real iEEG recordings from patients with seizures are presented. Synthetic data was generated using a nonlinear sigmoidal activation function, Eq. (4), and an activation function that was linearized about the firing threshold, Eq. (9), to test for errors resulting from linearization. Both isotropic and anisotropic connectivity functions in the forward model were tested.

The proposed estimator was also incorporated into a time-varying algorithm to estimate patient-specific dynamics in functional connectivity from real iEEG at the transition from normal activity to an epileptic seizure. In these examples, we demonstrate how information extracted from our estimation algorithm may be used to better understand the mechanisms underlying the brain's dynamics. The detailed steps of the algorithm are summarized in Table 2.

Example I: nonlinear and linearized models

To demonstrate the performance of the proposed algorithm, we reconstructed the underlying connectivity function from data generated using the one-dimensional Amari neural field model. The actual connectivity function is defined as a sum of Gaussian basis functions

$$w(r - r') = \sum_{i=0}^{n_\theta} \theta_i \psi_i (r - r'), \tag{19}$$

where θ_i is the weight and

$$\psi_i = exp\left(-\frac{(r - r' - \mu_i)^2}{\sigma_i^2}\right), \tag{20}$$

where μ_i and σ_i denote the basis function center and width, respectively. The parameters for each simulation are given in Table 3.

In each example, 250 s of data was generated over a one-dimensional field, $\Omega = [-30, 30]$, with periodic boundary condition and sampled at $T_s = 1$ ms (ten times faster than the membrane time constant (Stephan et al., 2008)) at 40 regularly spaced locations. The pick-up range of the sensors was defined by a Gaussian,

$$m(r - r') = exp\left(-\frac{(r - r')^2}{\sigma_m^2}\right), \tag{21}$$

Table 2
Algorithm for estimating the connectivity function. This algorithm shows the steps in the implementation of the connectivity function estimator.

1. Re-referencing:
$y^d_t = y^+_t - y^-_t$
2. Compute correlation terms and find the average over time:
$\bar{R}_{y^d_t y^d_t}$ and $\bar{R}_{y^d_t y^d_{t+1}}$
3. Normalize correlation terms, i.e., divide by the number of re-referenced observations.
4. Rearrange each $\bar{R}$ term such that the zero lag correlation is shifted to the first index using fftshift and flip commands (in Python or MATLAB).
5. Calculate $\mathcal{S}_{\varepsilon^d_t \varepsilon^d_t}(\nu)$. Note $R_{\varepsilon^d_t \varepsilon^d_t}(\boldsymbol{\tau})$ should have the same structure as the rearranged correlation terms in Step 4.
6. Compute the connectivity function $w(\cdot)$ using Eq. (15) and reverse Step 4.

Table 3
Parameters of the connectivity basis functions of the model used in Example I.

Parameter	Isotropic	Anisotropic-I	Anisotropic-II
θ_i	100, − 80, 5	80, − 80, 5, 15	200, − 200
σ_i	1.8, 2.4, 6	1.8, 2.4, 6, 2	2.4, 2.4
μ_i	0, 0, 0	0, 0, 0, − 3	− 0.5, 0.5

where σ_m sets the width and was chosen to be 0.9. The observation noise variance, σ_ε^2, was set to 0.1. The field disturbance covariance function, $\gamma(\cdot)$, was modeled by a Gaussian with width $\sigma_\gamma = 1.3$ (spatial) and scaled by the temporal standard deviation, $\sigma_d = 10$. The simulation parameters are summarized in Table 4.

Two sets of data were generated using the forward model, one with the nonlinear activation function given in Eq. (4) and the other with a linearized activation function given in Eq. (9), for each connectivity function. After re-referencing, there were $n_d = n_y - 1$ observation points.

The connectivity function estimates were obtained using Eq. (15), where the observation noise variance and synaptic decay terms were assumed to be known. The estimation results along with actual connectivities are shown in Fig. 2. The isotropic function in Fig. 2(A) forms a semi-compact support Mexican-hat shape commonly used in the neural field modeling (Amari, 1977; Atay and Hutt, 2005; Breakspear et al., 2010). The results demonstrate the ability of the proposed method to reconstruct the connectivity function from observed field, when either the nonlinear or linearized activation functions were used in the generative model. Note that the connectivity function could only be reconstructed at the observation locations (see Figs. 2(A–C)). This highlights the importance of the design of a measurement system such that it adequately samples the underlying field to avoid spatial aliasing (Yellott, 1982; Sanner and Slotine, 1992). The average (over time) spatial auto-correlation, $\bar{R}_{y_t^d y_t^d}$, and cross-correlation, $\bar{R}_{y_t^d y_{t+1}^d}$, of observed fields are also shown in Figs. 2(D–F). The figures show a subtle difference between average (auto-) cross-correlations amplitudes, all having a Mexican hat form. The Mexican hat form is an artifact of the re-referencing, and demonstrates that simple correlation analysis provides little indication of the true shape of the underlying connectivity function.

In the estimation algorithm, the observation noise variance was assumed to be known. An inaccurate guess of this parameter distorts the shape of the estimated connectivity function. To examine the effect of the observation noise, σ_ε^2, on the estimation result, we considered the data generated with varying noise levels using a forward model with

Table 4
Parameters of the model used to generate simulated data.

Symbol	Quantity	Value	Units
Domain			
Δ	Spatial discretization step	0.5	mm
T_s	Time step	0.001	s
T	Number of time steps	2.5×10^5	n.a.
Model			
ς	Slope of activation function	0.56 Wendling (2005)	mV^{-1}
v_0	Firing threshold	1.8 Marreiros et al. (2008)	mV
t_m	Membrane time constant	0.01 David and Friston (2003)	s^{-1}
σ_d	Temporal disturbance	10	mV
σ_γ	Disturbance covariance (spatial) width	1.3	mm
n_y	Number of sensors	40	n.a.
n_d	Number of observations after re-referencing	39	n.a.
Δ_y	Distance between sensors	1.5	mm
σ_m	Observation function width	0.9	mm
Σ_ε	Observation noise covariance	$0.1 \times \mathbf{I}_{n_y}$	mV^2

the anisotropic connectivity function shown in Fig. 2(C). An upper bound on the measurement noise, σ_ε^2, was calculated using Eq. (16), giving $0 \leq \sigma_\varepsilon^2 \leq 0.25$. If the observation noise variance was set to a value close to the upper or lower (zero) bounds, then the true shape of the connectivity function could not be recovered (see Figs. 3(A) and (E)). If the observation noise variance was close to the true value, the estimates captured the structural shape of the connectivity function (Figs. 3(B) and (D)). The estimated connectivity function using the actual value of measurement noise variance is plotted in Fig. 3(C), showing an accurate result.

Example II: intracranial EEG data

Intracranial EEG (iEEG) data from three patients were analyzed to demonstrate the utility of the method. In this section, we present the results from two of the three patients. The results from the third patient are presented in Appendix B. The electrographic seizures from Patients I and II are shown in Figs. 4(A) and (B), respectively. For the initial analysis, the time series was segmented into three time periods and given the labels: normal (background), electrographic seizure, and post-seizure periods. In the figures, the electrographic seizure onset is marked by the red line and the electrographic seizure end by the blue line. Note, our analysis identified a fourth time period that we have labeled pre-electrical onset, or pre-electrical for short. Before running the estimation algorithm, the two-dimensional (in space) data was transformed into one dimension as shown in Fig. 5(A), providing a higher number of regularly sampled observations.

For all analyses, a physiologically plausible value for the membrane time constant parameter of $t_m = 10$ ms was used. Varying the value of membrane time constant changes the value of variable ξ in Eq. (15), which only alters the connectivity function's amplitude at the origin (spatial lag zero) since the constant ξ in frequency domain is transformed to a delta function in the spatial domain.

In all examples, we assumed that the standard deviation of the measurement noise was the same at all electrodes and stationary across the normal, electrographic seizure, and post seizure periods. Following this, Eq. (16) was used to obtain an upper bound. As the iEEG envelope amplitude was non-stationary, we calculated upper bounds for the background, electrographic seizure, and post-seizure time periods separately. Note that since we did not expect the measurement noise to change over these time periods, computing the upper bounds on the segmented data provides the opportunity to obtain multiple estimates. Since we expect the measurement noise to be stationary over the entire duration of our datasets (approximately 6 min), we take the smallest of the three upper bounds as our best estimate. The corresponding bounds on the observation noise from Patient I for normal, electrographic seizure period, and post-seizure period were 8.6×10^{-4} mV^2, 8.8×10^{-3} mV^2, and 6×10^{-4} mV^2 respectively. The bounds for Patients II and III were of the same order of magnitude as Patient I. Following this analysis, a value slightly smaller than the minimum upper bound was used as an estimate for the observation noise, i.e., 1×10^{-4} mV^2 in subsequent analysis.

As demonstrated earlier, the shape of the estimated connectivity function is affected by the accuracy of observation noise variance, σ_ε^2. Therefore, we performed an analysis to test the sensitivity of this parameter and to check if we can tighten the bounds. To explore the effect of the observation noise, we estimated the connectivity functions over the normal, electrographic seizure, and post-seizure periods using 1000 different values in the range $0 \leq \sigma_\varepsilon^2 \leq 1 \times 10^{-3}$ mV^2 with step size 1×10^{-6} mV^2. The results are shown in Figs. 5(B–G). White dashed lines show the upper bound on the observation noise variance for the post-seizure period; any estimates above and slightly below this bound are distorted and must be rejected. For very small values of observation noise, no inhibitory connections are present, which cannot be physiologically plausible for at least pre-electrical onset and post-seizure regimes. Therefore, we considered a lower bound on the observation noise where the inhibition first appeared, around
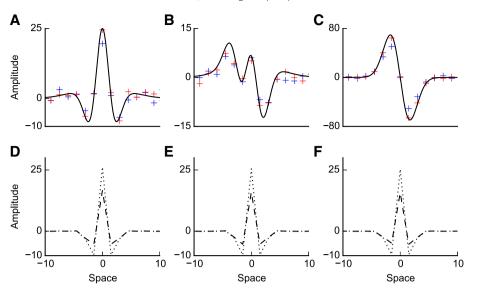
Fig. 2. Estimation of the connectivity function using synthetic data. (A–C) Estimated connectivity functions, from synthetic data for both nonlinear and linearized forward models, are shown with blue and red crosses, respectively. The actual connectivity functions are shown by solid lines. (D–F) The average (over-time) auto-correlations, $\bar{R}_{y_t^d y_t^d}$ (dotted line) and cross-correlations, $\bar{R}_{y_t^d y_{t+1}^d}$ (dashed line) of observations corresponding to the row above.

$\sigma_\varepsilon^2 = 1 \times 10^{-5}$ mV^2. Above this lower bound, the estimated connectivity functions had a weaker inhibition strength during the electrographic seizure period than the other time segments (as shown in Fig. 6(B)).

Following our analysis, the plausible range of observation noise was $1 \times 10^{-5} < \sigma_\varepsilon^2 < 6 \times 10^{-4}$ mV^2. For all subsequent analyses, a conservative value within this range was chosen, where $\sigma_\varepsilon^2 = 1 \times 10^{-4}$ mV^2. For this value, the estimated connectivity function had a Mexican-hat shape with short range excitation, lateral inhibition, and long range excitation over the normal and post-seizure regimes for Patient I (Figs. 6(A) and (C)). During the electrographic seizure, the connectivity function consisted of only short range excitation (Fig. 6(B)). For Patient II, a Mexican-hat shaped connectivity function was estimated with slightly different amplitudes for central excitation and surrounding inhibition over normal, electrographic seizure, and post-seizure regimes (Figs. 6(D–F)).

To detect temporal changes in the shape of the connectivity functions, we performed a sliding window-based analysis. The window size was set to four seconds (4000 samples) with a one second overlap. The overlap compensated for the losses in temporal resolution. For Patient I, this approach resulted in 223 quasi-stationary segments. The results for Patient I are shown in Fig. 7. In Fig. 7(A), the electrical onset of the seizure is marked by the red line (index 91) and the electrical end is marked by the blue line (index 123). From the analysis, it can be seen that the regular patterns in the shape of the connectivity function started to change at around the time marked by the magenta line (index 79), which we label pre-electrical onset. After electrical onset,

the lateral inhibition began to fluctuate and disappeared until the termination of the seizure. This suggests that the early sporadic spiking is driven by fluctuations in the lateral inhibition (see Fig. 4(A)). The excitation and lateral inhibition amplitudes are also shown in Fig. 7(B), highlighting a paradoxical increase in lateral inhibition prior to the electrical onset of the seizure. This is in agreement with other experimental and computational studies reported in the literature (Khalilov et al., 2002; Wendling et al., 2005; Freestone et al., 2013).

The logarithm of the normalized absolute ratio of the central excitation and lateral inhibition peaks is shown in Fig. 7(C). The figure clearly shows a relatively constant ratio between excitation and inhibition prior to and after the seizure, with a significant jump during the seizure period. Note that there was a decrease in the ratio during the pre-electrical onset period due to the increase in lateral inhibition. A scatter plot of the inhibition versus the excitation is shown in Fig. 7(D). This figure better shows the clusters for different activity regimes. It is interesting to see that the pre-electrical onset cluster is at the opposite side of the background cluster to the seizure.

The same analysis was performed for Patient II, and the results are presented in Fig. 8. An increase in lateral inhibition prior to the electrical onset of the seizure can be also observed (see Fig. 8(B)). The seizure cluster for Patient II is at a different location compared to Patient I (see Fig. 8(D)), due to the existence of surround inhibition during the seizure.

While we do not suggest that these results demonstrate that this approach gives seizure prediction, we do claim that this method is
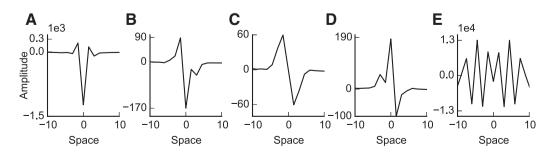

Fig. 3. Effect of the observation noise variance on the connectivity function estimation. Each panel shows the result of the algorithm using different values of measurement noise, σ_ε^2. (A) $\sigma_\varepsilon^2 = 0$, (B) $\sigma_\varepsilon^2 = 0.09$, (C) $\sigma_\varepsilon^2 = 0.1$ (corresponding to the actual value), (D) $\sigma_\varepsilon^2 = 0.11$, (E) $\sigma_\varepsilon^2 = 0.27$ (above the upper bound).
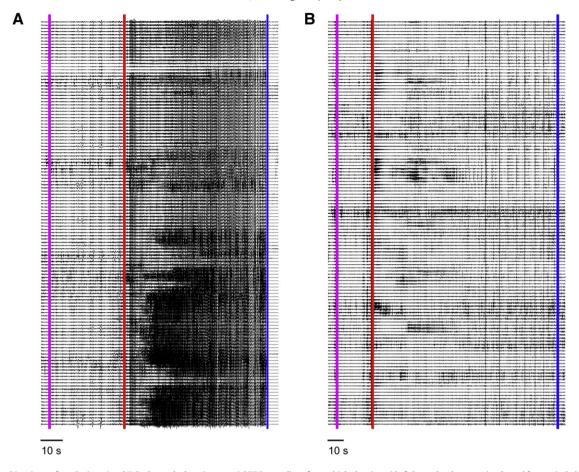
Fig. 4. Electrographic seizures from Patients I and II. Both panels show intracranial EEG recordings from a high-density grid of electrodes that were implanted for surgical planning. The red lines mark the earliest electrical evidence of seizure onset, as marked by clinicians blinded to the analysis. The blue lines mark the end of the seizures. The magenta lines mark the time points where large changes occurred in the shape of the estimated connectivity functions. (A) Data from Patient I. The pre-electrical onset interval for Patient I had a duration of 36 s. (B) Data from Patient II. The pre-electrical onset interval for Patient II had a duration of 24 s.
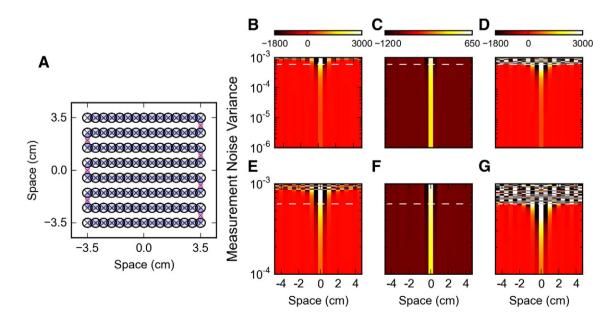

Fig. 5. Electrodes and the estimation results for different values of measurement noise variance for Patient I. (A) A grid macro-electrode being implanted in surgery, showing electrode arrangement with the distance between adjacent sensors 0.5 and 1 cm along x and y coordinates respectively. The blue line shows the conversion of the two dimensional grid into the one dimensional array where missing measurements along y axis is interpolated using their adjacent electrodes (shown by red crosses). Estimated connectivity function from iEEG data for normal, electrographic seizure and post-seizure regimes for different values of measurement noise, σ_ε^2, are shown in (B), (C), and (D), respectively. The white dashed lines show the upper bound on the observation noise variance (6×10^{-4} mV^2) obtained for the post-seizure period using Eq. (16). (E–G) A zoomed view of estimated connectivity functions for a more focused range of observation noise parameters, $10^{-4} \leq \sigma_\varepsilon^2 \leq 10^{-3}$, are re-plotted for a better visualization.
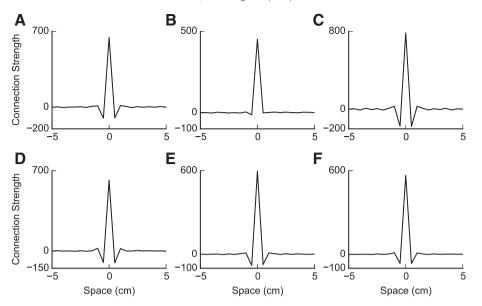
Fig. 6. Estimated connectivity functions from iEEG data for normal, electrographic seizure and post-seizure regimes for Patient I and Patient II. Estimated connectivity function using one dimensional iEEG for (A) normal, (B) electrographic seizure and (C) post-seizure regimes for $\sigma_\epsilon^2 = 1 \times 10^{-4}$ for Patient I. Estimated connectivity function using one dimensional iEEG for (D) normal, (E) electrographic seizure and (F) post-seizure regimes for $\sigma_\epsilon^2 = 1 \times 10^{-4}$ for Patient II.

capable of inferring hidden aspects of physiology that may prove to be useful in better understanding epilepsy. The question of whether this is useful for seizure prediction remains open, and we expect solutions to be highly patient-specific.

Discussion

We have presented an efficient and novel approach to identifying spatiotemporal cortical dynamics using electrophysiological recordings. In particular, our results demonstrate an approach that can be used to estimate intra-cortical connectivity from synthetic and real iEEG data.

Physiological relevance of connectivity estimates

Before discussing the physiological relevance of the connectivity estimates, we would like to stress that the contribution of this paper is a method. The results from patient data are included to demonstrate the application of the method. Further investigation is required to draw robust conclusions regarding the mechanisms that govern seizure initiation and spread. Nevertheless, the results are sufficiently interesting to warrant a small discussion.

The intra-cortical connectivity estimates from the iEEG data lead to several interesting insights. Each patient had a qualitative change in connectivity profile before any electrographic changes could be seen in the iEEG. During this pre-electrical onset phase, all patients had an
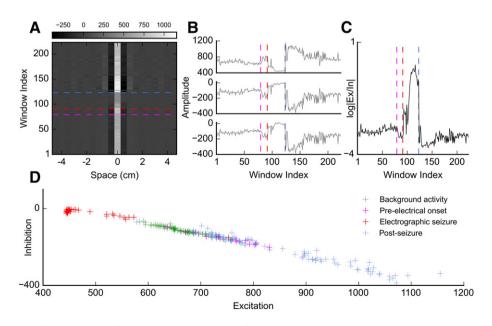

Fig. 7. Estimation of the connectivity function over time for Patient I. (A) The connectivity function estimates over 669 s (223 windows), where the gray scale shows the connectivity strength. The central section of the estimated connectivity function from − 5 to 5 cm is illustrated in the figure. (B) The central excitation and lateral inhibition amplitudes over 669 s (223 windows) are shown in top, middle and bottom panels. (C) The logarithm of the normalized absolute ratio of the central excitation to sum of lateral inhibitions of the estimated connectivity over time. The pre-electrical onset, electrical seizure start and electrical end are shown by magenta, red and blue lines respectively. (D) Scatter plot of inhibition versus excitation of the estimated connectivity function during normal (background), pre-electrical onset, electrographic seizure and post-seizure regimes.
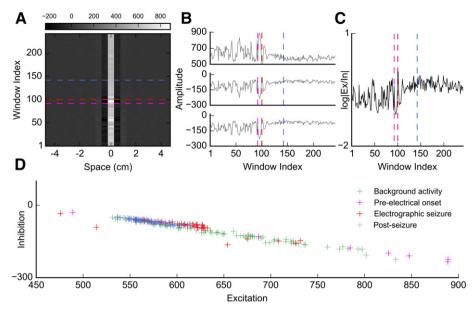
Fig. 8. Estimation of the connectivity function over time for Patient II. (A) The connectivity function estimates over 725 s (242 windows), where the gray scale shows the connectivity strength. The central section of the estimated connectivity function from − 5 to 5 cm is illustrated in the Figure. (B) The central excitation and lateral inhibition amplitudes over 725 s (242 windows) are shown in top, middle and bottom panels. (C) The logarithm of the normalized absolute ratio of the central excitation to sum of lateral inhibitions of the estimated connectivity over time. The pre-electrical onset and electrical seizure onset and the electrical end of the seizure are shown by magenta, red and blue lines respectively. (D) Scatter plot of inhibition versus excitation of the estimated connectivity function during normal (background), pre-electrical onset, electrographic seizure and post-seizure regimes.

increase in both local excitation and surrounding inhibition. Most interestingly, the change in surrounding inhibition strength was dominant during this pre-electrical onset time period. However, at electrographic seizure onset, the local excitatory connections overtook the lateral inhibitory effects and dominated the dynamics. Increased surround inhibitory activity at seizure onset has recently been observed at the microscopic scale and has been dubbed the ictal penumbra (Schevon et al., 2012). We expect the onset seizures to be observed earlier at the microscopic scale, compared to the mesoscopic measurements of iEEG. Therefore, the increased inhibitory drive that was observed prior to the electrographic onset may be related to the aforementioned phenomena. Correspondence between the microscopic observations and our data-driven model-based results would suggest that the intra-cortical connectivity estimates enable a physiologically relevant interpretation. However, the global connectivity function enforces the seizures to arise as an emergent phenomenon of the entire field and the notion of a discrete focus is lost.

Patient I and Patient III had seizures that spread electrographically across the array, while Patient II's seizure did not. The connectivity estimates provide a possible explanation for differences in seizure spread, where Patient I and III both experienced severe loss of lateral inhibition during the events, where Patient II maintained a relatively normal connectivity structure. This result points to the loss of surround inhibition as the mechanism for seizure spread. Evidence of lateral inhibition being the mechanism for containing epileptic events has previously been reported (Prince and Wilder, 1967). Although the spatial resolution of their work was limited, Prince and Wilder (1967) demonstrated the spatial scale of lateral inhibition extended to a spatial scale 8 mm, which is in agreement with our findings (5 mm). These findings provide further evidence of the physiological relevance of the connectivity estimates.

Advantages over other estimation algorithms

The estimation procedure presented in this paper has several advantages over previously published methods. For example, although the derivation is complex, the closed-form estimator presented in Eq. (15) is algorithmically straightforward when compared to alternative algorithms (Dewar et al., 2009; Scerri et al., 2009; Freestone et al., 2011; Aram et al., 2012). Furthermore, by using the closed-form solution, one can see how errors in the knowledge of other parameters of the model that are assumed to be known will affect
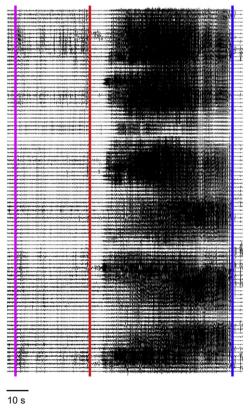

Fig. 9. Electrographic seizure from Patient III. Intracranial EEG recording from a high-density grid of electrode that was implanted for surgical planning. The red line marks the earliest electrical evidence of seizure onset, as marked by clinicians blinded to the analysis. The blue line marks the electrical end of the seizure. The magenta lines mark the time point where large changes occurred in the shape of the estimated connectivity function. The pre-electrical onset interval for Patient III had a duration of 36 s.
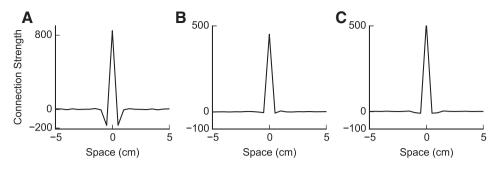
Fig. 10. Estimated connectivity functions from iEEG data for normal, electrographic seizure and post-seizure regimes for Patient III. Estimated connectivity function using one dimensional iEEG for (A) normal, (B) electrographic seizure, and (C) post-seizure regimes for observation noise $\sigma_\epsilon^2 = 1 \times 10^{-4}$.

the estimates. The closed-form solution also leads to new insights into the integro-differential-based neural fields models by formally relating other parameters of the model to the connectivity structure. For example, it can be seen that the time constant parameter not only affects the temporal frequency of the neural field but also its spatial frequency, which is a uniform shift in power across all frequencies.

Another key contribution of this estimation method is the introduction of differential electrode referencing in the estimation scheme. Explicitly accounting for this re-referencing scheme is normally overlooked in data-driven modeling, particularly with intracranial electrophysiological measurements. Other methods that map from the measurement space to the source space in scalp EEG take care of re-referencing via source localization algorithms (Mosher and Leahy, 1998). It is critically important to account for the common referencing effects of electrophysiology amplifiers, even with intracranial measurements. By accounting for the differential montage, the algorithm transforms the problem closer to the source space without requiring source localization algorithms.

Limitations of the framework

The estimation results are sensitive to observation noise. Assuming an incorrect value of the noise level can result in a substantial distortion

of the estimated connectivity function. For the real iEEG data, we derived and calculated a theoretical upper bound on the observation noise variance. We then performed numerous experiments by finely sampling the permitted range to study the consistency of our estimates. While the resulting estimates show promise, we cannot rule out the possibility that the resultant connectivity function is influenced and distorted by the observation noise. It should also be noted that the model is an extreme simplification of neural interactions. The results should be taken as a proof-of-principal of the method derived in this paper, rather than a direct proof of epilepsy-related functional connectivity changes. Nevertheless, the model-based framework proposed in this paper enables meaningful estimates to track neural activity transitions from normal to abnormal states.

We assume the membrane time constant does not change over the entire estimation period, while an earlier study showed that it varies during normal and abnormal activities (Freestone et al., 2013), resembling that of a conductance-based synaptic mechanism (rather than the current-based mechanism used in this framework). In the study by Freestone et al. (2013), intracranial EEG data was fit to the neural mass model of Jansen and Rit (1995), and the dynamical evolution of the synaptic gains and time constants were estimated. The reciprocal of both the inhibitory and excitatory time constants increased at seizure onset. Therefore, the increase in the local excitation at electrographic
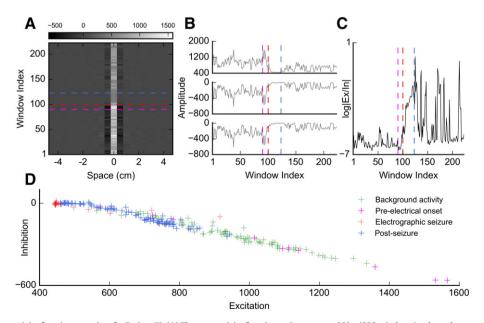

Fig. 11. Estimation of the connectivity function over time for Patient III. (A) The connectivity function estimates over 669 s (223 windows), where the gray scale shows the connectivity strength. The central section of the estimated connectivity function from −5 to 5 cm is illustrated in the Figure. (B) The central excitation and lateral inhibition amplitudes over 669 s (223 windows) are shown in top, middle and bottom panels. (C) The logarithm of the normalized absolute ratio of the central excitation to sum of lateral inhibitions of the estimated connectivity over time. The pre-electrical onset, electrographic seizure, and the electrical end of the seizure are shown by magenta, red and blue lines respectively. (D) Scatter plot of inhibition versus excitation of the estimated connectivity function during normal (background), pre-electrical onset, electrographic seizure, and post-seizure regimes.

seizure onset could be a result of a change of the membrane time constant, which shifts the level of central excitation of the connectivity function in our model (see Eq. (15)). Similar arguments can be made for the activation function parameters. However, this is less likely to be an issue since the activation function slope is multiplicative to the connectivity function (the ratio of excitation to inhibition is not effected) and the threshold parameter cancels in the differential scheme.

We assumed instantaneous action potential propagation in the neural field model. However, Ghosh et al. (2008) emphasized that time delays play a significant role in determining the dynamics of neural fields. Therefore, the space-time structure of the brain connectivity (Sanz-Leon et al., 2015) must be accounted for to capture completely the physiologically accurate connectivity. Nevertheless, the proof-of-principal provided in this paper demonstrates that an approximate model, with its computational simplicity, may lead to useful insights.

It has been shown that subcortical structures play an important role in initiation of epileptic seizures (Norden and Blumenfeld, 2002). For example, the strength of excitatory cortico-thalamic interactions has been shown to be crucial in transitions from healthy to abnormal states in absence seizures (Marten et al., 2009a,b). In the current study, such mechanisms were not included in the model. However, the disturbance covariance can be thought of as an input.

Extensions to the framework

We assume that the dynamics are dominated by homogeneous connectivity and, hence, the long range heterogeneous cortico-cortical connections cannot be recovered. Nevertheless, insights into the short range intra-cortical structure can be estimated. Algorithms that account for the addition of heterogeneous connectivity should be explored, such as two-point connection structure (Qubbaj and Jirsa, 2007). Frameworks currently exist for estimating and tracking inter-regional functional connectivity using mesoscopic neural models (Freestone et al., 2014). However, the aforementioned framework is far more complicated than the method introduced in this paper.

Another extension to this framework would be to use a more accurate post-synaptic response function with a finite rise time, leading to a second-order Markovian dependency. The resultant model would be a second-order IDE, which can be written as two coupled first-order differential equations. This way, the synaptic dynamics can be better approximated (Van Rotterdam et al., 1982), possibly leading to more physiologically realistic estimates.

The emergence of clusters in the excitation–inhibition space encourages the use of classification techniques to interpret the data and forecast future events. This is beyond the scope of this paper, which focuses on the methods, but the preliminary data warrants future investigation. For example, k-means clustering (Hartigan and Wong, 1979) may be applied to identify regions in the excitation–inhibition space where there is a greater likelihood of imminent seizures.

The ability to track the ratio of excitation to inhibition is of great importance for developing new therapies, where the ratio can be used as a control variable. For example, therapies based on electrical stimulation are evaluated by how well they reduce seizure frequencies. An analogy of this control framework would be to test a system for automatically controlling the speed of a car (i.e., cruise control) based on how often the driver gets a speeding ticket. Clearly, this is not a good idea for controlling speed, nor is it a good idea to measure the effect of a therapy by accessing seizure frequency. A better approach is to track the root cause of seizures, which has been strongly linked to the balance of excitation and inhibition.

The data that was used to demonstrate the proof-of-principal was spatially under-sampled. Ideally, the distance between sensors should be less than 1.25 mm to capture the spatial frequencies of mesoscopic neural dynamics (Freeman et al., 2000). The distance between the sensors in our case was 5 mm, so the higher spatial frequencies were lost. Data from voltage sensitive dyes or more densely spaced electrodes would be well suited to this method. A difficulty in using electrode arrays is in having a sufficient number of recording sites to obtain an accurate estimate of the correlation structure. Further work should be directed towards applying and validating the framework on more subjects with higher fidelity data, such as that used in other recent studies with implanted Utah arrays (Blackrock Microsystems) in epilepsy patients (Truccolo et al., 2011).

The bigger picture

The correlation analysis introduced in this work provides an estimate of the connectivity function. It does not, however, provide an estimate of the dynamics of the neural field. Nevertheless, if one is interested in tracking the dynamics of the field, the solutions provided in this paper can be used to inform priors of other more complicated iterative algorithms (Freestone et al., 2011; Aram et al., 2012), which is an extremely challenging task. In our earlier work, the neural field and the connectivity function were represented by a set of weighted basis functions, where the weights were considered unknown. The estimate from this paper can be used as a guide to constrain the placement of basis functions of the connectivity function. Given an initial estimate from the closed-form equation provided in this current paper, the corresponding coefficients of the connectivity basis functions can be corrected using an iterative state-parameter estimation algorithm for a more accurate representation of the connectivity structure.

Appendix A. Derivation of the connectivity function estimator

For the derivation of the connectivity function estimator of the neural field equations, we switch to a more compact notation to define convolution and correlation operators. The spatial convolution for arbitrary functions $a(\cdot)$ and $b(\cdot)$ are denoted as

$$\int_{\boldsymbol{\Omega}} a(\mathbf{r} - \mathbf{r}')b(\mathbf{r}')d\mathbf{r}' = a(\mathbf{r}) * b(\mathbf{r}) = (a * b)(\mathbf{r}) \tag{A.1}$$

and the spatial cross-correlation is denoted as

$$\int_{\boldsymbol{\Omega}} a(\mathbf{r})b(\mathbf{r} + \boldsymbol{\tau})d\mathbf{r} = a(\boldsymbol{\tau}) \star b(\boldsymbol{\tau}) = (a \star b)(\boldsymbol{\tau}), \tag{A.2}$$

where $\boldsymbol{\tau}$ is the spatial shift.

Without loss of generality, we will derive the estimator under the assumption that the electrode locations are infinitesimally close. This assumption will simplify the notation. Here, we define the infinitesimally close differential observations as $y_t^d(\mathbf{r})$. Further, we will define the underlying neural fields and components that contribute to the measurements $y_t^+(\mathbf{r})$ and $y_t^-(\mathbf{r})$ as $v_t^+(\mathbf{r})$ and $v_t^-(\mathbf{r})$, respectively, such that

$$y_t^+(\mathbf{r}) = (m * v_t^+)(\mathbf{r}) + \mathbf{c}_t^+ + \varepsilon_t^+(\mathbf{r}), \tag{A.3}$$

and

$$y_t^-(\mathbf{r}) = (m * v_t^-)(\mathbf{r}) + \mathbf{c}_t^- + \varepsilon_t^-(\mathbf{r}), \tag{A.4}$$

where the common-mode signals $\mathbf{c}_t^+ = \mathbf{c}_t^-$. Following this, the differential electrophysiological measurement at time $t + 1$ can be written in the form

$$y_{t+1}^d(\mathbf{r}) = (m * (v_{t+1}^+ - v_{t+1}^-))(\mathbf{r}) + \varepsilon_{t+1}^+(\mathbf{r}) - \varepsilon_{t+1}^-(\mathbf{r}), \tag{A.5}$$

noting that the common-mode signals have canceled. By substituting Eq. (7) for the $v_{t+1}^{+,-}(\mathbf{r})$ terms, we obtain

$$y_{t+1}^d(\mathbf{r}) = (m * (\xi(v_t^+ - v_t^-) + T_s(g_t^+ - g_t^-) + e_t^+ - e_t^-))(\mathbf{r}) \\ + \varepsilon_{t+1}^+(\mathbf{r}) - \varepsilon_{t+1}^-(\mathbf{r}), \tag{A.6}$$

where

$$g_t^{+,-}(\mathbf{r}) = \left(w * \hat{f}_t^{+,-}\right)(\mathbf{r}) \tag{A.7}$$

is the weighted collection of firing rates influencing each point in the field from all other points.

The spatial cross-correlation between re-referenced consecutive (in time) observations is defined as

$$R_{y_{t+1}^d y_t^d}(\boldsymbol{\tau}) = \left(y_{t+1}^d \star y_t^d\right)(\boldsymbol{\tau}). \tag{A.8}$$

The goal of the derivation is to substitute the model equations into the auto-correlation and cross-correlation of the observations and to separate the resulting expressions using the convolution theorem. In this way, we show that a closed-form expression can be derived that relates correlation terms to the connectivity function. In the next step, we substitute Eq. (A.6) for $y_{t+1}^d(\mathbf{r})$ and expand to give

$$R_{y_{t+1}^d y_t^d}(\boldsymbol{\tau}) = \left(\xi(m * (v_t^+ - v_t^-)) \star y_t^d\right)(\boldsymbol{\tau}) \\ + \frac{T_s \varsigma}{4}\left(m * w * (v_t^+ - v_t^-) \star y_t^d\right)(\boldsymbol{\tau}) \\ + \left(m * (e_t^+ - e_t^-) \star y_t^d\right)(\boldsymbol{\tau}) \\ + \left((\varepsilon_{t+1}^+ - \varepsilon_{t+1}^-) \star y_t^d\right)(\boldsymbol{\tau}). \tag{A.9}$$

The last two terms of Eq. (A.9) are zero since the disturbance at time t and measurement noise at time $t + 1$ are assumed to be independent of the observations at time t. Substituting $y_t^d - \varepsilon_t^+ + \varepsilon_t^-$ for $m * (v_t^+ - v_t^-)$ we have

$$R_{y_{t+1}^d y_t^d}(\boldsymbol{\tau}) = \xi\left(\left(y_t^d - \varepsilon_t^+ + \varepsilon_t^-\right) \star y_t^d\right)(\boldsymbol{\tau}) \\ + \frac{T_s \varsigma}{4}\left(w * \left(y_t^d - \varepsilon_t^+ + \varepsilon_t^-\right) \star y_t^d\right)(\boldsymbol{\tau}). \tag{A.10}$$

To isolate the connectivity function, the order of the convolution and cross-correlation is reversed by recognizing that the property $(a * b)(\boldsymbol{\tau}) \star c(\boldsymbol{\tau}) = a(-\boldsymbol{\tau}) * (b \star c)(\boldsymbol{\tau})$ applies (see Appendix C). Therefore,

$$R_{y_{t+1}^d y_t^d}(\boldsymbol{\tau}) = \xi\left(\left(y_t^d - \varepsilon_t^+ + \varepsilon_t^-\right) \star y_t^d\right)(\boldsymbol{\tau}) \\ + \frac{T_s \varsigma}{4} w(-\boldsymbol{\tau}) * \left(\left(y_t^d - \varepsilon_t^+ + \varepsilon_t^-\right) \star y_t^d\right)(\boldsymbol{\tau}). \tag{A.11}$$

The common term on the right hand side of Eq. (A.11) can be written as

$$R_{y_t^d y_t^d}(\boldsymbol{\tau}) + R_{\varepsilon_t^d \varepsilon_t^d}(\boldsymbol{\tau}) = \left(\left(y_t^d - \varepsilon_t^+ + \varepsilon_t^-\right) \star y_t^d\right)(\boldsymbol{\tau}) \tag{A.12}$$

where for a sufficiently large number of observations or after averaging over a sufficiently large number of samples (since the measurement noise is spatially and temporally white) we can write

$$R_{\varepsilon_t^d \varepsilon_t^d}(\boldsymbol{\tau}) = \sigma_\varepsilon^2 \delta(\tau_x - 1)\delta(\tau_y) - 2\sigma_\varepsilon^2 \delta(\tau_x)\delta(\tau_y) \\ + \sigma_\varepsilon^2 \delta(\tau_x + 1)\delta(\tau_y). \tag{A.13}$$

Note, the two dimensional Dirac Delta function (since we are assuming continuous space for derivation) is written in a separable form where the spatial shifts along x and y directions are denoted by τ_x and τ_y respectively. The non-zero terms with the spatial shift are due to re-referencing. A factor of two appears at zero spatial shift, $\boldsymbol{\tau} = 0$, since y^d consists of the difference between adjacent observations that both contribute to the noise variance. Note that the result for differential re-referencing along y-direction can be easily obtained by swapping τ_x with τ_y in Eq. (A.13). To minimize the effects of the observation noise and process disturbance, the average over time of the spatial auto-correlation and cross-correlation can be used; i.e.,

$$\overline{R}_{y_{t+1}^d y_t^d}(\boldsymbol{\tau}) = \frac{1}{T-1}\sum_{t=1}^{T-1} R_{y_{t+1}^d y_t^d}(\boldsymbol{\tau}) \tag{A.14}$$

$$\overline{R}_{y_t^d y_t^d}(\boldsymbol{\tau}) = \frac{1}{T}\sum_{t=1}^{T} R_{y_t^d y_t^d}(\boldsymbol{\tau}). \tag{A.15}$$

Now by substituting Eq. (A.12) into Eq. (A.11) using the average quantities, $\overline{R}_{y_{t+1}^d y_t^d}(\boldsymbol{\tau})$ and $\overline{R}_{y_t^d y_t^d}(\boldsymbol{\tau})$, taking Fourier transform, and then rearranging, we have

$$w(\boldsymbol{\tau}) = \frac{4}{T_s \varsigma} \mathcal{F}^{-1} \overline{\left(\frac{\mathcal{S}_{y_{t+1}^d y_t^d}(\boldsymbol{\nu})}{\mathcal{S}_{y_t^d y_t^d}(\boldsymbol{\nu}) + \mathcal{S}_{\varepsilon_t^d \varepsilon_t^d}(\boldsymbol{\nu})} - \xi\right)} \tag{A.16}$$

$$= \frac{4}{T_s \varsigma} \mathcal{F}^{-1}\left(\frac{\mathcal{S}_{y_t^d y_{t+1}^d}(\boldsymbol{\nu})}{\mathcal{S}_{y_t^d y_t^d}(\boldsymbol{\nu}) + \mathcal{S}_{\varepsilon_t^d \varepsilon_t^d}(\boldsymbol{\nu})} - \xi\right) \tag{A.17}$$

where

$$\mathcal{S}_{y_t^d y_{t+1}^d}(\boldsymbol{\nu}) = \mathcal{F}\left\{\overline{R}_{y_t^d y_{t+1}^d}(\boldsymbol{\tau})\right\}, \tag{A.18}$$

$$\mathcal{S}_{y_t^d y_t^d}(\boldsymbol{\nu}) = \mathcal{F}\left\{\overline{R}_{y_t^d y_t^d}(\boldsymbol{\tau})\right\}, \tag{A.19}$$

$$\mathcal{S}_{\varepsilon_t^d \varepsilon_t^d}(\boldsymbol{\nu}) = \mathcal{F}\left\{R_{\varepsilon_t^d \varepsilon_t^d}(\boldsymbol{\tau})\right\} \\ = 2\sigma_\varepsilon^2(\cos(2\pi\boldsymbol{\nu}) - 1), \tag{A.20}$$

are the power spectral densities and $\boldsymbol{\nu}$ is the spatial frequency.

The over-bar denotes the complex conjugate operator. As the connectivity is a real function, the complex conjugate operator essentially swap the order of the correlation in the numerator of Eq. (A.16). It should be noted that all the formulations hold in one-dimensional space, however, Eq. (A.13) simplifies to

$$R_{\varepsilon_t^d \varepsilon_t^d}(\tau) = \sigma_\varepsilon^2 \delta(\tau - 1) - 2\sigma_\varepsilon^2 \delta(\tau) + \sigma_\varepsilon^2 \delta(\tau + 1). \tag{A.21}$$

Appendix B. Results for Patient III

In this section, the results that demonstrate the utility of the proposed algorithm for Patient III are provided. The electrographic seizure from Patient III is shown in Fig. 9. In the figure, the pre-electrical onset is marked by magenta line, the electrographic seizure onset is marked by the red line and the electrical end by the blue line. From Fig. 10(C) it can be seen that the post-seizure connectivity exhibits very weak

inhibition, which perhaps led to sporadic spikes after the seizure stopped (see Fig. 9). An increase in lateral inhibition prior to the electrical start of the seizure can be also observed (Fig. 11(B)). Another explanation for the sporadic spikes in the post-seizure period for Patient III could be large peaks in the excitation to inhibition ratio after the seizure stopped (Fig. 11(C)). The clusters for different activity regimes are shown in Fig. 11(D). The sporadic spikes in the ratio of excitation to inhibition in the post-seizure period moved the post-seizure cluster for Patient III towards the seizure region.

Appendix C. Convolution and correlation

In this section, the property of the cross-correlation and the convolution used in Appendix A is derived. To show

$$(a * b)(\boldsymbol{\tau}) \star c(\boldsymbol{\tau}) = a(-\boldsymbol{\tau}) * (b \star c)(\boldsymbol{\tau}), \tag{C.1}$$

first note that cross-correlation function is related to the convolution by Yarlagadda (2009),

$$(a \star b)(\boldsymbol{\tau}) = a(-\boldsymbol{\tau}) * b(\boldsymbol{\tau}). \tag{C.2}$$

Therefore, Eq. (C.1) can be written as

$$\begin{aligned}
(a * b)(\boldsymbol{\tau}) \star c(\boldsymbol{\tau}) &= (a * b)(-\boldsymbol{\tau}) * c(\boldsymbol{\tau}) \\
&= a(-\boldsymbol{\tau}) * (b(-\boldsymbol{\tau}) * c(\boldsymbol{\tau})) \\
&= a(-\boldsymbol{\tau}) * (b \star c)(\boldsymbol{\tau}).
\end{aligned} \tag{C.3}$$

References

Amari, S., 1977. Dynamics of pattern formation in lateral-inhibition type neural fields. Biol. Cybern. 27, 77–87.

Aram, P., Freestone, D., Dewar, M., Scerri, K., Jirsa, V., Grayden, D., Kadirkamanathan, V., 2012. Spatiotemporal multi-resolution approximation of the Amari type neural field model. Neuroimage 66, 88–102.

Atay, F., Hutt, A., 2005. Stability and bifurcations in neural fields with finite propagation speed and general connectivity. SIAM J. Appl. Math. 65, 644–666.

Breakspear, M., Heitmann, S., Daffertshofer, A., 2010. Generative models of cortical oscillations: neurobiological implications of the Kuramoto model. Front. Hum. Neurosci. 4.

David, O., Friston, K., 2003. A neural mass model for MEG/EEG: coupling and neuronal dynamics. Neuroimage 20, 1743–1755.

Dewar, M., Scerri, K., Kadirkamanathan, V., 2009. Data-driven spatio-temporal modeling using the integro-difference equation. IEEE Trans. Signal Process. 57, 83–91.

Freeman, W., Rogers, L., Holmes, M., Silbergeld, D., 2000. Spatial spectral analysis of human electrocorticograms including the alpha and gamma bands. J. Neurosci. Methods 95, 111–121.

Freestone, D.R., Aram, P., Dewar, M., Scerri, K., Grayden, D.B., Kadirkamanathan, V., 2011. A data-driven framework for neural field modeling. Neuroimage 56, 1043–1058.

Freestone, D., Kuhlmann, L., Chong, M., Nesic, D., Grayden, D., Aram, P., Postoyan, R., Cook, M., 2013. Patient-specific neural mass modelling — stochastic and deterministic methods. World Scientificpp. 63–82 (chapter 5).

Freestone, D.R., Karoly, P.J., Nešić, D., Aram, P., Cook, M.J., Grayden, D.B., 2014. Estimation of effective connectivity via data-driven neural modeling. Front. Neurosci. 8.

Friston, K.J., Harrison, L., Penny, W., et al., 2003. Dynamic causal modelling. Neuroimage 19, 1273–1302.

Ghosh, A., Rho, Y., McIntosh, A., Kötter, R., Jirsa, V., 2008. Noise during rest enables the exploration of the brain's dynamic repertoire. PLoS Comput. Biol. 4, e1000196.

Gorzelic, P., Schiff, S., Sinha, A., 2013. Model-based rational feedback controller design for closed-loop deep brain stimulation of Parkinson's disease. J. Neural Eng. 10, 026016.

Hartigan, J.A., Wong, M.A., 1979. Algorithm as 136: a k-means clustering algorithm. Appl. Stat. 100–108.

Jansen, B., Rit, V., 1995. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. Biol. Cybern. 73, 357–366.

Khalilov, I., Hirsch, J., Cossart, R., Ben-Ari, Y., 2002. Paradoxical anti-epileptic effects of a glur5 agonist of kainate receptors. J. Neurophysiol. 88, 523–527.

Marreiros, A., Daunizeau, J., Kiebel, S., Friston, K., 2008. Population dynamics: variance and the sigmoid activation function. Neuroimage 42, 147–157.

Marten, F., Rodrigues, S., Benjamin, O., Richardson, M., Terry, J., 2009a. Onset of polyspike complexes in a mean-field model of human electroencephalography and its application to absence epilepsy. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 367, 1145.

Marten, F., Rodrigues, S., Suffczynski, P., Richardson, M.P., Terry, J.R., 2009b. Derivation and analysis of an ordinary differential equation mean-field model for studying clinically recorded epilepsy dynamics. Phys. Rev. E 79, 021911.

Mosher, J.C., Leahy, R.M., 1998. Recursive music: a framework for EEG and MEG source localization. IEEE Trans. Biomed. Eng. 45, 1342–1354.

Norden, A.D., Blumenfeld, H., 2002. The role of subcortical structures in human epilepsy. Epilepsy Behav. 3, 219–231.

Pinotsis, D.A., et al., 2013. Dynamic causal modelling of lateral interactions in the visual cortex. Neuroimage 66, 563–576.

Prince, D.A., Wilder, B.J., 1967. Control mechanisms in cortical epileptogenic foci*: surround inhibition. Arch. Neurol. 16, 194–202.

Qubbaj, M., Jirsa, V., 2007. Neural field dynamics with heterogeneous connection topology. Phys. Rev. Lett. 98, 238102.

Rasmussen, C., Williams, C., 2005. Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning). The MIT Press, Cambridge MA.

Sanner, R., Slotine, J., 1992. Gaussian networks for direct adaptive control. IEEE Trans. Neural Netw. 3, 837–863.

Sanz-Leon, Paula, 2015. Mathematical framework for large-scale brain network modeling in The Virtual Brain. NeuroImage 111, 385–430.

Scerri, K., Dewar, M., Kadirkamanathan, V., 2009. Estimation and model selection for an IDE-based spatio-temporal model. IEEE Trans. Signal Process. 57, 482–492.

Schevon, C.A., Weiss, S.A., McKhann Jr., G., Goodman, R.R., Yuste, R., Emerson, R.G., Trevelyan, A.J., 2012. Evidence of an inhibitory restraint of seizure activity in humans. Nat. Commun. 3, 1060.

Schiff, S.J., 2011. Neural Control Engineering: The Emerging Intersection Between Control Theory and Neuroscience. MIT Press.

Schiff, S., Sauer, T., 2008. Kalman filter control of a model of spatiotemporal cortical dynamics. J. Neural Eng. 5, 1–8.

Sedigh-Sarvestani, M., Schiff, S.J., Gluckman, B.J., 2012. Reconstructing mammalian sleep dynamics with data assimilation. PLoS Comput. Biol. 8, e1002788.

Stephan, K., Kasper, L., Harrison, L., Daunizeau, J., den Ouden, H., Breakspear, M., Friston, K., 2008. Nonlinear dynamic causal models for fMRI. Neuroimage 42, 649–662.

Truccolo, W., Donoghue, J., Hochberg, L., Eskandar, E., Madsen, J., Anderson, W., Brown, E., Halgren, E., Cash, S., 2011. Single-neuron dynamics in human focal epilepsy. Nat. Neurosci. 14, 635–641.

Turner, B.M., Forstmann, B.U., Wagenmakers, E.J., Brown, S.D., Sederberg, P.B., Steyvers, M., 2013. A Bayesian framework for simultaneously modeling neural and behavioral data. Neuroimage 72, 193–206.

Ullah, G., Schiff, S., 2010. Assimilating seizure dynamics. PLoS Comput. Biol. 6, 35–45.

Van Rotterdam, A., Lopes da Silva, F., Van den Ende, J., Viergever, M., Hermans, A., 1982. A model of the spatial-temporal characteristics of the alpha rhythm. Bull. Math. Biol. 44, 283–305.

Wendling, F., 2005. Neurocomputational models in the study of epileptic phenomena. J. Clin. Neurophysiol. 22, 285.

Wendling, F., Hernandez, A., Bellanger, J.J., Chauvel, P., Bartolomei, F., 2005. Interictal to ictal transition in human temporal lobe epilepsy: insights from a computational model of intracerebral EEG. J. Clin. Neurophysiol. 22, 343.

Yarlagadda, R., 2009. Analog and Digital Signals and Systems. Springer Verlag.

Yellott, J.I., 1982. Spectral analysis of spatial sampling by photoreceptors: topological disorder prevents aliasing. Vis. Res. 22, 1205–1210.