Journal of Neuroscience Methods 308 (2018) 48–61

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

# A procedure to increase the power of Granger-causal analysis through temporal smoothing

E. Spencer[a], L.-E. Martinet[b], E.N. Eskandar[b,c], C.J. Chu[b], E.D. Kolaczyk[d], S.S. Cash[b], U.T. Eden[d], M.A. Kramer[d,*]

[a] Graduate Program in Neuroscience, Boston University, United States
[b] Department of Neurology, Massachusetts General Hospital, United States
[c] Department of Neurological Surgery, Albert Einstein College of Medicine, Montefiore Medical Center, United States
[d] Department of Mathematics and Statistics, Boston University, United States

ARTICLE INFO

*Keywords:*
Functional connectivity
Granger causality
Autoregressive modeling
Time series analysis

ABSTRACT

*Background:* How the human brain coordinates network activity to support cognition and behavior remains poorly understood. New high-resolution recording modalities facilitate a more detailed understanding of the human brain network. Several approaches have been proposed to infer functional networks, indicating the transient coordination of activity between brain regions, from neural time series. One category of approach is based on statistical modeling of time series recorded from multiple sensors (e.g., multivariate Granger causality). However, fitting such models remains computationally challenging as the history structure may be long in neural activity, requiring many model parameters to fully capture the dynamics.

*New Method:* We develop a method based on Granger causality that makes the assumption that the history dependence varies smoothly. We fit multivariate autoregressive models such that the coefficients of the lagged history terms are smooth functions. We do so by modelling the history terms with a lower dimensional spline basis, which requires many fewer parameters than the standard approach and increases the statistical power of the model.

*Results:* We show that this procedure allows accurate estimation of brain dynamics and functional networks in simulations and examples of brain voltage activity recorded from a patient with pharmacoresistant epilepsy.

*Comparison with Existing Method:* The proposed method has more statistical power than the Granger method for networks of signals that exhibit extended and smooth history dependencies.

*Conclusions:* The proposed tool permits conditional inference of functional networks from many brain regions with extended history dependence, furthering the applicability of Granger causality to brain network science.

## 1. Introduction

The human brain consists of a vast network of interacting elements. Understanding how these elements interact to support cognition, behavior and perception remains a fundamental challenge in neuroscience. One approach to address this challenge is through the analysis of the brain's anatomical networks and functional connectivity - how separate brain regions interact via transient coordination of activity (de Pasquale et al., 2010; Park and Friston, 2013). Anatomical and functional networks, summarized via tools from network analysis (Bassett and Sporns, 2017; Bullmore and Sporns, 2009) provide insight into cognition (Braun et al., 2015; de Pasquale et al., 2010; Kabbara et al., 2017; Petersen and Sporns, 2015; Telesford et al., 2016), learning (Bassett et al., 2015, 2011; Singer, 1993) and neurological disorders such as autism, stroke, schizophrenia, and epilepsy (De Vico Fallani et al., 2014; Kramer and Cash, 2012; Lynall et al., 2010; Matlis et al., 2015).

There exist many ways to estimate functional connectivity, inspired – in part – by theories of how brain regions communicate. One of the most established theories posits that oscillatory neuronal activity supports communication between brain regions (Bastos and Schoffelen, 2016; Bonnefond et al., 2017; Fries, 2015; Uhlhaas et al., 2010). This theory motivates the application of many functional connectivity measures that characterize coupling between the phase and/or amplitude of rhythmic brain signals (Bastos and Schoffelen, 2016; Greenblatt et al., 2012; Lachaux et al., 1999). Other popular approaches for

* Corresponding author at: Department of Mathematics and Statistics, 111 Cummington Mall, Boston, MA, 02215, United States.
*E-mail address:* mak@bu.edu (M.A. Kramer).

https://doi.org/10.1016/j.jneumeth.2018.07.010
Received 26 April 2018; Received in revised form 6 July 2018; Accepted 14 July 2018
Available online 19 July 2018

measuring functional connectivity - although less physically motivated - have been adopted from other scientific fields. One of the most popular methods is conditional Granger causality, which provides a direct quantification of how much the history of one brain area can predict the activity of another (Granger, 1969; Seth, 2010). There are many advantages to the Granger causality approach. First, it can be used to infer the direction of information flow, also known as directed functional connectivity (Barrett et al., 2012; Ding et al., 2006). Second, it is a model-based approach that is rooted in stochastic process theory (Bastos and Schoffelen, 2016; Cohen, 2014). Third, when the models are conditioned on the whole network and thus use all of the observed data, Granger causality limits the impact of indirect coupling - an important confound in functional network inference (Seth, 2010).

While Granger causality has been successfully applied to analyze multivariate neural activity (Seth et al., 2015), implementation can be computationally difficult on large brain networks (Seth, 2010; Valdés-Sosa et al., 2005). Two features contribute to this computational difficulty: the number of observed brain regions, i.e. nodes, and the duration for brain signal transmission. Modern recordings now support observations from hundreds, or even thousands, of sensors (Jäckel et al., 2017; Viventi et al., 2011). Interhemispheric communication can range between 5–300 ms depending on the myelination of the fiber commissures (Ringo et al., 1994), whereas intrahemispheric communication ranges between 5–10 ms (de Pasquale et al., 2010; Smith et al., 2011). Fitting Granger causality models to data with many nodes and long history dependence generates computational challenges due to the large number of parameters to estimate. Because brain signals are highly nonstationary, the amount of data available to fit large brain networks is typically limited (Barnett and Seth, 2014; Cohen, 2014), making networks inferred via conditional Granger causality subject to overfitting (Seth, 2010) and highly underpowered (Kelley and Maxwell, 2003). Furthermore, the more parameters included in the model, the longer the computation time required to estimate the model and deduce the functional network.

Common approaches to address this computational challenge rely on reducing the model size, for example, by only performing analysis on a subset of nodes, by downsampling the time series which removes the high frequency content from the signal, or by using smaller model orders chosen via parsimonious model selection techniques (Barnett and Seth, 2014; Barrett et al., 2012; Seth, 2010). However, these approaches are not always desirable. For instance, a common approach to infer larger networks is to only implement Granger causality pairwise on nodes (Seth, 2010), which may result in networks confounded by indirect influences (Bastos and Schoffelen, 2016; Ding et al., 2006); conditioning the models on the entire observed network reduces the effects of this confound (Ding et al., 2006). Another approach is to limit the size of the history dependence included in the model. Determining the optimal model order for brain signals is a difficult, but important, problem since over- and underspecified models are not informative (Cohen, 2014; Seth, 2010). It has been proposed that the history dependence of brain signals is on the order of tens to hundreds of milliseconds (Barrett et al., 2012; Kabbara et al., 2017; Ringo et al., 1994); using a smaller model order than the true history dependence may result in a poor representation of the brain functional network (Bressler and Seth, 2011). It has additionally been shown that low model orders do not necessarily capture all complexities in the signal spectrum (Bressler and Seth, 2011; Cohen, 2014). To infer accurate and informative functional networks, it is important to incorporate all nodes that are relevant to the particular phenomenon in question, and to include the appropriate history dependence.

In this paper, we propose a method to reduce the number of parameters needed for model estimation in the Granger causality framework, thus permitting inference of larger brain networks with longer history dependence. Existing approaches involve priors assuming that network connections are sparse (Seth, 2010; Valdés-Sosa et al., 2005). Here, instead of imposing assumptions on the structure of the network, we impose assumptions on the shape of the history dependence between nodes, i.e. the coefficients of the lagged time indices of the history dependent autoregressive (AR) model. The assumption is suitable for brain signals, but is less general than the classic AR model. We express the model coefficients in a spline basis, which imposes a smooth structure on the coefficients of the history dependence. In doing so, we assume the coefficients change smoothly and gradually from one time point to the next. We refer to this model as the *spline AR model*, and when applied to determine functional connectivity, the *spline-Granger method.* These serve as extensions, respectively, of what we will refer to as the traditional *standard AR model* and *the standard-Granger method.*

If the spline AR model is in fact an appropriate representation of the signal, there are many benefits to using this method. First, expressing the data in a spline basis reduces the dimensionality of the model, without removing data (i.e. downsampling) and without ignoring potentially critical nodes in the system (i.e. fitting only pairwise on nodes). Also, by using fewer parameters to estimate the same number of observations, we have more data to fit each parameter and thus can expect more statistical power, in the case that the true history dependence of the signal is well represented by a smooth function. That is, we are more precise in our coefficient estimates and are better able to detect small, nonzero effects between nodes. Thus, the spline-Granger method has more statistical power to detect true connections between nodes. Fewer parameters make computing large functional networks feasible; an appropriate ratio of parameters to observations - given stationarity constraints - can be achieved, resulting in more certainty in our model inference (Kelley and Maxwell, 2003). Therefore, the spline-Granger procedure permits inference of more precise and representative networks that cover a larger spatial and temporal scale.

In what follows we compare the performance and accuracy of the spline AR model to the standard AR model on simulations and real data. Through single-node simulations and in vivo recordings, we show that the spline AR model can reconstruct signals with accuracy comparable to the standard AR model, while the spline AR model requires fewer parameters to do so. Then, through simulation of nine-node networks, we show how appropriate choice of model order, or amount of history dependence included, improves the accuracy of the networks inferred. Finally, we implement both measures on a 26-node network of cortical data recorded preceding a seizure. We show that, for the 26-node networks, the spline-Granger method successfully infers functional networks for a smaller time window due to estimating fewer parameters and providing greater confidence in the inferences. We show that, because we have more statistical power and precision in estimating the model coefficients when fitting the spline AR models, the spline-Granger method is more sensitive to detecting true edges, or true positives, between nodes while preserving the same false positives.

## 2. Methods

### 2.1. Implementation of the standard AR model and standard-Granger method

We employ the traditional conditional Granger causality measure for directed functional connectivity, which we refer to as the *standard-Granger method* (Barnett and Seth, 2014). This method is a model-based approach to determine if the activity in one brain region - or node - drives activity in another brain region. If the history of activity at node *A* significantly reduces the amount of unexplained variance in the model of activity at another node *B*, then we conclude that node *A* has predictive power over - or drives - node *B* (Granger, 1969). For example, consider inference of the directed functional connectivity of a three-node network comprised of signals $x_t$, $y_t$, $z_t$, where $t$ indicates a discrete time index. To test the hypothesis that $y_t$ drives $x_t$, we build two autoregressive (AR) models of the activity of $x_t$: a full model including the history of all nodes in the network, and a nested (or restricted) model that includes the history of all nodes except that of $y_t$. We will

refer to these models as the *standard AR model*, and the equations are, respectively:

$$x_t = x^T_{(t-1:t-p)}\beta_{xx} + y^T_{(t-1:t-p)}\beta_{xy} + z^T_{(t-1:t-p)}\beta_{xz} + \varepsilon_{x,t}\,, \tag{1a}$$

$$x_t = x^T_{(t-1:t-p)}\beta_{xx} + z^T_{(t-1:t-p)}\beta_{xz} + \tilde{\varepsilon}_{x,t}\,. \tag{1b}$$

In these equations, $x_t$, $y_t$ and $z_t$ are column vectors containing the history dependence of their respective signals, i.e. for a model of order $p$, $x^T_{(t-1:t-p)} = [x_{t-1}...x_{t-p}]$, where $T$ indicates transpose, and $\beta_{xx}$, $\beta_{xy}$, $\beta_{xz}$ are column vectors containing the corresponding coefficients of the history dependence, i.e. $\beta_{xx} = [\beta_{xx,1}...\beta_{xx,p}]^T$.

Choosing the best model order $p$ is a challenging task as there are several statistically guided techniques, such as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), which may yield different recommendations. Furthermore, different time windows of data and different pairs of electrodes may also yield different recommendations. Because model order can significantly affect results, we follow the recommendation to fix $p$ across all electrode pairs and time windows (Cohen, 2014). We also choose the same $p$ for both full and nested models as is conventionally done when estimating causality in neuroscience research (Barnett and Seth, 2014; Bressler and Seth, 2011; Ding et al., 2006; Granger, 1969; Karalis et al., 2016; Luo et al., 2012; Mitra et al., 2018; Place et al., 2016; Schmitt et al., 2017; Seth, 2010; Smith et al., 2015; Uddin et al., 2011; Zagha et al., 2015).We use the *fitglm* function in MATLAB to determine the model coefficients and compute AIC.

The *standard-Granger method* refers to the procedure for determining a functional connectivity network from multiple signals by fitting the full and nested *standard AR models* in (1) for every signal pair. To determine if the proportion of variance explained by introducing the variables associated with $y_t$ in (1a) is significant, we construct an F-test, comparing the residuals of the full and nested models (Barnett and Seth, 2014). The null hypothesis is that the regions are not functionally connected, or that, for our trivariate example, $y_t$ does not significantly improve model residuals. The F-statistic is defined as follows:

$$F = \frac{(\mathrm{RSS}_{restricted} - \mathrm{RSS}_{full})/p}{\mathrm{RSS}_{full}/(N - k^*p)}\,,$$

where $\mathrm{RSS}_{restricted}$ and $\mathrm{RSS}_{full}$ are the residual sum of squared errors for the restricted (1b) and full models (1a), respectively. The F-statistic accounts for the number of free parameters in each model; $p$ is the model order, $k$ is the number of nodes in the network, and $N$ is the number of observations used to fit the models. The F-statistic is compared to an F-distribution with parameters $d_1 = p$ and $d_2 = N - k^*p$.

Both equations in (1) are conditioned on the history of $z_t$ to remove the possibility of a spurious result in the case that $z_t$ acts to confound the relationship between $x_t$ and $y_t$, often referred to as the common input problem (Bastos and Schoffelen, 2016). That is, suppose $z_t$ has predictive power for both $x_t$ and $y_t$, and that $x_t$ and $y_t$ are conditionally independent. In this scenario, if we did not include the history of $z_t$ in either of the models, we would likely conclude that $y_t$ drives $x_t$, or $x_t$ drives $y_t$, due to the indirect effect of both signals being correlated with $z_t$. However, by conditioning each model on the rest of the network, here by including the history of $z_t$, we would correctly identify no direct interaction between$x_t$ and $y_t$. This example illustrates the importance of conditioning autoregressive models on the network as a whole rather than computing only pairwise comparisons.

In what follows, we compute the F-statistic and its associated p-value for every pair of nodes in the network. When doing so, we use the False-Discovery-Rate (FDR) to control for multiple testing, with q = 0.05 (i.e. an upper bound of 5% on the expected proportion of false positives among all declared edges in the inferred network) (Benjamini and Hochberg, 1995).

To estimate confidence bounds for the model coefficients (e.g. $\beta_{xx}$ in (1)) we employ a bootstrapping procedure. To do so, we use the observed coefficient estimates and their estimated covariance to generate 10,000 normally distributed samples of the coefficients. From the resulting distribution, we determine the 0.025 and 0.975 quantiles of estimated model coefficients. In this way, we use the surrogate distribution to define the 95% confidence interval for the history dependence estimates.

### 2.2. Implementation of the spline AR model and the spline-Granger method

The method we develop in this paper is a modified version of the standard-Granger method, which we refer to as the *spline-Granger method.* The difference between the two methods is that, in the spline-Granger method, we rewrite the standard AR model (1), such that the coefficients of the lagged variables are written in a Cardinal spline basis (Hearn and Baker, 1996). This spline basis fits a smooth curve to the data via local interpolation. It does so by estimating third-degree polynomials between specified points, called knots, in the data such that connections at the knots are smooth. To estimate a third-degree polynomial between two knots, we require that the tangent line at each knot is determined by the slope of the surrounding points whose shape is controlled by a tension parameter, $\tau$, which we set to 0.5. That is, if we are estimating a polynomial, $f(x)$, between the two discretely indexed knots at $x = u_2$ and $x = u_3$ from the set of knots $u_1, u_2, ..., u_\ell$, then we impose: $f'(u_2) = \tau(f(u_3) - f(u_1))$ and $f'(u_3) = \tau(f(u_4) - f(u_2))$. The resulting estimated curves are continuous and have continuous first derivatives. By using a spline basis to estimate the history dependence in the Granger model, we reduce the number of model parameters to estimate from $p$ (the number of history terms in the standard-AR model (1)) to the number of knots. This method allows us to estimate directed functional connectivity with the same duration of history dependence as the standard-Granger model while reducing the number of parameters in the system.

We express the model coefficients in (1) in a spline basis via the transformation $\beta_{ij} = M\alpha_{ij}$, where $M$ is a matrix of size $p$ by $\ell$, the number of knots, and whose columns consist of the spline basis vectors. We will refer to this model as the *spline AR model* with equations:

$$x_t = x^T_{(t-1:t-p)}\, M\, \alpha_{xx} + y^T_{(t-1:t-p)}\, M\, \alpha_{xy} + z^T_{(t-1:t-p)}\, M\, \alpha_{xz} + \varepsilon_{x,t}\,, \tag{2a}$$

$$x_t = x^T_{(t-1:t-p)} M\, \alpha_{xx} + z^T_{(t-1:t-p)} M\alpha_{xz} + \tilde{\varepsilon}_{x,t}\,. \tag{2b}$$

We choose the number of knots in the system, $\ell$, to be less than $p$ in the standard-Granger model. We note that for $\ell = p$, the number of knots equals the order, and the history dependent model in the spline basis and standard basis are equivalent.

Many alternative approaches exist to reduce the number of parameters required in a multivariate autoregressive model of neural activity. Common approaches include downsampling the time series data (Cohen, 2014; Seth et al., 2015, 2013) or using regularization techniques (Smith et al., 2011; Valdés-Sosa et al., 2005) which makes the assumption that network connectivity is sparse. Both downsampling and an assumption of sparse network connectivity are inappropriate for some types of neural data, for example seizure activity, during which high frequency rhythms (Frauscher et al., 2017) and dense functional network connectivity are common (Burns et al., 2014; Jiruska et al., 2013; Kramer and Cash, 2012; Schindler et al., 2007). In addition, interhemispheric interactions may occur with prolonged delays (de Pasquale et al., 2010; Ringo et al., 1994), and developing a model that omits these delayed interactions may result in inaccurate functional networks (Bressler and Seth, 2011).

In what follows, we develop an approach to estimate multivariate autoregressive models of neural activity with prolonged history dependence and many nodes, yet maintain a relatively small number of parameters. Many continuous systems can be well approximated with low-order AR models because the system exerts itself instantaneously. However, in neural systems, there exist delays between the field signals

that are influenced by the neural mechanisms themselves, such as the neural tissue and myelination of the fibers. We suspect that higher-order AR models are needed to fully capture these delays. Therefore, we choose to use a spline basis to aid in capturing the effects at longer lags, as previously used in the point process framework, where splines have been shown to efficiently capture the effects of a spike on a point process through some delay (Deng et al., 2013, Eden et al., 2012). The spline basis helps us achieve the goal of maintaining the spatial scale, i.e. the size of the network, while extending our temporal scale, i.e. including longer lags."

To fit the spline AR model equations, we must first choose the location of the knots. Although procedures exists to select knot locations, an optimal knot placement is difficult to determine (Dimatteo et al., 2001). We choose here to place knots uniformly spaced every five indices starting at zero. When the sampling rate of the data is 500 Hz, this corresponds to a knot every 10 ms. Because there is missing information at the first and last boundary knots, we include one (invisible) knot at −200 ms before the first time index, and we require that the first derivative of the spline function be zero at the last knot. That is, at the last knot, $x = u_\ell$, we fix $f'(u_\ell) = 0$. In doing so, we assume that the history dependence does not change dramatically at long delays. In Fig. 1, we plot an example set of basis functions for an example history dependence spanning 60 ms (sampling rate 500 Hz). We show the eight fitted basis functions (for knots at −200 ms, 0 ms, 10 ms, 20 ms, 30 ms, 40 ms, 50 ms, and 60 ms).

In what follows, we refer to the *spline-Granger method* as the procedure for determining a functional connectivity network from multiple signals by fitting the full and nested *spline AR models* (2) for every signal pair, and applying the same F-test as described for the standard-Granger method. We estimate the confidence bounds on the coefficients via the same bootstrapping procedure as described for the standard-Granger method.

### 2.3. Goodness-of-fit: Grenander and Rosenblatt test

As one measure of the model goodness-of-fit, we compare the spectra of the estimated model signals to the observed signal. To do so, we implement a modification of the Grenander and Rosenblatt test of the integrated spectrum, as described in (Priestley, 1981). This test compares the true cumulative spectrum of the signal, $H(\omega)$, with the cumulative spectrum of a new realization of the estimated model $\hat{H}(\omega)$.

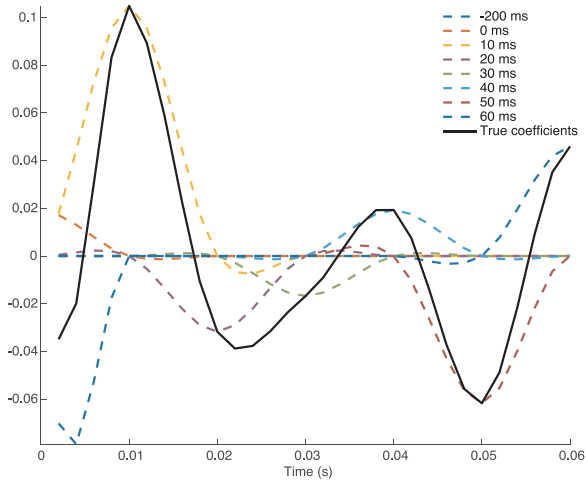


**Fig. 1.** Example approximation of history dependence with Cardinal spline basis functions. Example history dependence (black, solid) spanning 60 ms, and its approximation using eight Cardinal spline basis functions (colored, dashed) with knots at −200 ms, 0 ms, 10 ms, 20 ms, 30 ms, 40 ms, 50 ms, and 60 ms. A multiplicative factor scales each spline basis function, so that sum of all eight basis functions approximates the history dependence.

Because we are most interested in observations that consist of a single, noisy realization of a process, we replace the true cumulative smoothed spectrum with the spectrum estimated from the observed data. In this way, the test compares the observed spectrum with the spectrum simulated from the model. We estimate the spectrum using the multitaper method with a frequency resolution of 2 Hz. Doing so reduces the variability of the spectral estimates, compared to other approaches (Bokil et al., 2010). The 95% confidence intervals for the cumulative spectra are computed using the following two equations:

$$\hat{H}(\omega) - a\sqrt{\frac{8\pi\hat{G}(\pi)}{N}} \le H(\omega) \le \hat{H}(\omega) + a\sqrt{\frac{8\pi\hat{G}(\pi)}{N}}\,,$$

$$\hat{G}(\pi) = \frac{1}{4\pi}\sum_{s=-(N-1)}^{N-1}\hat{R}^2(s)\,,$$

where $N$ is the number of samples in the signal from which the spectrum is estimated, $\hat{R}(s)$ is the sample autocovariance function, and $a = 2.2414$ specifies the 95% confidence interval. Grenander and Rosenblatt define the statistic: $k_{GR} = \max_\omega \sqrt{N}\,|\hat{H}(\omega) - H(\omega)|$, which here is the weighted absolute difference between the observed cumulative spectrum and the estimated model spectrum. The model is considered a good fit with 95% confidence if the value of the statistic $k_{GR}$ is less than 2.2414 (Priestley, 1981). We determine the p-value associated with this statistic using the table in (Grenander and Rosenblatt, 1984), and correct for multiple comparisons using the False Discovery Rate with a significance level of 0.05 (Benjamini and Hochberg, 1995).

### 2.4. Goodness-of-fit: Durbin-Watson test

As a second measure of the model goodness-of-fit, we analyze the model residuals by computing the Durbin-Watson statistic. The Durbin-Watson test checks for serial correlation of the model residuals, as described in (Durbin and Watson, 1950). The Durbin-Watson statistic is defined as:

$$k_{dw} = \frac{\sum_{t=2}^{N}(e_t - e_{t-1})^2}{\sum_{t=1}^{N} e_t^2}\,,$$

where $e_t$ is the residual value at time $t$, and $N$ is the number of observations. To compute a p-value for this statistic, we use the approximation method described in (Durbin and Watson, 1950) and commonly implemented, such as in the Multivariate Granger Causality Toolbox (Barnett and Seth, 2014), and correct for multiple comparisons using the False Discovery Rate with a significance level of 0.05 (Benjamini and Hochberg, 1995).

### 2.5. Generation of synthetic signals: single-node simulations

We generate a signal with a long history dependent structure to examine the performance of the standard AR and spline AR models. To generate the signal, we use an autoregressive model of order 20, AR (20):

$$x_t = x_{(t-1:t-20)}^T\beta_{xx} + \varepsilon_{x,t}\,, \tag{4}$$

where we choose the coefficients $\beta_{xx}$ to create a signal dominated by high frequency activity, $\beta_{xx} = [-0.023, 0.100, 0.050, -0.160, -0.170, -0.160, -0.123, -0.086, -0.008, 0.056, 0.083, 0.079, 0.056, 0.027, 0.005, 0.002, 0.003, 0.013, 0.021, 0.019]^T$. We note that these coefficients establish a relatively smooth history dependence (for example, see Fig. 3c). The last term in the model, $\varepsilon_{x,t}$, is a normal random variable (mean = 0, variance = 0.0625) and $t$ corresponds to a sample every $1/f$ of a second, where $f$ is the sampling frequency. In all simulations, we set the sampling frequency to 500 Hz. To fit the models, we simulate 8 s of data, and analyze the last 2 s (1000 samples) of data to avoid the effects of initial transients. In autoregressive modeling, it is

required that the signals are weakly stationary, meaning that the mean and variance do not change over time (Cohen, 2014). To assess the stationarity of our simulated data, we apply to each simulation the KPSS test implemented in the Multivariate Granger Causality toolbox (Barnett and Seth, 2014; Seth et al., 2015). All simulations passed these tests at significance level $\alpha = 0.05$.

To examine the impact of downsampling the signal as an alternative means of reducing the number of parameters, we first simulate a signal at 500 Hz as described above, and then downsample this signal using the MATLAB function decimate, which first low-pass filters the data to prevent aliasing, and then downsamples the signal. We downsample by a factor of 5 such that the new sampling rate is 100 Hz.

### 2.6. Generation of synthetic signals: nine-node simulations

To simulate activity in the nine-node networks, we implement a multivariate autoregressive (MVAR) model of order 30. The model for the nine signals is:

$$\boldsymbol{u_t} = \sum_{i=1}^{30} \boldsymbol{B_i}\, \boldsymbol{u_{t-i}} + \boldsymbol{\varepsilon_{u,t}}\ , \tag{5}$$

where $\boldsymbol{u_t}$ is a $9 \times 1$ vector representing the value for each signal at time $t$, $\boldsymbol{B}$ is a multidimensional array with dimensions $9 \times 9$ x $p$, and $\boldsymbol{\varepsilon_{u,t}}$ is a $9 \times 1$ vector of normal random variables (mean = 0, variance = 0.0625). $B_i$ contains the model coefficients at time $t-i$, where the $n^{th}$, $m^{th}$ entry, $B_i(n, m)$, is the coefficient of influence of signal $m$ at time $t-i$ on signal $n$ at time $t$. To generate $B_{1:30}(n, m)$ we first define two smooth functions (black lines in Fig. 2) as references for the history dependence of nodes connecting to themselves, $n = m$ (self history; Fig. 2a) and for nodes connected to other nodes, $n \neq m$ (cross history; Fig. 2b). Then, for each $B$ coefficient in the network (i.e. for all $n$, $m$, and $i$), we add random noise (mean = 0, variance = 0.01) to each value to create different (noisy) functions for the history dependencies. The resulting values of $B_i(n, m)$ approximate the smooth functions, but are jagged (i.e. not smooth). Example history dependencies for the self-history $B_{1:p}(n, n)$ and cross-history $B_{1:p}(n, m)$, with $n \neq m$ are plotted in Fig. 2. In this way, we disadvantage the spline-Granger method because the true history dependence violates the spline AR model assumption that the history dependence is smooth.

We implement two simulations with this MVAR model. In the first, we fix the estimated model order $p = 30$, simulate 8 s of data, and analyze the last 2 s of data to avoid the effects of initial transients. In the second simulation, we fit two models with different model orders, $p = 5$ and $p = 30$, and analyze 2 s, 4 s and 8 s of data after simulating an initial 6 s of data to avoid initial transients. In all simulations, we set the sampling frequency to 500 Hz. All simulations passed the KPSS test for stationarity (Barnett and Seth, 2014).

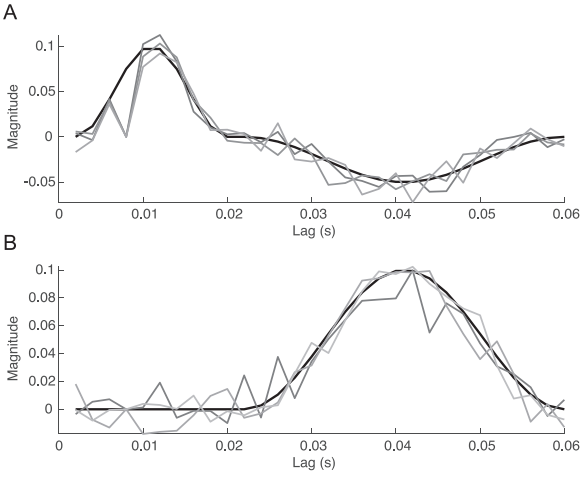


**Fig. 2.** Illustration of model coefficients for network simulations. Plots of example functions used for history dependence in the nine-node simulations. The network model coefficients (gray curves, three examples shown) were created by adding noise to the smooth functions (thick black curves) for the self-influence terms (a), and cross-influence terms (b).

The single-node and nine-node simulations (include values of $\boldsymbol{B_i}$) are provided for reuse and further development at the repository: https://github.com/Mark-Kramer/spline-granger-causality.

### 2.7. Calculation of network accuracy

We represent the functional networks inferred from the signals as binary matrices in which 1 defines a Granger-caused connection (or edge) between two nodes and 0 defines a non-edge, i.e. where there is no evidence for functional connectivity between two nodes. We define the accuracy as the proportion of correctly identified edges, i.e. true positives (TP), and non-edges, i.e. true negatives (TN):

$$\frac{TP + TN}{N^2},$$

where $N^2$ is the total number of edges in the network. We note that here we allow self-connections (i.e. an edge from a node to itself) which represent significant self-history dependence. We also note that computing the accuracy requires knowledge of where edges exist, and therefore only applies to simulated data.

### 2.8. Calculation of computation time

All simulations were run on a standard personal computer (64-bit Mac OX 10.10.5, two Quad-Core Intel Xeon 2.27 GHz CPUs, 32 GB memory). Computation time was calculated using the *tic* and *toc* functions and averaged over multiple simulation trials (as described below) in MATLAB version 8.6.0.

### 2.9. In vivo recordings from a human subject

One patient (male age 45 years) with medically intractable focal epilepsy underwent clinically indicated intracranial electroencephalogram (ECoG) recordings for epilepsy monitoring. The recordings were performed using a standard clinical recording system (XLTEK, subsidiary of Natus Medical) with a 500 Hz sampling rate. A two-dimensional subdural electrode array grid (Adtech Medical) was placed on the pia to confirm the hypothesized seizure focus, and locate epileptogenic tissue in relation to essential cortex, thus directing surgical treatment. The reference electrode was a strip of electrodes placed outside the dura and facing the skull at a region remote from the electrode array grid. Clinical electrode implantation, positioning, duration of recordings and medication schedules were based purely on clinical need as judged by an independent team of physicians without reference to this research. Analysis of these data was performed retrospectively under protocols monitored by the local Institutional Review Boards according to NIH guidelines.

In what follows, we analyze a 10 s interval of these data selected to occur before seizure onset. Before applying the Granger methods, we filter these data to remove 60 Hz line noise (Butterworth filter, order 2, with passband [59,61] Hz). All signals analyzed passed the KPSS test for stationarity (Barnett and Seth, 2014).

### 2.10. Simulation and model fitting code

The single-node and nine-node simulations (including values of $\beta_{xx}$ in Eq. (4) and $\mathbf{B_i}$ in Eq. (5)) and code for implementing both the spline-Granger and standard-Granger network inference procedures are provided for reuse and further development at the repository: https://github.com/Mark-Kramer/spline-granger-causality.

## 3. Results

We begin with simulations to examine the performance of two approaches to history dependent modelling, one in which the model history is estimated in a standard basis, and another in which the history is estimated in a Cardinal spline basis. We show that, in both cases, the spline AR model accurately estimates the history dependence of the simulated signal, with more certainty than the standard AR model, and fits the data well. Next, through simulation of a nine-node network, we show that the spline-Granger network inference procedure accurately infers the functional connectivity, and has more statistical power to detect true edges. Finally, we apply the model fitting and network inference procedures to an example multi-electrode voltage recording from human cortex. We show that both the spline-Granger and standard-Granger procedures infer similar functional connectivity networks, while the spline-Granger method can be implemented on shorter duration datasets and has more power to detect edges. These results illustrate the accuracy and efficiency of the proposed spline-Granger procedure.

### 3.1. Simulation: single node

We begin by examining the performance of the spline AR model on two seconds of activity simulated from a single node. To construct the simulated signal, we implement an autoregressive (AR) model of order 20, corresponding to 40 ms, (see Methods; Eq. (4)) with activity dominated by frequencies in the 15–20 Hz range (Fig. 3a,b). From this simulated signal, we then estimate history dependent autoregressive models, with history dependence expressed in the spline and standard bases, and include up to 60 ms of history. We choose the estimated history dependence to exceed the true dependence because, in practice, estimating model order from noisy neural data is not straightforward, and a common approach is to choose a model order between 5 and 100 ms (Barrett et al., 2012; de Pasquale et al., 2010; Smith et al., 2011). We estimate the parameters of the history dependent AR model in two ways. First, we estimate the coefficients $\beta_{xx}$ of the standard AR model at each integer lag, in the univariate case where $\beta_{xy}$ and $\beta_{xz}$ are zero vectors (Eq. (1a)) (Barnett and Seth, 2014). For 60 ms of history, this corresponds to 30 parameters to estimate where one coefficient is estimated for every lag. Second, we transform the history of the data into a spline basis and estimate the coefficients $\alpha_{xx}$ of the spline AR model, in the univariate case where $\alpha_{xy}$ and $\alpha_{xz}$ are zero vectors (Eq. (2a)). For 60 ms of history, this corresponds to 8 parameters in the spline basis where coefficients are estimated for each of the 8 basis functions (see example of basis functions in Fig. 1). Both the standard and the spline AR models estimate 60 ms of history dependence, but the spline AR model does so with fewer parameters.

We find that both models accurately estimate the true coefficients used to generate the simulated data (Fig. 3c). For both methods, the coefficient estimates approximate the true model coefficients. However, those estimated in the standard AR model lack the smooth dependence that exists in the true model coefficients used to generate the signal, and have large confidence bounds that frequently include 0, meaning there is no evidence of influence at that lag. Visual inspection of Fig. 3c reveals that the bounds of the estimate from the spline AR model are much tighter, and frequently exclude 0; by using fewer parameters and fitting on the same amount of data, we increase the amount of data used to estimate each parameter. Thus, we increase our effective degrees of freedom, gain statistical power to detect nonzero influences, and gain more certainty in the parameter estimates. We note that the spline AR model accurately captures the non-zero influence of history for broad intervals of time near 0.01 s and 0.03 s, while the standard AR model fails to do so (Fig. 3c). At larger delays beyond the true model order, the estimated coefficients of both models fluctuate around zero.

To investigate further the performance of both models, we calculate two complementary goodness-of-fit measures that evaluate different aspects of the model's agreement with the data. First, we compare estimates of the integrated spectrum calculated from the simulated signal, and estimated from the two history dependent models (Fig. 3d, see Methods). Visual inspection of Fig. 3d suggests that both the standard AR model and the spline AR model produce signals with similar spectra to the observed data spectrum. Consistent with this observation, we calculate the Grenander and Rosenblatt test statistic ($k_{GR}$) and find that both models pass the Grenander and Rosenblatt test in this realization of the process. This test provides no evidence of a significant difference between the spline AR model and the observed signal ($k_{GR} = 1.21,\ p = 0.45$), and no evidence of a significant difference between the standard AR model and the observed signal ($k_{GR} = 1.89,\ p = 0.12$). Second, we analyze the model residuals and compute the Durbin-Watson test statistic ($k_{DW}$) for autocorrelation of the residuals (Fig. 3e) (Barnett and Seth, 2014). With the exception of the residual at 0 lag, we observe no evidence of significant autocorrelation in the residuals of the standard AR model ($k_{DW} = 2.00,\ p = 0.99$) or the spline AR model ($k_{DW} = 2.00,\ p = 0.99$). In this case, we find that both the spline AR model and standard AR model pass both goodness-of-fit tests.

Repeating this analysis for 1000 realizations of the simulated data and model estimates, we find consistent results. The mean and standard deviation of $k_{GR}$ for the standard AR model is 2.28 and 0.79, respectively, and for the spline AR model is 2.21 and 0.81, respectively. A good model fit with 95% confidence is indicated when $k_{GR}$ is below 2.2414 as computed in (Grenander and Rosenblatt, 1984). The standard AR model passed the GR test in 55% of the realizations, and the spline model passed the GR test in 59% of the realizations (Fig. 3f). The mean $k_{DW}$ for the standard AR model is 1.998 (95% confidence interval [1.9978, 1.9983]) and for the spline AR model is 1.993 (95% confidence interval [1.992, 1.994]). Both the standard and spline AR models passed the DW test in 100% of the realizations (see Methods; Fig. 3g) (Barnett and Seth, 2014; Seth, 2010). In addition, we find that the computation time required to estimate the standard AR model (mean 94.0 ms, 95% confidence interval [93.8, 94.2] ms) is significantly greater ($p < 10^{-15}$, t-test) than that required for the spline AR model (mean 61.0 ms, 95% confidence interval [60.6, 61.3] ms, Fig. 3h). However, we note that the reduction in computation time (approximately 33 ms) is of little practical importance for the univariate data analyzed here. We conclude that both models are rapidly estimated for these univariate data, and tend to pass the two goodness-of-fit tests in the same way.

However, we find that the standard AR and spline AR model results differ in two ways. First, for the 1000 realizations of the simulated data, we compute the AIC values for model orders up to 30 for both the spline AR model and the standard AR model (Fig. 3i). For the standard AR model, we conclude that 16 parameters minimize the AIC, corresponding to a history dependence of 32 ms, while for the spline AR model, we conclude that 4 parameters minimize the AIC, corresponding to a history dependence of 20 ms (knots placed at −200 ms, 0 ms, 10 ms, 20 ms). We conclude that, for these simulated data, the model order that minimizes the AIC is smaller in the spline AR model than the standard AR model; the spline AR model requires fewer parameters to optimize model quality. Second, for the 1000 realizations of the simulated data, we determine the width of the 95% confidence interval at 10 ms, and how often this interval correctly excludes the null value, for both models. We find that the widths of the confidence bands at 10 ms are tighter for the spline AR model (mean width 0.081, 95% confidence interval [0.075, 0.088]) than for the standard AR model (mean width 0.128, 95% confidence interval [0.123, 0.133]). The tighter confidence intervals of the spline AR model result in more accurate detection of the non-zero effect at this lag. In the spline AR model, the confidence intervals correctly exclude zero for 99.9% of the realizations, while for the standard model only 84.5% of the realizations exclude zero (Fig. 3j). We conclude that the spline AR model has more power to detect small nonzero influences in the history dependence not detected

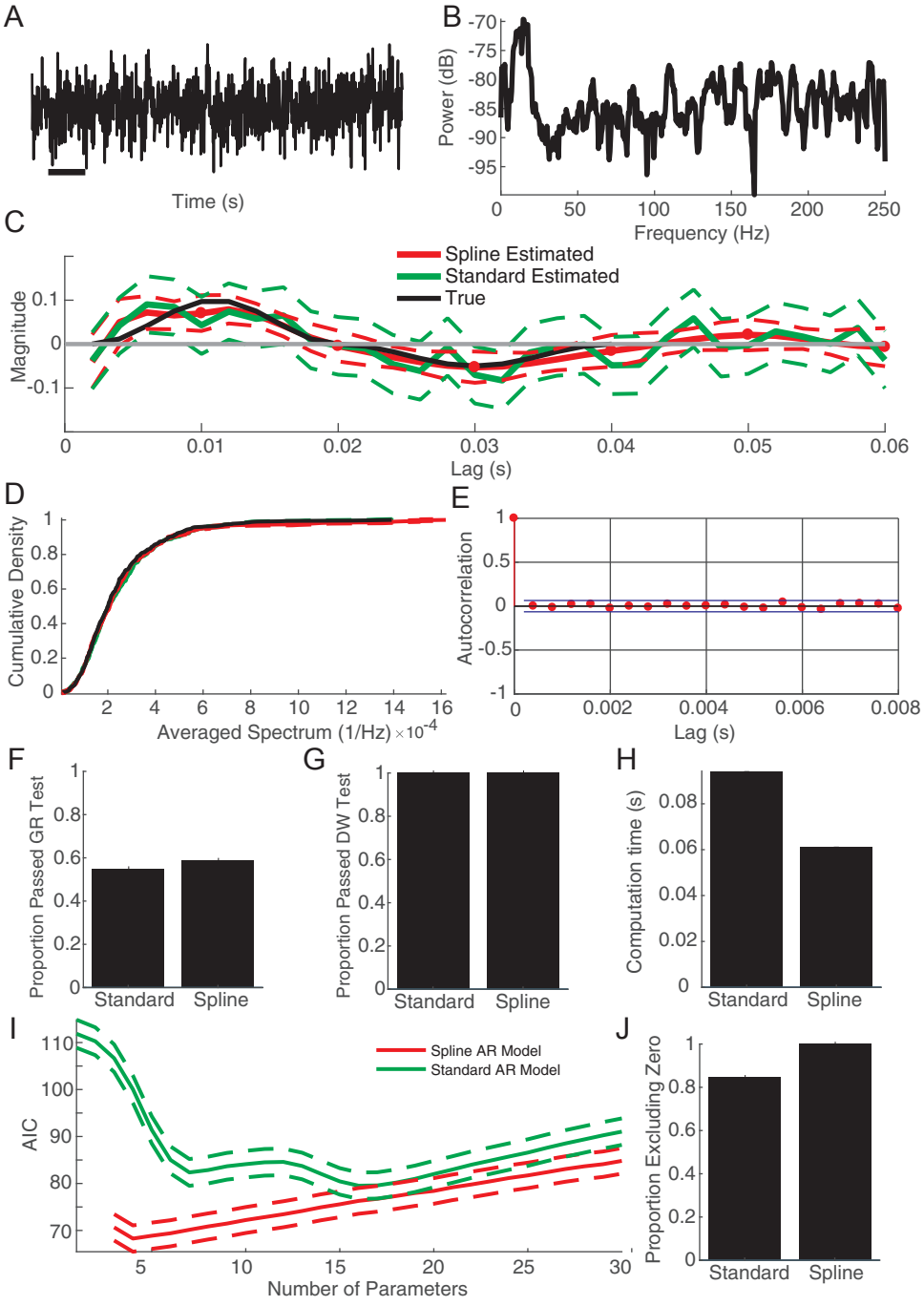


**Fig. 3.** Both modeling approaches perform well for a single node with a dominant high frequency spectral peak and long, smooth history dependence.
(a,b) Example trace (a) and spectrum (b) of the simulated signal. Rhythms near 15 Hz dominate the activity. Scale bar in (a) indicates 200 ms.
(c) The true autoregressive coefficients (black curve) estimated using the spline AR model (solid red curve, 95% confidence interval in dashed red) and using the standard AR model (green curve, 95% confidence interval in dashed green).
(d) Illustration of the Grenander and Rosenblatt goodness-of-fit test of the integrated spectrum for the data in (a). The cumulative distribution of the observed signal power (black) and both estimated averaged signals' power (standard AR in green, spline AR in red) overlap. Dashed lines indicate 95% confidence intervals.
(e) Illustration of the Durbin-Watson goodness-of-fit test for the spline AR model for the data in (a). The autocorrelation of residuals remains small for all nonzero lags.
(f–h) Population results (1000 instances) of the simulation for the (f) Grenander and Rosenblatt statistic, (g) Durbin-Watson statistic, and (h) computation time. The computation time (h) is significantly smaller for the spline AR model; error bars indicate two standard errors of the mean. For both models, approximately 60% of the simulations pass the Grenander and Rosenblatt test (f), and nearly 100% of the simulations pass the Durbin-Watson test (g).
(i) Averaged AIC values for 1000 instances of the simulation for a range of parameters for both the spline AR (red) and standard AR (green) models; solid curves indicate the mean, and dashed curves represent two standard errors of the mean.
(*j*) Population results (1000 instances) of the simulation for the exclusion of the null value at 10 ms delay. The confidence bounds at 10 ms correctly exclude the null value in approximately 80% of the simulations for the standard AR model, and nearly all simulations for the spline AR model.

by the standard AR model.

An alternative approach to reduce the number of parameters when estimating an AR model is to downsample the data (Seth et al., 2015). In some cases, downsampling may be advantageous by allowing longer history dependence with fewer model parameters. However, downsampling below the timescale of interactions between nodes hinders detection of causality, and downsampling typically involves filtering, which is generally ill-advised, as it has been to shown to obfuscate underlying causal structure (Barnett and Seth, 2014; Seth et al., 2013). The spline-Granger method with uniformly spaced knots could be considered a type of downsampling. To verify that this is not the case, we examine the standard AR model fit to a downsampled and filtered signal (see Methods). We downsample the data so that the sampling frequency is reduced from 500 Hz to 100 Hz; in the resulting signal, each sample is separated by 10 ms, which matches the interval between knots in the spline AR model (Fig. 4a). We then fit the standard AR model to the downsampled data using the same history dependence of 60 ms, as in Fig. 3. We find that the estimated coefficients for the downsampled signal provide a poor estimate of the true history of the signal (Fig. 4b). We note that, for this example, the standard AR model fails the GR test ($k_{GR}$ statistic is 7.84, $p < 10^{-15}$) and passes the DW test ($k_{DW}$ statistic is 1.83 $p = 0.22$).

Repeating this analysis of downsampled data for 1000 realizations, we find consistent results. The mean $k_{GR}$ for the standard AR model is 8.25 (95% confidence interval, [8.20, 8.30]), passing the GR test for none of the realizations. The mean $k_{DW}$ for the standard AR model is 1.86 (95% confidence interval [1.86, 1.87]), passing the DW test in 97% of the realizations (Barnett and Seth, 2014; Seth, 2010). Thus, we conclude that the standard AR model is a poor fit to the downsampled and filtered signal, and not equivalent to the spline AR model, which

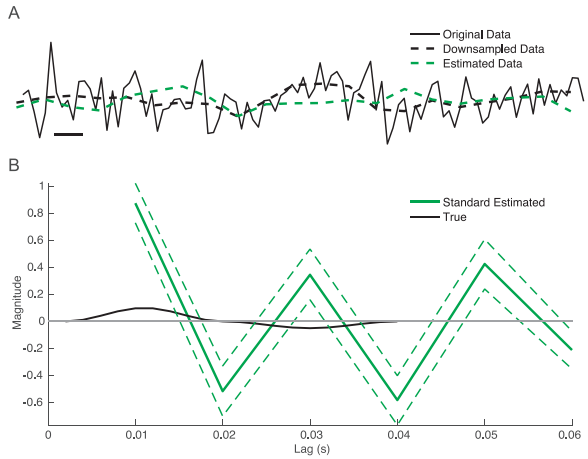


**Fig. 4.** The spline-Granger method is not equivalent to filtering and downsampling the signal.
(a) Example trace of a simulated signal (black), filtered and downsampled signal (dashed black) and signal estimated from the downsampled signal using the standard AR model (green). Scale bar indicates 10 ms.
(b) The true autoregressive coefficients (black curve), and estimated coefficients using the standard AR model (green curve, 95% confidence interval in dashed green).

passes both model fitting tests, and more accurately captures the true history dependence of the data. We note that the spline AR model acts to downsample the history dependence, not directly the signal itself.

### 3.2. Simulation: nine-node network

In the previous section, we showed that, in a single node simulation, estimation of an AR model in two ways (at each integer lag, and with a spline interpolation of history) performed similarly; both models adequately reconstructed the data (i.e., tended to pass two goodness-of-fit tests), and accurately captured the history dependence. In this section, we examine the performance of the spline-Granger and standard-Granger methods with a more complicated network of signals. To do so, we simulate a nine-node network using a multivariate autoregressive (MVAR) model with history dependent effects extending to 60 ms (see Methods). In these simulated data, the activity (Fig. 5a,b) at each node depends on its own history (up to 60 ms), and may depend on the history of activity at other nodes (again, up to 60 ms). We show a representation of the true network connectivity, defined as nonzero influence in the MVAR, in Fig. 5c. In this figure, we indicate the maximal influence (across all lags) between nodes; darker colors indicate stronger connectivity between nodes. We expect an accurate network inference to mimic this connectivity pattern.

We begin by estimating the functional connectivity given two seconds of data simulated from the nine-node network. We determine the functional connectivity in the traditional way by implementing Granger causality to assess the predictive power of each node on all nodes (including self-influences) (Barnett and Seth, 2014); see Methods. We do this in two ways: the standard-Granger causality method, and the modified spline-Granger causality method. In the standard-Granger approach, we fit AR models and estimate the history dependence at every time delay for each node, which requires estimation of a large number of parameters; each node in the nine-node network of order 30 (i.e. 60 ms history dependence) requires estimation of 270 parameters (9 nodes * 30 parameters) for the full model fit (see Methods, Eq. (1a) and (1b)). In the spline-Granger approach, we fit AR models whose coefficients are written in a lower dimensional spline basis. For this example, we use eight spline basis functions that span the specified 60 ms (see Methods, Eq. (2a) and (2b)). Doing so reduces the number of parameters to estimate for each node to 72 (9 nodes * 8 knots).

We fit both the standard AR model (Eq. (1a) and (1b) and the spline AR model (Eq. (2a) and (2b)) to the simulated data. For the example data in Fig. 5a, we find that the inferred spline-Granger networks accurately capture the true network connectivity (Fig. 5d). In this example for one realization of the network, the accuracy, or the proportion of correctly identified edges and non-edges (see Methods), is 100% for the spline-Granger network (Fig. 5d) and is 96.3% for the standard-Granger method (i.e. the inferred functional networks using the standard-Granger method correctly identify 78 of the 81 edges and non-edges in the true network). Repeating this MVAR simulation 1000 times with different noise instantiations, we find that nodes more strongly connected in the true network (Fig. 5c) appear more frequently in the inferred networks with less variability (mean and standard deviation of networks inferred across the 1000 realizations using the spline-Granger method in Fig. 5e and f, respectively). Computing the accuracy between the standard- and spline-Granger networks, we find 96.33 +/− 2.00%. This indicates that functional networks inferred using standard-Granger and spline-Granger are similar.

To examine the impact of model order on network inference, we repeat the MVAR simulation and estimation of the spline and standard AR models with history dependence up to: (1) 10 ms, corresponding to 5 parameters in the standard AR model and 3 parameters in the spline AR model, and (2) 60 ms, corresponding to 30 parameters in the standard AR model and 8 parameters in the spline AR model. We chose these two history dependencies to test the hypothesis that, if the true history dependence of a system is long, then models estimated using a longer history dependence infer more accurate networks. We compare models fit using 5 integer lags in the standard AR, a typical model order chosen for standard AR models in the literature (Barnett and Seth, 2014; Barrett et al., 2012) with models fit using 30 integer lags in the standard AR model. The latter corresponds to the true model order, but is traditionally not computationally feasible for larger-scale biological networks. The goal of these simulations is to compare how network inference is affected by the amount of history dependence included in the model. In both cases, we infer functional networks from simulated data of duration 2 s, 4 s, and 8 s. For each combination of model order and data duration, we simulate 1000 instances of the MVAR process, and compare the accuracy of the inferred networks and computation time of both the spline-Granger and standard-Granger methods.

We find that both the spline-Granger and standard-Granger methods infer more accurate functional networks when the model order is higher and closer to the true model order for all durations of data analyzed, as expected. For both models, the accuracy is higher when estimating 60 ms of history compared to 10 ms of history (compare pink and gray bars, Fig. 5g,h). We also observe that, for the longer history dependence of 60 ms, the accuracy increases as the duration of data analyzed increases (compare gray bars in Fig. 5g,h); additional data provides more evidence to estimate the additional parameters in the model with 60 ms of history. We conclude that the most accurate estimation of the functional network occurs when the model order is large and consistent with the simulated system, and the duration of data analyzed is large.

For the network fitting scenarios with less observed data and more parameters to estimate (i.e. 2 s of data, and a model history dependence of 60 ms), the spline-Granger method is significantly more accurate ($p < 10^{-15}$, *t*-test; Fig. 5g,h), capturing 98.69% (mean for 1000 simulations, 95% confidence interval [0.9861, 0.9878]) of the true network, while the standard-Granger captures 96.96% (mean for 1000 simulations, 95% confidence interval intervals [0.9686, 0.9706]). We note that the spline-Granger method produces more accurate functional networks, despite the fact that the true simulated history dependence was not smooth. This is because the spline AR model has more statistical power to correctly infer edges, i.e. correctly reject the null hypothesis that there is no correlation between two nodes. The difference is more pronounced when there is less available data to estimate the model parameters.

To illustrate further the increased statistical power of the spline AR

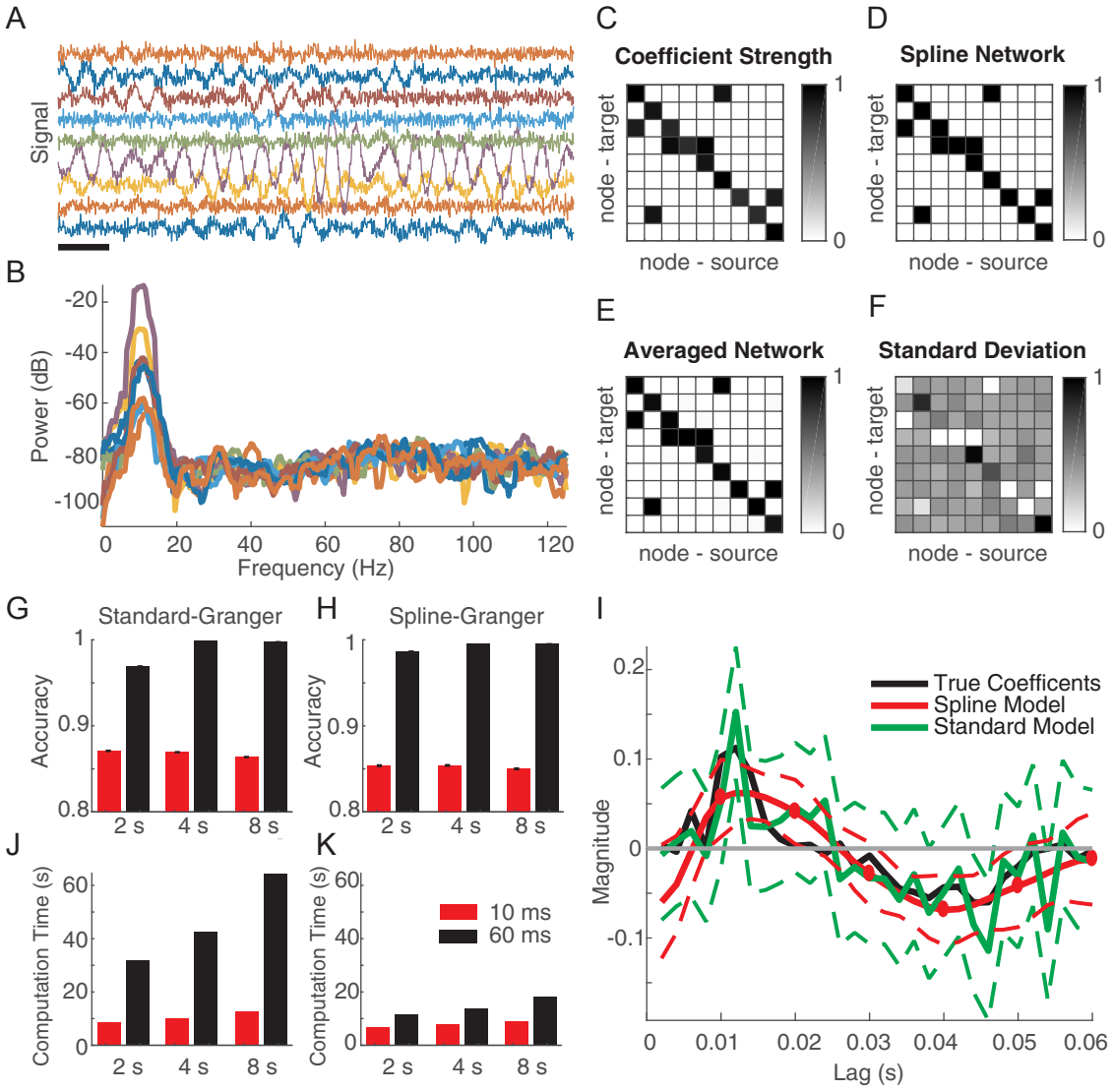


**Fig. 5.** Increases in estimated model order produce accurate functional networks in minimal computation time using the spline-Granger method.
(a,b,c) Example simulated signals (a) and spectra (b) of the nine-node network with the connectivity strength between nodes shown in (c). Scale bar in (a) indicates 200 ms.
(d,e,f) Corresponding networks inferred using the spline-Granger method (d), the average network estimated from the spline-Granger method with history dependence 60 ms (e), and the variability of those estimates (f) for 1000 realizations of the simulated network data. In each figure, darker shades indicate larger values.
(g,h) Accuracy for 1000 realizations of the nine-node network for the standard-Granger (g) and spline-Granger (h) models with history dependence extending to 10 ms (red) and 60 ms (gray), and data of duration 2 s, 4 s, and 8 s.
(i) An example of the estimated history dependence for which the standard-Granger method fails to detect an edge, while the spline-Granger method detects an edge. Inferred coefficients for standard AR model (green) and spline AR model (red) and true model coefficients (black); solid curves indicate the mean, and dashed curves indicate the 95% confidence intervals.
(*j*,k) Computation time for 1000 realizations of the nine-node network for the standard-Granger (*j*) and spline-Granger (k) models with history dependence extending to 10 ms (red) and 60 ms (gray), and data of duration 2 s, 4 s, and 8 s.

model to detect edges, we compare the estimated coefficients from the standard AR and the spline AR models for an example node pair at which the two methods produced different results; the standard-Granger method failed to detect an edge, while the spline-Granger method correctly detected the edge (Fig. 5i). In this case, consistent with the single node simulation (Fig. 3c), the confidence bounds on the coefficients of the standard AR model fluctuate around zero and are much larger than those of the spline AR model. Thus, there is not enough evidence to reject the null hypothesis in the standard-Granger method, while in the spline-Granger method the decreased uncertainty in the model coefficients correctly identifies the nonzero effects, resulting in an edge between the two nodes.

For both methods, the computation time increases as the model order and duration of data analyzed increase (Fig. 5j,k). However, the computation time for the spline-Granger method is less than that required for the standard-Granger method for all combinations of model order and data duration considered ($p < 10^{-15}$ in all cases). We note that this difference is most pronounced when the model order and data duration are large; i.e. in the scenario of the most accurate network inference. Specifically, for fitting models with 60 ms of history using 8 s of data, the standard-Granger method takes 72% longer to compute on average compared to the spline-Granger method.

This example highlights the utility of the spline-Granger approach; the method permits more accurate network inference because (1) we can fit models with longer history dependencies – which are more reflective of neural systems – with fewer parameters, and (2) we gain

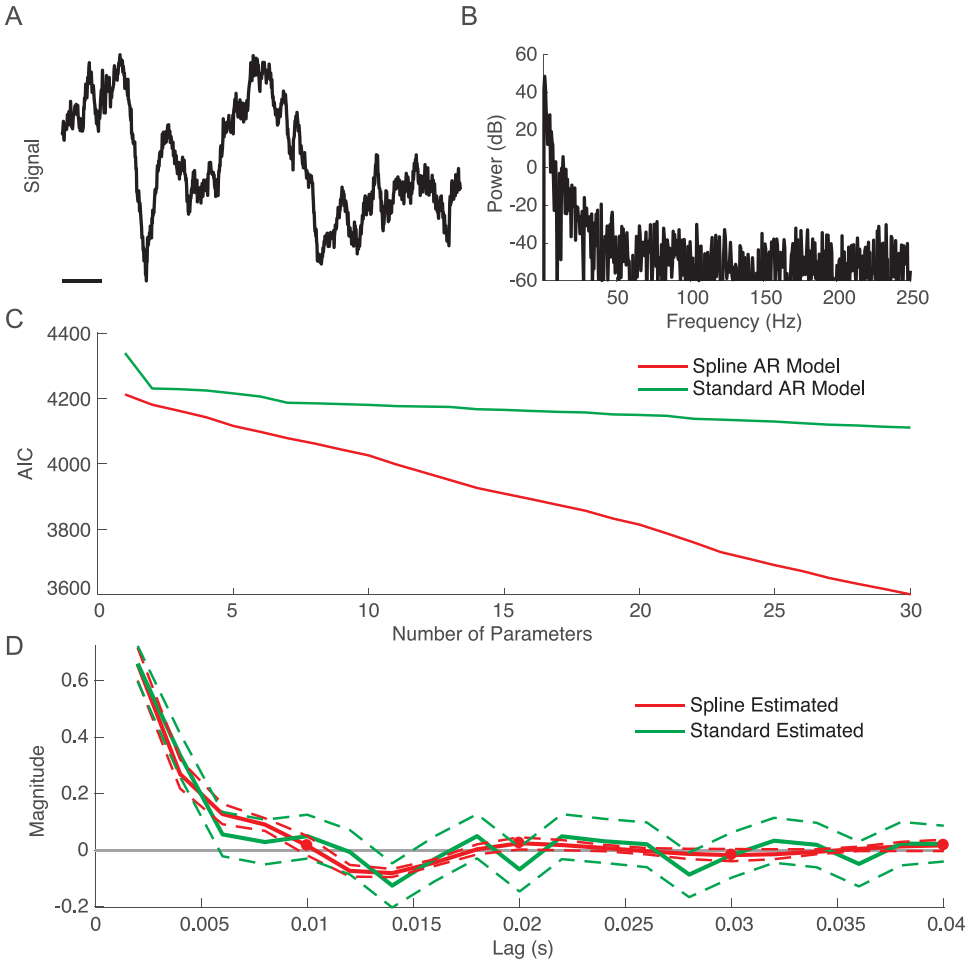


**Fig. 6.** The spline AR and standard AR models produce similar results for in vivo data recorded from a single node.
(a,b) Example signal (a) and spectrum (b) recorded from human cortex preceding a seizure. Black bar in (a) indicates 200 ms.
(c) AIC computed for both the spline AR and standard AR models up to order 30.
(d) Fits to the history dependence using the spline AR model (red) and standard AR model (green); solid curves indicate the mean, and dashed curves indicate the 95% confidence intervals.

more statistical power to correctly identify edges.

### 3.3. In vivo data: single node

To examine the performance of the spline AR model on physiological data, we first consider in vivo brain voltage activity recorded from a single electrode. We show in Fig. 6a an example voltage trace from human cortex at a time preceding seizure onset. The power spectrum of these data exhibits a 1/f structure, common in recordings of brain voltage activity (Fig. 6b) (He et al., 2010). Although the true history dependence is unknown for these in vivo data, we may still compare the performance of the spline AR model with the standard AR model. First, we compute the AIC values for both models (Fig. 6c). We find for model orders up to 30, no absolute minimum occurs, as expected for 1/f signals which theoretically require infinite model order (Tang et al., 2017). Because AIC is not an informative way to select model order for brain signals of this type, and other model selection procedures tend to produce variable results (Cohen, 2014), we choose to fit the univariate spline AR and standard AR models with history dependence up to 40 ms, which is within the range of history dependencies reported in brain signals (Barrett et al., 2012; de Pasquale et al., 2010; Smith et al., 2011). In the standard AR model this corresponds to estimating 20 parameters at each integer lag, and in the spline AR model this corresponds to estimating 6 parameters for the coefficients of 6 spline basis functions (with knots at −200 ms, 0 ms, 10 ms, 20 ms, 30 ms, and 40 ms). The coefficients inferred for both models are similar (Fig. 6d), but the coefficients for the spline AR model are smoother, as expected. Consistent with the single node simulations, we find tighter confidence bounds around the estimated coefficients for the spline AR model, and significant effects in the history dependence. The estimated coefficients of the standard AR model fluctuate around zero after 7 ms, whereas in the spline AR model, we find weak but significant nonzero effects up to 20 ms. The ability to detect these weak effects results from the gain in statistical power in the spline-Granger model, which uses the same amount of data to estimate fewer parameters. The $k_{DW}$ statistics for the standard AR and the spline AR models are 2.00 ($p = 0.97$) and 2.00 ($p = 0.96$), respectively, indicating no evidence for autocorrelation of the model residuals. We find that both the standard AR and the spline AR models fail the GR test ($p < 10^{-15}$) in both cases the models fail to capture all of the observed signal power at low frequencies. Comparing the structure of the model coefficients and the model residuals, we conclude that both methods similarly reconstruct the data, although the spline-Granger method provides more statistical power to detect weak – but significant – effects.

### 3.4. In vivo data: multiple nodes

Finally, we apply both network inference methods to in vivo brain data recorded from a 26-electrode array. We chose to use only 26 signals of an 8 × 8 electrocorticography grid because while still a

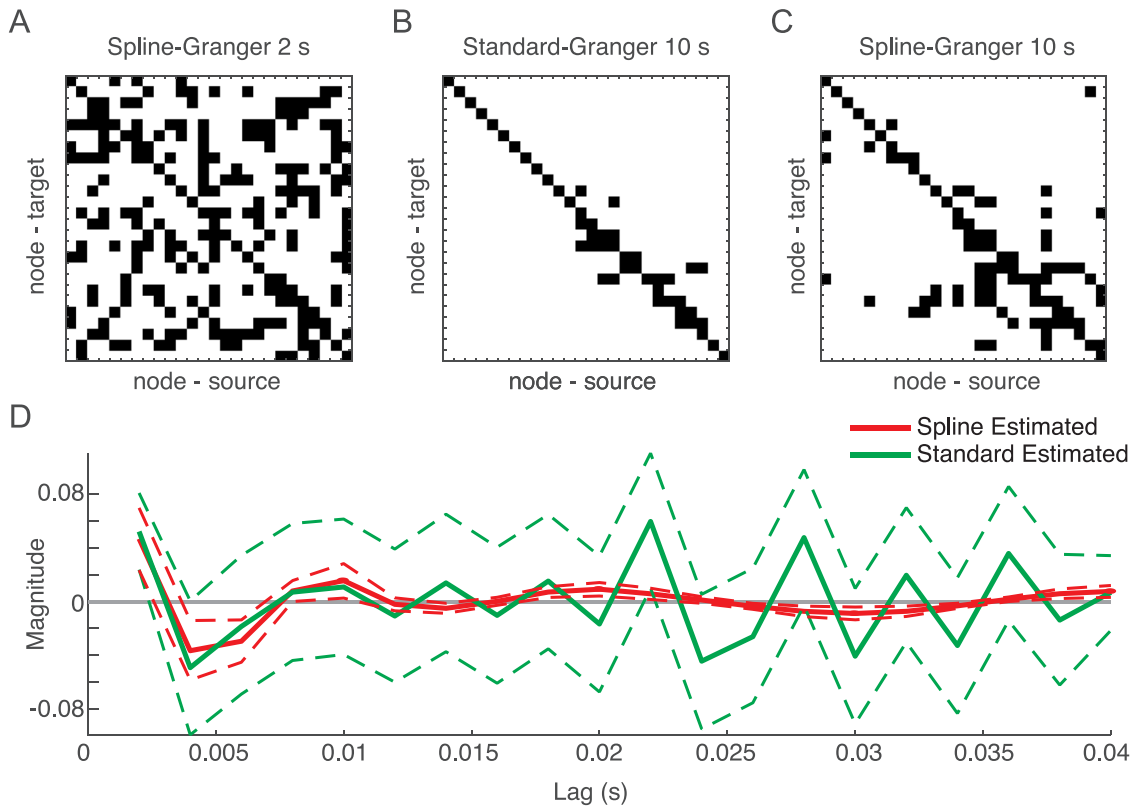


**Fig. 7.** The spline-Granger method makes inference of a 26-node network computationally tractable and more sensitive.
(a) Network inferred applying the spline-Granger method to 26 voltage recordings of duration two seconds. The standard-Granger method requires estimation of too many parameters given the amount of data observed.
(b,c) Networks inferred using standard-Granger (b) and spline-Granger (c) methods applied to voltage recordings from 26 electrodes of duration ten seconds.
(d) Estimated history dependence for a signal pair for which a connection was detected using the spline-Granger method (solid red curve, 95% confidence interval in dashed red), but not using the standard-Granger method (solid green curve, 95% confidence interval in dashed green).

sizable number of signals, it is small enough to fit, and thus compare, both methods. Typically, computing networks of this scale using the standard-Granger approach is infeasible due to the limited amount of stationary data available relative to the number of parameters required for model inference. To demonstrate the utility of the spline-Granger approach, we fit models using 2 s and 10 s of data, and a model history dependence up to 40 ms. For 2 s of data, estimation of the standard AR model is not appropriate; the number of parameters to estimate for each node (26 nodes * 20 lags = 520) is nearly half the number of observations per node (2 s * 500 Hz −20 lags = 980 observations). In general, a good model fit requires a ratio of the number of observations to the number of parameters much greater than 1 (Harrell, 2001). For the standard AR model, this ratio is 1.9, and too small to support confident results. However, because the spline AR model requires estimation of only 156 parameters for each node (26 nodes * 6 knots), this ratio is 6.3, and large enough to perform the model fit. We note that estimating the functional network using the spline-Granger method requires 2.90 min of computation time (Fig. 7a).

While, ideally, we would not fit AR models on 10 s of data due to the nonstationary nature of brain signals, we do so here as a means of comparison for both network inference procedures; despite expecting the underlying brain system generating the observed activity to lack stationarity over a 10 s interval, we note that all signals analyzed passed the KPSS test for stationarity. For 10 s of data, the number of observations is sufficient to infer network structure with both methods. We find that, while the standard-Granger method requires 25.4 min to compute (Fig. 7b), the spline-Granger method requires 5.93 min (Fig. 7c). Excluding the self-influence terms, both methods infer similar functional networks; the overlap in edges and non-edges is 93.1%. Additionally, we note that the spline-Granger method infers more edges than the standard-Granger method (compare Fig. 7b and c). As we illustrated previously, we compare the estimated coefficients for an instance in which the standard-Granger method fails to detect an edge and the spline-Granger method detects an edge (Fig. 7d). Although both methods detect a similar mean effect, the 95% confidence bounds on the estimated coefficients for the spline AR model are tighter than those for the standard AR model, consistent with the simulation results from the nine-node network (Fig. 5i). Thus, the influence between the two signals is nonzero, and we infer that the two nodes are functionally connected. We conclude that the reduction in parameters when using spline-Granger method permits larger functional networks to be fit using shorter durations of data, often required when analyzing nonstationary data, and that it has more statistical power to detect edges.

## 4. Discussion

Granger causality is a powerful approach to infer functional networks from multi-sensor recordings of brain activity. However, the standard-Granger approach requires estimation of many model parameters, limiting its applicability to small or otherwise restricted networks. To address this limitation, we developed a modification of the standard-Granger approach by assuming the model coefficients of the time lagged history dependent terms are smooth functions. This assumption reduces the number of model parameters to estimate and makes inference of functional networks from stationary intervals of multi-sensor recordings computationally tractable.

In univariate and multivariate simulations, we showed that the spline AR model accurately reconstructs the signals and closely estimates the true model parameters, and that the spline-Granger method accurately infers functional network structure. We also showed that

when the true history dependence of a system is long (e.g., 60 ms, consistent with brain signals), accuracy of functional network inference improves when models approximate the true history dependence of the system. We applied the proposed method to in vivo brain data and found that a univariate brain signal can be modeled by history dependencies that extend to 40 ms, and that multivariate functional networks can be inferred using the spline-Granger method on shorter intervals of data. Additionally, in the simulations and in vivo brain data, we found that using the same amount of data to estimate fewer parameters in the spline-Granger approach produced more precise parameter estimates, i.e. smaller confidence bounds around the parameter estimates, and ultimately detection of weak – but present – connections in the functional networks.

Two challenges limit the applicability of the standard-Granger approach to large brain networks. First, the computation time required for the standard-Granger model rapidly increases with network size. Second, brain signals are highly nonstationary which limits the amount of data and leads to statistically underpowered models. The problem of computation time may be addressed through brute force approaches given powerful computers and efficient algorithms. The problem of statistically underpowered models may be addressed through appropriate choice of model. Combining too many model parameters with too few observations leads to Type II errors, meaning many true functional connections between nodes will be missed (Kelley and Maxwell, 2003). By assuming that the history dependence of the model is smooth, we express the MVAR models in a lower dimensional basis, thus reducing the number of parameters to estimate. Compared to the standard AR model, the spline AR model proposed here reduces the number of parameters by a factor of approximately 5, allowing reliable models estimates from shorter data intervals, in the scenario that our model assumptions are appropriate, and reducing computation time.

Because Granger methods are model-based, there are many ways to assess goodness-of-fit and confidence in the model results. Since different goodness-of-fit tests assess different aspects of model performance, we chose to implement two tests here. The Durbin-Watson statistic analyzes the model fit in the time domain by checking for serial correlations in the model residuals of the reconstructed signal. The Grenander-Rosenblatt statistic analyzes the model fit in the frequency domain. This test is derived from stochastic process theory and compares a known spectrum to a model estimated spectrum. Ultimately, this test assesses the model's ability to generate signals with appropriate spectral properties. We chose these two methods because the Durbin-Watson test is commonly used in MVAR modeling (Barnett and Seth, 2014) and because the Grenander-Rosenblatt test assess the spectrum - an important characteristic of brain data (Buzsáki et al., 2004).

In the case of real data, there is no ground truth to assess which model performs more accurately. Therefore, we test our models on simulated datasets that are multivariate and include an extended timescale of interactions between nodes, consistent with in vivo brain activity. Motivated by these simulated results, and our analysis of the ECoG data, we speculate that the spline-Granger method supports more accurate and representative functional networks. By allowing estimation simultaneously from more brain regions, the method permits inference of larger functional networks, i.e. networks with more brain regions or nodes, and therefore accounts for more indirect interactions. And, by permitting estimation with longer history dependence, the method captures longer scale brain dynamics. The fundamental modification in this model compared to the standard-Granger approach is the smooth history dependent structure. The benefit of the spline AR model is that spline functions are sufficiently flexible to reflect variations in the history dependence while using fewer parameters than the standard AR model. We emphasize that neither method is more correct, but that the spline-Granger method can operate on more brain regions with smaller intervals of data than the standard-Granger method. In the case that the model assumptions are met, the effective degrees of freedom in the spline AR model are greater, giving the model more statistical power to detect weak interactions. However, we note that this assumption is poor when the history dependence changes quickly, such that the history dependence is not smooth, and that this assumption is not particularly useful when the true history dependence is short.

While we cannot know the true history dependence that drives observed neural signals, nor necessarily interpret the Granger coefficients as a representation of the underlying mechanisms (Barrett and Barnett, 2013), there exist two observations that motivate an assumption of smoothness for the model coefficients. First, comparing the standard AR to the spline AR model coefficients fit to in vivo data (Figs. 6 and 7), the standard AR coefficients fluctuate around the spline AR coefficients. While these fluctuations may contain meaningful information, we hypothesize that – instead – the rapid fluctuations of the standard AR coefficients represent non-informative noise. By smoothing these rapid fluctuations, the spline AR model both reduces the impact of this noise and requires less parameters to estimate. Second, in the nine node simulations, we designed a system such that critical information appeared at longer lags, and the history dependencies between signals were not smooth. Yet, despite an incorrect assumption of smoothness, the spline-Granger model accurately captured the network structure. In this case, the omission of rapid fluctuations in the history dependence did not impact network inference. Importantly, because the spline-Granger model required fewer parameters, this approach also permits more accurate network inference on a smaller duration of data compared to the standard-Granger model. Ultimately, our analysis suggests that the spline basis acts as a reasonable approximation for the system.

In our model selection procedure, there remain three challenges: determining the model order, the choice of knot placement, and the number of knots. Common approaches to determining model order, such as AIC, BIC, and partial autocorrelation functions, aim to select the most parsimonious model. However, parsimony may not be desirable; in brain signals, lags up to 100 ms can drive the network (Barrett et al., 2012; Smith et al., 2011; Tang et al., 2017) and models that omit these influences may miss important features. We also note that AIC in particular may be a poor choice for brain data because neural signals are dominated by pink noise, which requires an infinite model order (Tang et al., 2017). Determining the optimal model order for a multivariate system, rather than a univariate signal, is an even more challenging task. For example, the optimal model order may differ for each signal (Cohen, 2014). Because of the smoothness induced by the spline basis, small fluctuations in the history dependence near zero are damped to zero, consistent with a weak neural influence. We speculate that the spline bases could be used *a posteriori* to determine model order as the last non-zero term in the history dependence.

The transformation into a spline basis also depends heavily on the choice of knot placement and the number of basis functions. We chose to place the knots uniformly every 10 ms to standardize our models across our analysis and to limit how much prior information we impose on the models. Other work has explored principled approaches, including Bayesian techniques, to determine the optimal number and location of splines (Dimatteo et al., 2001), which may potentially yield more accurate models.

Continuing research to reduce the number of parameters and choose appropriate models that reflect the system in these approaches remain important goals to accurately infer functional networks. Ideally, studies are designed to detect features with enough power to obtain confidence intervals - of any size - that correctly exclude the null value. However, when the feature values are nonzero but small, model estimates need to be more precise, i.e. the confidence intervals need to be small to correctly exclude the null value. Thus, not only is it important to design studies that have enough power to determine the significance of parameters, but also to apply methods that support more precise parameter estimates. By obtaining tighter confidence bounds, we achieve more accurate parameter estimates and thus obtain more power to correctly identify weak, but present, functional connections between nodes (Kelley and Maxwell, 2003). Approaches exist to design studies that

achieve both a desired statistical power level and a desired confidence band width, such as accuracy in parameter estimation analysis (Kelley and Maxwell, 2003). Knowing the optimal number of parameters to achieve desired power and precision might provide insight when choosing the number of nodes and knots to include in spline-Granger network models. For instance, if the optimal number of parameters can be predetermined, these parameters could then be distributed uniformly across all node pairs in the network.

In conclusion, the spline-Granger method provides a flexible and useful tool for network inference of large models. Because accurate network inference is limited by the stationarity of the data available, we develop an approach that reduces the number of parameters to estimate. The number of parameters in the model is a function of the number of brain regions included and the amount of history dependence estimated between these brain regions. To account for the possible confounding influences of indirect interactions between brain regions, we develop an approach that does not reduce the number of brain areas, but instead simplifies the estimated history dependence. We showed that we gain statistical power and precision in our parameters estimates at the cost of making assumptions on how the influence of past values changes over time. By fitting more statistically powerful models with more nodes and longer history dependence, the method can produce more precise and more informative functional networks.

## Acknowledgements

This work was supported by the National Science FoundationDMS #1451384, and the National Institutes of Health NINDS R01NS095369.

## References

Barnett, L., Seth, A.K., 2014. The MVGC multivariate Granger causality toolbox: a new approach to Granger-causal inference. J. Neurosci. Methods 223, 50–68. https://doi.org/10.1016/j.jneumeth.2013.10.018.

Barrett, A.B., Barnett, L., 2013. Granger causality is designed to measure effect, not mechanism. Front. Neuroinform. 7, 6–7. https://doi.org/10.3389/fninf.2013.00006.

Barrett, A.B., Murphy, M., Bruno, M.A., Noirhomme, Q., Boly, M., Laureys, S., Seth, A.K., 2012. Granger causality analysis of steady-state electroencephalographic signals during propofol-induced anaesthesia. PLoS One 7. https://doi.org/10.1371/journal.pone.0029072.

Bassett, D.S., Sporns, O., 2017. Network neuroscience. Nat. Neurosci. 20, 353–364. https://doi.org/10.1038/nn.4502.

Bassett, D.S., Wymbs, N.F., Porter, M.A., Mucha, P.J., Carlson, J.M., Grafton, S.T., 2011. Dynamic reconfiguration of human brain networks during learning. Proc. Natl. Acad. Sci. 108, 7641–7646.

Bassett, D.S., Yang, M., Wymbs, N.F., Grafton, S.T., 2015. Learning-induced autonomy of sensorimotor systems. Nat. Neurosci. 18, 744–751. https://doi.org/10.1038/nn.3993.

Bastos, A.M., Schoffelen, J.-M., 2016. A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. Front. Syst. Neurosci. 9, 1–23. https://doi.org/10.3389/fnsys.2015.00175.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B 289–300. https://doi.org/10.2307/2346101.

Bokil, H., Andrews, P., Kulkarni, J.E., Mehta, S., Mitra, P.P., 2010. Chronux: a platform for analyzing neural signals. J. Neurosci. Methods 192, 146–151. https://doi.org/10.1016/j.jneumeth.2010.06.020.

Bonnefond, M., Kastner, S., Jensen, O., 2017. Communication between brain areas based on nested oscillations. eNeuro 4https://doi.org/10.1523/ENEURO.0153-16.2017. ENEURO.0153-16.2017.

Braun, U., Schäfer, A., Walter, H., Erk, S., Romanczuk-Seiferth, N., Haddad, L., Schweiger, J.I., Grimm, O., Heinz, A., Tost, H., Meyer-Lindenberg, A., Bassett, D.S., 2015. Dynamic reconfiguration of frontal brain networks during executive cognition in humans. Proc. Natl. Acad. Sci. 112, 11678–11683. https://doi.org/10.1073/pnas.1422487112.

Bressler, S.L., Seth, A.K., 2011. Wiener-Granger causality: a well established methodology. Neuroimage 58, 323–329. https://doi.org/10.1016/j.neuroimage.2010.02.059.

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10, 186–198. https://doi.org/10.1038/nrn2575.

Burns, S.P., Santaniello, S., Yaffe, R.B., Jouny, C.C., Crone, N.E., 2014. Network dynamics of the brain and influence of the epileptic seizure onset zone. Proc. Natl. Acad. Sci. U. S. A. 111, E5321–5330. https://doi.org/10.1073/pnas.1401752111.

Buzsáki, G., Draguhn, A., Buzsaki, G., 2004. Neuronal oscillations in cortical networks. Science 304 (80), 1926–1929. https://doi.org/10.1126/science.1099745.

Cohen, M.X., 2014. Granger prediction. Analyzing Neural Time Series Data: Theory and Practice. MIT Press, pp. 371–388.

de Pasquale, F., Della Penna, S., Snyder, A.Z., Lewis, C., Mantini, D., Marzetti, L., Belardinelli, P., Ciancetta, L., Pizzella, V., Romani, G.L., Corbetta, M., 2010. Temporal dynamics of spontaneous MEG activity in brain networks. Proc. Natl. Acad. Sci. 107, 6040–6045. https://doi.org/10.1073/pnas.0913863107.

De Vico Fallani, F., Richiardi, J., Chavez, M., Achard, S., 2014. Graph analysis of functional brain networks: practical issues in translational neuroscience. Philos. Trans. R. Soc. B Biol. Sci. 369https://doi.org/10.1098/rstb.2013.0521. 20130521–20130521.

Deng, X., Eskandar, E.N., Eden, U.T., 2013. A point process approach to identifying and tracking transitions in neural spiking dynamics in the subthalamic nucleus of Parkinson's patients. Chaos 23. https://doi.org/10.1063/1.4818546.

Dimatteo, I., Genovese, C.R., Kass, R.E., 2001. Bayesian curve-fitting with free-knot splines. Biometrika 88, 1055–1071. https://doi.org/10.1093/biomet/88.4.1055.

Ding, M., Chen, Y., Bressler, S.L., 2006. Granger causality: basic theory and application to neuroscience. Handb. Time Ser. Anal. Recent Theor. Dev. Appl. 437–460. https://doi.org/10.1002/9783527609970.ch17.

Durbin, J., Watson, G.S., 1950. Trust Testing for Serial Correlation in Least Squares Regression I, vol. 37. Oxford Univ. Press, pp. 159–177.

Eden, U.T., Gale, J.T., Amirnovin, R., Eskandar, E., 2012. Characterizing the spiking dynamics of subthalamic nucleus neurons in Parkinson's disease using generalized linear models. Front. Integr. Neurosci. 6, 28. https://doi.org/10.3389/fnint.2012.00028.

Frauscher, B., Bartolomei, F., Kobayashi, K., Cimbalnik, J., van 't Klooster, M.A., Rampp, S., Otsubo, H., Höller, Y., Wu, J.Y., Asano, E., Engel, J., Kahane, P., Jacobs, J., Gotman, J., 2017. High-frequency oscillations: the state of clinical research. Epilepsia 1316–1329. https://doi.org/10.1111/epi.13829.

Fries, P., 2015. Rhythms for cognition: communication through coherence. Neuron 88, 220–235. https://doi.org/10.1016/j.neuron.2015.09.034.Rhythms.

Granger, C.W.J., 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37 (424). https://doi.org/10.2307/1912791.

Greenblatt, R.E., Pflieger, M.E., Ossadtchi, A.E., 2012. Connectivity measures applied to human brain electrophysiological data. J. Neurosci. Methods 207, 367–402. https://doi.org/10.1146/annurev-immunol-032713-120240.Microglia.

Grenander, U., Rosenblatt, M., 1984. Statistical Analysis of Stationary Time Series, 2nd ed. Chelsea Publishing Company.

Harrell, F., 2001. Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis, 2nd ed. Springer.

He, B.J., Zempel, J.M., Snyder, A.Z., Raichle, M.E., 2010. The temporal structures and functional significance of scale-free brain activity. Neuron 66, 353–369. https://doi.org/10.1086/498510.Parasitic.

Hearn, D.D., Baker, M.P., 1996. Computer Graphics: C Version, 2nd ed. Pearson Education.

Jäckel, D., Bakkum, D.J., Russell, T.L., Müller, J., Radivojevic, M., Frey, U., Franke, F., Hierlemann, A., 2017. Combination of High-density microelectrode array and patch clamp recordings to enable studies of multisynaptic integration. Sci. Rep. 7. https://doi.org/10.1038/s41598-017-00981-4.

Jiruska, P., de Curtis, M., Jefferys, J.G.R., Schevon, C.A., Schiff, S.J., Schindler, K., 2013. Synchronization and desynchronization in epilepsy: controversies and hypotheses. J. Physiol. 787–797. https://doi.org/10.1113/jphysiol.2012.239590.

Kabbara, A., El Falou, W., Khalil, M., Wendling, F., Hassan, M., 2017. The dynamic functional core network of the human brain at rest. Sci. Rep. 7, 2936. https://doi.org/10.1038/s41598-017-03420-6.

Karalis, N., Dejean, C., Chaudun, F., Khoder, S.R., Rozeske, R., Wurtz, H., Bagur, S., Benchenane, K., Sirota, A., Courtin, J., Herry, C., 2016. 4-Hz oscillations synchronize prefrontal-amygdala circuits during fear behavior. Nat. Neurosci. 19, 605–612. https://doi.org/10.1038/nn.4251.

Kelley, K., Maxwell, S.E., 2003. Sample size for multiple regression: obtaining regression coefficients that are accurate, not simply significant. Psychol. Methods 8, 305–321. https://doi.org/10.1037/1082-989X.8.3.305.

Kramer, M.A., Cash, S.S., 2012. Epilepsy as a disorder of cortical network organization. Neuroscientist 18, 360–372. https://doi.org/10.1177/1073858411422754.

Lachaux, J.P., Rodriguez, E., Martinerie, J., Varela, F.J., 1999. Measuring phase synchrony in brain signals. Hum. Brain Mapp. 8, 194–208. https://doi.org/10.1002/(SICI)1097-0193(1999)8:4<194::AID-HBM4>3.0.CO;2-C.

Luo, C., Guo, Z., wei, Lai, Y., xiu, Liao, W., Liu, Q., Kendrick, K.M., Yao, D., zhong, Li, H., 2012. Musical training induces functional plasticity in perceptual and motor networks: insights from resting-state fMRI. PLoS One 7. https://doi.org/10.1371/journal.pone.0036568.

Lynall, M.-E., Bassett, D.S., Kerwin, R., McKenna, P.J., Kitzbichler, M., Muller, U., Bullmore, E., 2010. Functional connectivity and brain networks in schizophrenia. J. Neurosci. 30, 9477–9487. https://doi.org/10.1523/JNEUROSCI.0333-10.2010.

Matlis, S., Boric, K., Chu, C.J., Kramer, M.A., 2015. Robust disruptions in electroencephalogram cortical oscillations and large-scale functional networks in autism. BMC Neurol. 15, 1–17. https://doi.org/10.1186/s12883-015-0355-8.

Mitra, A., Kraft, A., Wright, P., Acland, B., Snyder, A.Z., Rosenthal, Z., Czerniewski, L., Bauer, A., Snyder, L., Culver, J., Lee, J.M., Raichle, M.E., 2018. Spontaneous infra-slow brain activity has unique spatiotemporal dynamics and laminar structure. Neuron 98, 297–305. https://doi.org/10.1016/j.neuron.2018.03.015. e6.

Park, H.-J., Friston, K., 2013. Structural and functional brain networks: from connections to cognition. Science 342 (80). https://doi.org/10.1126/science.1238411.1238411–1238411.

Petersen, S.E., Sporns, O., 2015. Brain networks and cognitive architectures. Neuron 88, 207–219. https://doi.org/10.1016/j.neuron.2015.09.027.

Place, R., Farovik, A., Brockmann, M., Eichenbaum, H., 2016. Bidirectional prefrontal-

hippocampal interactions support context-guided memory. Nat. Neurosci. 19, 992–994. https://doi.org/10.1038/nn.4327.
Priestley, M.B., 1981. Spectral Analysis and Time Series. Academic Press, London.
Ringo, J.L., Don, R.W., Demeter, S., Simard, P.Y., 1994. Time is of the essence: a conjecture that hemispheric specialization arises from interhemispheric conduction delay. Cereb. Cortex 331–343.
Schindler, K., Elger, C.E., Lehnertz, K., 2007. Increasing synchronization may promote seizure termination: evidence from status epilepticus. Clin. Neurophysiol. 118, 1955–1968. https://doi.org/10.1016/j.clinph.2007.06.006.
Schmitt, L.I., Wimmer, R.D., Nakajima, M., Happ, M., Mofakham, S., Halassa, M.M., 2017. Thalamic amplification of cortical connectivity sustains attentional control. Nature 545, 219–223. https://doi.org/10.1038/nature22073.
Seth, A.K., 2010. A MATLAB toolbox for Granger causal connectivity analysis. J. Neurosci. Methods 186, 262–273. https://doi.org/10.1016/j.jneumeth.2009.11.020.
Seth, A.K., Chorley, P., Barnett, L.C., 2013. Granger causality analysis of fMRI BOLD signals is invariant to hemodynamic convolution but not downsampling. Neuroimage 65, 540–555. https://doi.org/10.1016/j.neuroimage.2012.09.049.
Seth, A.K., Barrett, A.B., Barnett, L., 2015. Granger causality analysis in neuroscience and neuroimaging. J. Neurosci. 35, 3293–3297. https://doi.org/10.1523/JNEUROSCI.4399-14.2015.
Singer, W., 1993. Synchronization of cortical activity and its putative role in and learning. Annu. Rev. Physiol. 55, 349–374. https://doi.org/10.1146/annurev.ph.55.030193.002025.
Smith, S.M., Miller, K.L., Salimi-Khorshidi, G., Webster, M., Beckmann, C.F., Nichols, T.E., Ramsey, J.D., Woolrich, M.W., 2011. Network modelling methods for FMRI. Neuroimage 54, 875–891. https://doi.org/10.1016/j.neuroimage.2010.08.063.
Smith, E.H., Banks, G.P., Mikell, C.B., Cash, S.S., Patel, S.R., Eskandar, E.N., Sheth, S.A., 2015. Frequency-dependent representation of reinforcement-related information in the human medial and lateral prefrontal cortex. J. Neurosci. 35, 15827–15836. https://doi.org/10.1523/JNEUROSCI.1864-15.2015.
Tang, W., Liu, H., Douw, L., Kramer, M.A., Eden, U.T., Hämäläinen, M.S., Stufflebeam, S.M., 2017. Dynamic connectivity modulates local activity in the core regions of the default-mode network. Proc. Natl. Acad. Sci. 114, 201702027. https://doi.org/10.1073/pnas.1702027114.
Telesford, Q.K., Lynall, M.-E., Vettel, J., Miller, M.B., Grafton, S.T., Bassett, D.S., 2016. Detection of functional brain network reconfiguration during task-driven cognitive states. Neuroimage 142, 198–210. https://doi.org/10.1016/j.neuroimage.2016.05.078.
Uddin, L.Q., Supekar, K.S., Ryali, S., Menon, V., 2011. Dynamic reconfiguration of structural and functional connectivity across core neurocognitive brain networks with development. J. Neurosci. 31, 18578–18589. https://doi.org/10.1523/JNEUROSCI.4465-11.2011.
Uhlhaas, P.J., Roux, F., Rodriguez, E., Rotarska-Jagiela, A., Singer, W., 2010. Neural synchrony and the development of cortical networks. Trends Cogn. Sci. 14, 72–80. https://doi.org/10.1016/j.tics.2009.12.002.
Valdés-Sosa, P.A., Sanchez-Bornot, J.M., Lage-Castellanos, A., Vega-Hernández, M., Bosch-Bayard, J., Melie-García, L., Canales-Rodríguez, E., Sa, J.M., Valde, P.A., Bosch-Bayard, J., Melie-garcı, L., Canales-rodrı, E., 2005. Estimating brain functional connectivity with sparse multivariate autoregression. Philos. Trans. R. Soc. Lond. B Biol. Sci. 360, 969–981. https://doi.org/10.1098/rstb.2005.1654.
Viventi, J., Kim, D.H., Vigeland, L., Frechette, E.S., Blanco, J.A., Kim, Y.S., Avrin, A.E., Tiruvadi, V.R., Hwang, S.W., Vanleer, A.C., Wulsin, D.F., Davis, K., Gelber, C.E., Palmer, L., Van Der Spiegel, J., Wu, J., Xiao, J., Huang, Y., Contreras, D., Rogers, J.A., Litt, B., 2011. Flexible, foldable, actively multiplexed, high-density electrode array for mapping brain activity in vivo. Nat. Neurosci. 14, 1599–1605. https://doi.org/10.1038/nn.2973.
Zagha, E., Ge, X., McCormick, D.A., 2015. Competing neural ensembles in motor cortex gate goal-directed motor output. Neuron 88, 565–577. https://doi.org/10.1016/j.neuron.2015.09.044.