Measuring connectivity in spike train data

Tuoxing Liu

Abstract

Understanding neural connectivity is pivotal for investigating the mechanisms underlying brain function and behavior. This review aims to provide a comprehensive introduction to the procedures of neural connectivity analysis. It explores connectivity analysis within the domain of spike train data, emphasizing the distinction between functional and effective connectivity. It highlights data acquisition methods, including Neuropixels 2.0 and tetrodes, for capturing extracellular neural activity with high spatial and temporal resolution. It evaluates dynamic causal modeling (DCM), Granger causality (GC), and transfer entropy (TE) as principal methodologies for measuring neural interactions. Each method's theoretical underpinnings, practical applications, and associated challenges, such as handling nonlinear interactions, data stationarity, and computational demands, are critically examined. This review emphasizes the importance of interpreting results within the context of their methodological constraints and explores emerging advancements that aim to enhance the reliability and applicability of these approaches. By integrating technical insights with practical considerations, this work serves as a guide for selecting appropriate methodologies and advancing future research in neural connectivity analysis using spike train data.

Introduction

Understanding the mechanisms of neural processes requires a deep exploration of connections—specifically, how discrete elements within the brain communicate and coordinate to facilitate behaviour and cognition. To study how individual neurons interact during the physiological activities of animals, it is essential to employ specific concepts to interpret the data. One such concept is functional connectivity, defined as statistical dependencies among remote neurophysiological events. Another such concept is effective connectivity, which is defined as “the influence one neuronal system exerts over another”[1].

This review will focus exclusively on connectivity observed through spike train data obtained after the processing of action potentials collected from extracellular multi-unit recordings of individual neurons. These recordings capture the action potentials which have a frequency of 300 Hz ~ 7 kHz (for a single neuron)[2], at a high temporal resolution of 30 kHz[2] and a high spatial resolution of around 0.10 mm[2] (cellular level). Note that after emitting a spike, the neuron is unable to generate another spike within about 2 ms[3, p. 2]. Non-invasive recording methods with a relatively low spatial or temporal resolution like EEG and fMRI are not included.

This review will start by briefly introducing the concepts of functional and effective connectivity and the techniques used for spike train data acquisition, including microelectrode arrays (Neuropixel, tetrodes) and optotagging, which enable studies at finer temporal and spatial resolutions. Following this, this review will introduce three key methods used to study neural interactions: Dynamical Causal Modelling, Granger Causality, and Transfer Entropy, and discuss the issues they face when dealing with spike train data.

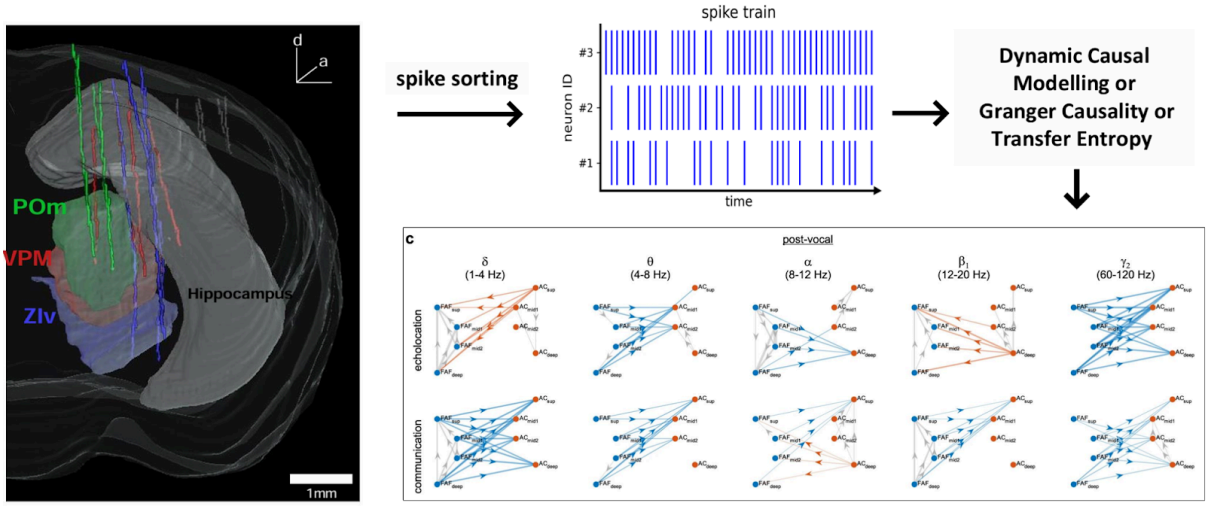

Figure 1. Flowchart of the procedures of neural connectivity analysis. First, data is acquired through multi-unit extracellular recordings. Next, spike sorting algorithms are applied to extract spike trains. Finally, connectivity measurements are used to provide a quantitative or qualitative description of the connections. The neural recording figure on the left comes from Filippo Heimburg's poster for Society for Neuroscience 2022. The graph visualization in the lower-right corner is cropped from a paper[4] licensed under Creative Commons Attribution 4.0 International License.

Functional connectivity and effective connectivity

By definition, functional connectivity refers to statistical dependencies among remote neurophysiological events. It is essentially a set of information-theoretic measures as functions of probability distributions over observed multivariate processes (e.g. Each neuron as a variable)[5]. The measures are usually model free (except Granger causality) and can be directed ($m_{X->Y} \neq m_{Y->X}$) or undirected ($m_{X->Y} = m_{Y->X}$). These include the Pearson correlation coefficient, mutual information, Granger causality, and transfer entropy.

Effective connectivity means how one neuronal system exerts another, which is essentially a generative model mapping from causes to consequences[5] (e.g. from stimuli and inputs from other neurons to activities of target neurons through hidden neurophysiological states). Typically we use differential equation models to perform this mapping for interpretability. The most prevalent approaches to effective connectivity analysis are dynamic causal modelling, structural equation modeling, and Granger causality[5].

As one might notice, Granger causality can be considered as both a measure of functional and effective connectivity. This will be explained later in the Granger causality section.

The distinction between functional and effective connectivity emerged as an attempt to disambiguate the effects of a (shared) stimulus-evoked response from those induced by neuronal connections between two units[5]. Functional connectivity is correlation-based, accounting for common inputs (except for transfer entropy). For example, regions within the Default Mode Network exhibit highly correlated activity in resting-state fMRI when a person is at rest and not focused on the external environment, although the underlying cause of this phenomenon is not specified. In contrast, effective connectivity is causality-based,

discounting influences outside the targeted neuronal systems. For example, research suggests that perception and imagination both engage neuronal representations in the same visual areas, but the differences between perception and imagery in terms of their effect on the strength of top-down and bottom-up coupling (effective connectivity) were unknown. An effective connectivity analysis revealed that the modulation of bottom-up connections (increased effective connectivity) was specific to perception, while top-down coupling was increased during both perception and imagery[6].

From my understanding, functional and effective connectivity can be seen as two names for the results of two different approaches towards the understanding of underlying neural mechanisms. If you want to use interpretable models to explain the behaviour/change of neural systems, what you will get about the connectivity is called effective connectivity. When you analyse the data without building models and only using statistics, you will get functional connectivity. However, when we want to make further inferences about the specific mechanism of the neural system, for example, the long-term impact of experimental manipulation on the coupling between particular brain regions, functional connectivity might fail because it only measures the short-term (real time) relationship.

To further explain this and similar issues faced in effective connectivity, I will use an analogy of a group of communicating computers as a server cluster. These computers cooperate through a pre-installed server framework, a software that runs on every server to provide a unified service. A service is a function that a server provides to its users. The computers use the internet to receive requests from other computers (outside server), and cooperate together to fulfill the request. In this analogy, the computers are neurons (possibly in different brain areas). The server framework is the long-term effective neural coupling we want to investigate. The messages among the computers through the internet are the spikes (considering spike train data). The network throughput (input and output throughput together) we can measure on computers is the spike rate. The service provided by the server cluster is the function of the neural network.

The effective connectivity models the server framework characterized by network throughputs at each time period and uses parameters to describe how the (past) value of all throughputs affects the current throughput at the target computer. It models a measure of the outputs of the server framework. When the task performed by the computer is simple (e.g. relaying messages), it is plausible to do so. However, when the computation is highly nonlinear, which is also the case in neural systems, this approach could encounter difficulty. For instance, researchers have determined that at least five layers in an artificial neural network (corresponding to a minimum of five nested nonlinear functions) are necessary to accurately model the input–output spikes of an L5 pyramidal cell[7]. This indicates that low-order approximations, such as the second-order method used in DCM, provide only severely limited and inaccurate results. Even if the approximation can be achieved through high-order methods, the resulting models will lack interpretability. A similar problem in computer science is called interpretable machine learning[8], which is still unsolved.

Meanwhile, the functional connectivity measures the correlation of the throughputs between these two computers during the service. When the service is changed (i.e. performing a different task), it's reasonable to see the correlation change without changing the server framework. Thus, we cannot compare the correlation of throughputs between different tasks

to infer the change of the server framework. From functional connectivity, we can only tell whether the correlation (or information transferred) among neurons of interest is changed, which does not imply that the effective neural coupling is changed. Thus, functional connectivity measures the real-time activity, not the underlying mechanism. In the original case, functional connectivity is measuring the dependency of spike rates (throughputs), not the change/dependency of the neural coupling (server framework) that causes the spike rates (throughputs). This can be a problem, but it can be bypassed by introducing a new concept for interpretation (e.g. "temporary correlation", “temporary coupling” or "temporary functional connectivity"), which means we are only describing the differences of short-term properties of the neural system instead of the change in the long-term mechanism.

According to Karl Friston’s paper in 2011[5], effective connectivity was more commonly used in basic neuroscience, whereas functional connectivity was more frequently applied in clinical and translational settings. I am analyzing the data recorded from the thalamocortical system of learning mice. It is well established that the activity of cortical layers and the thalamus is correlated, with dense bidirectional synaptic connectivity linking the two regions[9], [10], [11]. Therefore I am focusing on their directional relationship excluding the common input.

Data acquisition

Spike train data

A spike train s can be represented as an ordered set of non-negative real values where each value indicates the time when a spike occurs: $s = \{t_i \in [0, T] \mid i \in 1, 2, \ldots, n\}$, with $0 \leq t_1 < t_2 < \ldots < t_i \leq T$, where t_i are the spike times, n is the number of spikes recorded, $T \in R$ is the duration of the recording period. The spike times are derived from intra- or extracellular recordings. These spikes reflect the electrical signal transmitted across the membrane. Spikes also indicate when neurons initiate the synaptic release of neurotransmitters, facilitating communication with other neurons. When we treat spikes as identical brief events (about 1 ms), the spikes are characterized by the time they occurred[12], therefore we can use a spike train to represent the output behaviour of a neuron.

In practice, when using extracellular recording methods, the extracted (also called “sorted”) spike times are the collective electrical activity generated by action potentials propagating through the neuron and its surroundings. Essentially, an entire neuron is treated as a single point with a unified voltage measurement.

Compared with electric signals in raw waveform (or converted into frequency domain), the spikes are the signals that are transmitted to neurons over long distances, excluding the local computation or communication, which makes them more suitable for the analysis focusing on connectivity. However, since the information in spike train data is all contained in the presence and the location of the events, the usual operations focusing on the amplitude of the signal are not directly applicable to the analysis of spike train data[13, p. 8].

Multi-unit extracellular microelectrode recordings

Multi-unit extracellular microelectrode recordings refer to the methods that put microelectrodes in the extracellular space near neurons in vivo to record the electrical activity of multiple neurons at the same time. These type of methods do not need to penetrate the cell membrane of the recorded neurons, therefore it does less damage to the tissue and makes long-term recordings in behaving animals possible. Mixed spikes from different neurons recorded are then separated and identified using a class of techniques known as spike sorting[14].

Neuropixels 2.0

The Neuropixels 2.0 neural probe[15] is an advanced silicon CMOS digital integrated microsystem introduced in 2021. It enables long-term monitoring and dense sampling (at a 30 kHz sampling rate) of single-cell activity, including spikes and local field potentials, within large neuron populations. This is achieved across 4 parallel, tightly aligned ~1 mm x 10 mm planes, positioned perpendicular to the brain surface, and can be applied in both head-fixed and tethered freely moving animals. It has 1280 low-impedance TiN recording sites densely tiled along each 10 mm-long straight shank, making it 5120 electrodes divided over 4 shanks for each probe. Each probe can track hundreds of neurons for up to two weeks with a success rate exceeding 90%, and for up to two months with a success rate exceeding 80%. One headstage can connect up to 2 probes, which is in total 10240 recording sites, 768 of which are recordable simultaneously through a single headstage. Implants that can be carried without impeding the behavior of a mouse must weigh less than ~3 g. Two probes and their headstage weigh ~1.1 g.

Tetrode

A tetrode[16] is an array of bundled four microwire electrodes (e.g., tungsten microwires). The four electrodes are wrapped in insulating material individually, and only the tips of the electrodes are exposed. The goal of using multiple electrodes simultaneously to record neural activity in the same area is to leverage the slight differences between each channel, which arise from their varying distances to distinct signal sources, to classify signals originating from different neurons. This idea was first implemented in 1983, and it is still widely used in many neurophysiology laboratories, partly because they can be fabricated by hand[2]. A single tetrode can record up to 20 neurons. Usually, 20-40 tetrodes are inserted into the brain which can record hundreds of neurons simultaneously[17].

Spatial and temporal resolution

Compared with non-invasive recording methods like electroencephalography (EEG) and functional magnetic resonance imaging (fMRI), multi-unit extracellular microelectrode recordings have higher spatial (cellular level) and temporal (30 kHz) resolutions. EEG has a high temporal resolution of up to 30 kHz[18], though it is typically at 256 or 512 Hz since the frequency of the clinical EEG signal ranges from 0.5Hz to 70Hz[19]. Considering the Nyquist-Shannon sampling theorem, the sampling rate higher than twice of the signal frequency would be enough. However, EEG has a low spatial resolution of 1 to 10 centimeters[20]. fMRI has a low temporal resolution at 2~3s for each volume (full brain) or

~1s for each volume for part of the brain and a relatively low spatial resolution ranging between 2.8 and 3.5 mm (at 3T fields)[21].

Optotagging

To investigate the relationship among groups of neurons, one can record their activity during behavior and also use optogenetic manipulation to drive or inhibit certain populations of neurons in vivo. By observing how these manipulations affect other neurons, insights into network function can be gained.

Optotagging (optogenetic tagging) is specifically used to identify or "tag" which recorded units belong to a genetically defined neuron population. Genetic tools are employed to express light-gated ion channels (e.g., opsins) in a particular cell type; then, during multi-unit recordings, brief pulses of laser light evoke responses only in neurons that express the opsin, thereby revealing their identity among the larger population[22]. For example, layer 6 corticothalamic neurons can be selectively labeled (and thus tagged) in Ntsr1-Cre mice[23], enabling researchers to distinguish those specific cells from others in the recording based on their laser-evoked spiking.

Data processing

After obtaining neural activity data in spike train format, various methods can be employed to measure effective connectivity. These include dynamic causal modelling, Granger causality, and transfer entropy. Other approaches in this field, such as structural equation modelling and variants of the aforementioned methods, are not included in this review.

Dynamic Causal Modeling

The dynamic causal modelling (DCM)[24] was proposed in 2003 for fMRI data. This method sets a model using differential equations for the first derivative of the target signal (e.g. fMRI signal) and then uses Bayesian inference[25] to infer the parameter of the model from data.

To be specific, suppose the neural state $z_t = [z_{t1},..., z_{tn}]^T \in R^n$ is a vector containing the state of all n groups of neurons (or brain areas) we are interested in at time t. The model describing the changes of $z \in R^n$ would be $z' = F(z, u, \theta)$, where F is some nonlinear function describing the changes of neuronal states, $u \in R^m$ are the inputs to each group of neurons that exerts the activity and changes of connectivity within the interested groups, θ are the parameters of the model (function F) that needs to be inferred. For example, when the model is linear, we have $z' = \theta_1 z + \theta_2 u$.

In the original dynamic causal modelling, the first derivative z' was approximated by a bilinear form:

$$\frac{dz}{dt} \approx F(z, u, \theta)$$
$$= Az + \Sigma u_j B^j z + Cu$$
$$A = \frac{\partial F}{\partial z} = \frac{\partial z'}{\partial z} \in R^{n \times n}$$
$$B^j = \frac{\partial^2 F}{\partial z \partial u_j} = \frac{\partial}{\partial u_j}\frac{\partial z'}{\partial z} \in R^{n \times n}$$
$$C = \frac{\partial F}{\partial u} \in R^{m \times n}$$

Where $Az + Cu$ represents the total first-order derivative of F, and an interaction term $\Sigma u_j B^j z$ is added as the partial derivative term $\frac{\partial^2 F}{\partial z \partial u_j}$, which makes the function F nonlinear. A represents the connectivity among the groups of neurons in the absence of input. B^j are the changes in effective connectivity induced by the j-th input. C quantifies how external or experimental inputs affect the neuronal system. In practice, there's no need to write out the actual form of F, since we can always use the bilinear form or other approaches to approximate it.

After getting the form of the first derivative model F, DCM then uses Bayesian inference to estimate the parameter θ of the model. The priors of the inference can include principled prior knowledge like some values cannot be negative, and empirical knowledge from previous experiments. After obtaining the distributions of parameters θ, the researchers can optionally compare different models (different approximations) and select the one that balances best between fit to the data and model complexity using Bayesian model selection[26].

Conclusions can be drawn from the parameters and comparing different models. For example, larger values of parameters indicate more intensive coupling. When testing the hypothesis that isolation induces sensorimotor abnormalities in rats by increasing the amplitude of excitatory postsynaptic potentials elicited by presynaptic input, researchers found that the intrinsic coupling parameters of the DCM, derived from local field potential recordings, were significantly larger in the isolated group compared to the control group[27]. In another example, Kiebel et al. compared different models of DCM to disambiguate the direction of coupling between the hippocampus (CA1) and the lateral nucleus of the amygdala (LA) during the retrieval of negative memories in a fear-learning study in mice[27].

DCM for spike train data

To use dynamic causal modelling on spike train data, we need an interpretable generative model that describes the evolution of the hidden (latent) states over time. Since we need to compare the model parameters under different conditions to test hypotheses about the interaction of real neural networks, black box models using artificial neural networks are in general unusable. This type of generative modelling is still an ongoing research. In 2022, Williams et al. developed a point process model for neural sequences (PP-Seq) that characterizes fine-scale sequences at the level of individual spikes, and demonstrated their

advantages on experimental recordings from songbird higher vocal center and rodent hippocampus[28].

Granger Causality

In 1969, to decide the causality between two related variables, Clive Granger proposed a quantitative definition of causality, now known as Granger causality (GC)[29], to measure the causal lag and strength. In general, “We say that Y_t is causing X_t if we are better able to predict X_t using all available information than if the information apart from Y_t had been used”[29]. Granger then reduces the predictability to a decrease in the variance of prediction error.

The formal definition of X causing Y is as follows. Suppose X and Y are two stationary stochastic processes, which means the causal relationship between them does not change during the time the data is obtained. X_t is the state of X at time t. $U_{<t}$ is a set containing all time series related to X and Y up to time step $t-1$. Suppose $P(X_t|U_{<t})$ is the optimal prediction of X_t given $U_{<t}$. X is Granger causing Y if:

$$var[Y_t - P(Y_t|U_{<t})] < var[Y_t - P(Y_t|U_{<t} \backslash X_{<t})]$$

That is, the variance of the prediction error of Y_t is increased when excluding the history of X for prediction, which means the prediction is more accurate when including the history of X. The choice of using variance as a criterion to measure the closeness of predictions to true values is due to its ease of mathematical handling and simplicity of interpretation. Due to the variety of nonlinear forms in modelling, Granger used a linear estimator (a vector autoregressive model) to approximate the optimal estimator which is:

$$P(Y_t|Y_{<t}, X_{<t}) := \sum_{i=1}^{\infty} a_i X_{t-i} + \sum_{j=1}^{\infty} b_j Y_{t-j}$$

The parameters a_i and b_j are determined by minimising the sum of the squared estimation errors. The computational complexity of this estimator is $O(M^3 k^3 N)$ where N is the number of samples, M is the number of variables, k is the model order[30], which means the computational time grows linearly with the number of samples.

When people refer to Granger causality without further elaboration, they typically mean the linear approximation of the Granger causality estimator. However, the concept of GC also encompasses more complex nonlinear estimators.

After obtaining the fitted model, people usually do a F-test[31] or a permutation test for the significance of the result. In the permutation test, surrogate data is generated by permutation of original data under the null hypothesis that there is no causal relationship between X and Y. The statistics to be compared in the test is Granger causality strength, which is usually defined as $F_{X->Y} = ln\frac{\Sigma_1}{\Sigma_1}$,where Σ_1 is Y's unexplained variance in its autoregression, Σ_2 is Y

's unexplained variance in the joint (X and Y) regression. When the Granger causality strength of original data is significantly larger (defined by p-value) than the surrogate data, we conclude that X Granger causes Y. For example, Brovelli et al. recorded local field potentials of up to 15 sites in the hemisphere contralateral to the dominant hand of monkeys while they were performing a GONO-GO visual pattern discrimination task. They compared the peak Granger causality values of original data with 1,000 permuted versions and created a table summarizing the pairwise significant ($P < 0.005$) peak Granger causality values between sites[32].

Interpretation of the results

From the analysis of variance, we can know if the past X helps to predict the future Y, but how does this fit into the concept of functional or effective connectivity?

From the perspective of functional connectivity, the result of GC depends on the choice of statistical models used for prediction. The better the prediction, the closer the result is to the true GC (the model here can be a black box like artificial neural networks). Then the result should be interpreted as a measure of the functional connectivity parameterized by the used model (e.g. the linear model used in the original paper). Under different choices of the model, you will get functional connectivity under different definitions. For example, when you use linear models, you will get a GC measure of linear dependence.

From the perspective of effective connectivity, GC compares two models: one considers the impact of the source process's history, and the other does not. Through model comparison, which has a central role in effective connectivity, GC accept or rejects the hypothesis that the past X helps to predict the future Y.

Granger defined the strength of the causality from X to Y as $1 - \Sigma_2 / \Sigma_1$ where $\Sigma_1$and Σ_2 are the variance of the error of predictions without and with the history of X[33]. Similarly, John Geweke defined the measure of linear feedback as $ln(|\Sigma_1|/|\Sigma_2|)$ [34].

GC for spike train data

One may observe that a discrepancy remains between the continuous-valued states of the variables in the original GC and spikes in spike train data (i.e. points in the point process). This gap can be simplistically addressed by converting the spike train data into a histogram, treating it as a categorical time series where each unique spike count in the histogram represents a category. Due to the absolute refractory period of spiking neurons, when the bin size is less than 2 ms, there should only be two categories (0 and 1) in the histogram. Models for categorical time series include mixture transition distribution (MTD) model as well as multinomial logistic transition distribution (mLTD) model for autoregressive categorical processes and models focused on the high dimensional multivariate setting such as generalized linear autoregressive (GLAR) model, integer-valued autoregressive (INAR) model and integer-valued generalized autoregressive conditional heteroskedasticity (INGARCH) model[35].

However, the histogram-based approach inevitably loses information about the precise timing of spikes, motivating the adoption of conditional intensity function as a more informative representation. In 2011, Kim et al. proposed GC for point process data which uses conditional intensity function to fully describe the point process[36]. The conditional intensity function $\lambda(t \mid H_t)$ of a point process is the instantaneous probability that a point will be dropped (a spike will occur) at the time t, given the history up to this time H_t. Denote the event count during time interval $[t, t + dt)$ as $N(t, t + dt)$. The conditional intensity function is defined as:

$$\lambda(t \mid H_t) = \lim_{dt \to 0} \frac{\mathbb{E}[N(t, t + dt) \mid H_t]}{dt}$$

The probability that a neuron fires a single spike in a small interval $[t, t + dt)$ can be approximated as $\lambda(t \mid H_t)dt$. The paper then uses a generalised linear model (GLM) as the predictor and employs a likelihood ratio statistic to test the significance of detected causal interactions. To demonstrate, this method was applied to neural activity recorded from several points in the primary motor cortex (MI) of a cat during a reaching task[37] to investigate the causal relationship within MI. A causal connectivity map among 15 neurons was generated, revealing the presence of qualitative inhibitory and excitatory interactions. Ten years later, an improvement of this approach was proposed to overcome the problem of unobserved causes (e.g. exogenous temporal modulations) by modeling them as additional parameters in the GLM of the neuronal conditional intensity function[38].

The conditional intensity function representation is later used in GC for Hawkes processes. Hawkes process is a class of self-exciting point processes, which can be used as a model for neuronal activities. The probability of the occurrence of an event during $[t, t + dt)$ is described as:

$$\lambda(t)\, dt = \left(\mu(t) + \sum_{t_i : t_i < t} \phi(t - t_i) \right) dt$$

Where λ is the conditional intensity function, μ is the intensity of an underlying Poisson process, ϕ is a kernel function (some called influence function or interaction function). $\phi(t - t_i)$ quantifies the contribution level of an event occurring at time t_i to the intensity of the process at the current time t. When considering the interaction among n Hawkes processes, we get:

$$\lambda_i(t) = \mu_i(t) + \sum_{j=1}^{n} \sum_{t_i : t_i < t} \phi_{ij}(t - t_i),\ i = 1, ..., n$$

i-th process does not Granger cause j-th process if and only if $\phi_{ij}(t) = 0$ for all t. The conditional intensity functions λ_i will be estimated from the data. Then the kernel functions ϕ_{ij} can be interpreted to have excitatory or inhibitory effects based on its value above or under zero. For example, Casile et al. used their G-ETM (Granger causality with Exogenous

Temporal Modulations) model on spike train data from 57 trials of 12 neurons from the monkey premotor cortex (area F5) during the preparation of goal-directed motor acts. Consistent with previous studies, they found that the responses of neurons in area F5 were significantly modulated by the preparation of a motor act. From the kernel functions, they found both self- and mutual interactions are time-dependent with a general trend of being inhibitory at shorter time scales and excitatory at longer time scales[38].

Challenges and advances

Granger causality gave a quantitatively verifiable definition of causality, which addresses causality's direction between two correlated variables. This makes it applicable to various problems and can provide directional results. Apart from its pioneering contributions to causal inference, GC still faces several technical challenges. These include the assumption of stationarity in the data, the potential influence of unobserved variables, the determination of the appropriate lag for influence, the handling of diverse data types, and the handling of relationships among sets of variables[35]. One of the major issues related to the analysis of spike train data is the unobserved variables. In multiunit extracellular microelectrode recordings, only a tiny part of the closely (functionally or anatomically) related neurons can be reliably recorded. This means the assumption of knowing the history of all relevant variables is seriously violated, which may lead to spurious results. Another major issue is the comparability of the result. Since the strength of causality is defined on the comparison of variance of prediction error, there is no universal interpretation of that among different experiments, due to the different variances of the data and perhaps different models used for prediction. Considering three distinct brain areas A,B and C, this means $F_{A->B}$, the causality strength from brain area A to B, can be compared to $F_{B->A}$ but not to $F_{B->C}$. Similarly, the coefficient of determination (R^2) is only meaningful for comparing the fit of different models applied to the same dataset.

Relevant research is still going on today, focusing on overcoming the above-mentioned issues and approaching the optimal prediction for various types of data. Some approaches use artificial neural networks including multilayer perceptron (MLP), recurrent neural network (RNN) and long-short term memory network (LSTM), to capture nonlinear relationships in the data for better prediction[35].

Transfer entropy

A more recent method to measure functional connectivity is transfer entropy (TE), which was proposed by Thomas Schreiber in 2000[39]. It was designed to measure the amount of directional information transfer between two stationary random processes.

Information is defined as the decrease of uncertainty by Shannon in 1948, which can be quantified by Shannon entropy $H(X) := -\Sigma_{x\in\chi} P(x) log_2 P(x)$[40]. It measures the expected amount of information conveyed by a random variable X, by trying to optimally encode the outcome of random trials on X. It can be interpreted as the number of bits (when in the base

of 2) on average to encode a message that represents the state of X using an optimally efficient encoding scheme.

The transfer entropy from variable X to Y ($T_{X->Y}$) quantifies how much additional information X_{past} provides about Y_t when Y_{past} is already known. It coincides with Granger's idea of causality that if X_{past} helps predict Y_t then X is (at least part of) the cause of Y.

The transfer entropy during time T can be formally written as follows:

$$TE_{X\to Y} = H(Y_{t+1} \mid Y_t^{(k)}) - H(Y_{t+1} \mid Y_t^{(k)}, X_t^{(l)})$$

with X_{past}as $X_t^{(l)} = (X_t,\ldots, X_{t-l+1})$,$Y_{past}$as $Y_t^{(k)} = (Y_t,\ldots, Y_{t-k+1})$ represents the history of X and Y until time t with a length of l and k. The conditional Shannon entropy $H(X|Y) = H(X,Y) - H(Y)$ can be understood as the amount of additional information in X when Y is already known. Later, the definition of TE is extended by adding a delay τ from source to target yielding

$$TE_{X\to Y} = H(Y_{t+1} \mid Y_t^{(k)}) - H(Y_{t+1} \mid Y_t^{(k)}, X_{t-\tau}^{(l)}) = \sum_{t\in T} P(Y_{t+1}, Y_t^{(k)}, X_{t-\tau}^{(l)}) \log \frac{P(Y_{t+1} \mid Y_t^{(k)}, X_{t-\tau}^{(l)})}{P(Y_{t+1} \mid Y_t^{(k)})}$$

l, k and τ are three hyperparameters that need to be chosen by the user based on the properties of the data and computing resources available. The unit of TE depends on the base of the log function. If the base is two then TE is in bits. Sometimes TE values are normalized by the temporal duration of data, as longer periods generally encapsulate more information.

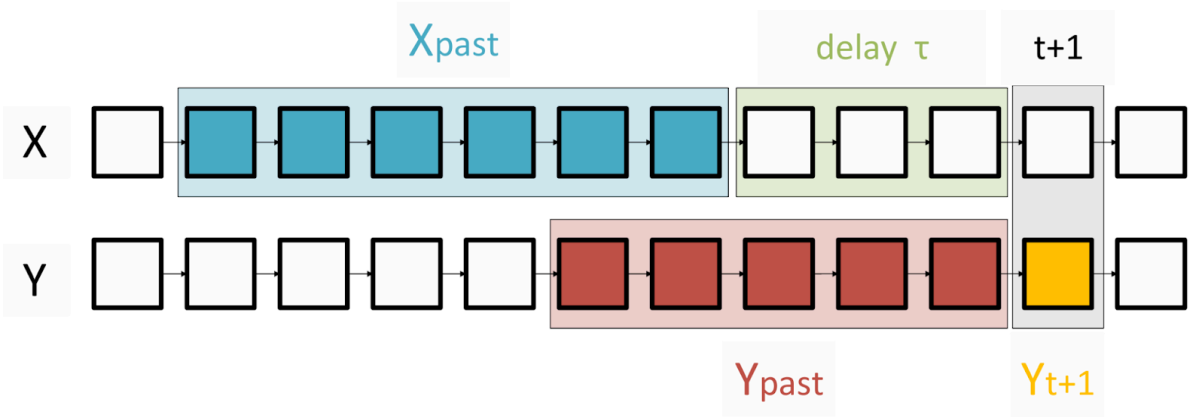

Figure 2. Diagram of transfer entropy $TE_{X->Y}$ with delay. In this case, $l = 6$, $k = 5$, $\tau = 3$.

One significant advantage of transfer entropy (TE) over Granger causality (GC) is that TE is model-free, meaning it does not require a predefined model to describe the interactions between variables for measurement. In contrast, the original Granger causality approach relies on a linear model to assess these interactions. Ideally, TE can capture linear and all possible nonlinear interactions. However, Ursino et al. recently showed that TE fails when interactions are highly nonlinear using a simulation of Rössler and Hénon attractors[41].

The choice of hyperparameters

The choice of hyperparameters is not a trivial thing. They considerably affect the outcome of the TE estimates. For instance[42], a small embedding length can be insufficient to represent the historical state and consequently degrade the meaning of TE measurement. While a too-large embedding length makes the estimators less accurate, especially for a fixed data length and increases the computing time.

Ideally, one would choose the combination of the three hyperparameters embedding length l , k and delay τ that yields the largest TE value and then tests the significance of the result. Currently, no theoretically proven method can ensure the optimal choice without evaluating all possible combinations. Some use heuristic rules like Ragwitz's or Cao's criterion to decide the embedding length and delay[43]. Some choose delay as the auto-correlation decay time or the first minimum of the auto-information[42].

In dynamical systems, where these two criteria were originally used, embedding lengths smaller or larger than the real dimension of data will lead to worse representation. For some spike train data, TE keeps increasing when l and k are increasing, since the longer the embedding length, the better it can represent the history of the process. A longer history can account for the long-term effect of neuron activities. If so, l and k should be chosen based on the computing capacity.

Despite the importance of the choice of hyperparameters, due to limited computing resources, many researchers use small (<10) embedding lengths[43] or even $l = k = 1$[44]. Researchers usually vary the delay with fixed embedding length to find the maximum TE, but it is essentially a combinatorial optimization problem.

Estimation of transfer entropy

Though we already have a clear definition of transfer entropy and a basic scheme to choose hyperparameters, there is still a technical but non-trivial issue in the calculation of transfer entropy. The challenge lies in the imperfect estimation of the joint or conditional distribution of data. This issue becomes particularly evident in the continuous domain, where we have only a finite number of discrete points but aim to estimate probabilities across the entire domain. An unbiased and efficient algorithm to estimate transfer entropy is the key for its application. Otherwise, we can only treat TE as a theoretical tool, as it becomes computationally infeasible to apply it to actual data.

Considering the difference in data type, there are two classes of estimators: one for discrete variables and the other class for continuous value variables. In reality, most signals are continuously valued, such as membrane potentials, chemical concentrations and blood oxygen level dependent (BOLD) signals. Meanwhile, some signals like categorical behaviors, uniform optogenetic stimulation, and neural spikes can be simplified as discrete valued variables. To be more specific, the timing of spikes are continuously valued, but the events (spikes) are discrete and uniform. A common approach to analyzing spike train data involves binning spikes into a histogram and using the spike counts within each time window as discrete variables. Recently, however, an estimator[45] was proposed that fully captures

the continuous values of inter-spike intervals, specifically designed for event-based data like spike train data.

Limited by the length of this review, only several estimators are introduced. Other estimators like linear, fixed-binning with ranking, kernel density, and adaptive partitioning are compared by Andrea et al.[46] .

Histogram-Based Estimator

Histogram-based estimator bins the continuous signals into histograms and then estimate the joint distribution of $p(x, y)$ in discrete value by $p(x_i, y_i) \approx \frac{count(x_i, y_i)}{N}$. Another similar approach is called symbolic estimator[47], which discretizes the continuous amplitudes into different discrete levels. These estimators are significantly faster than other estimators working in the continuous value domain, since they only need operations on discrete values and use simple schemes for estimation. However, they lose some information during discretization if the original data is continuously valued, which makes the result less accurate.

Kraskov-Stögbauer-Grassberger Estimator

The Kraskov-Stögbauer-Grassberger (KSG) estimator[48] was proposed in 2004 to estimate mutual information. Since transfer entropy can be expressed as conditional mutual information, this technique is also used for TE. The KSG estimator converts the original problem into a k-nearest neighbours (kNN) problem. For each data point x, It finds the k (k as a hyperparameter) nearest neighbours of that point and use their distances to x to estimate the entropy differences without reconstructing the full joint probability distribution.

KSG estimator adapts to local densities. It has higher resolution at dense areas, and lower resolution when data points are sparse. This property makes it scales much better to higher dimension ($O(D\ log\ D)$) compared with histogram-based estimator ($O(a^D)$).

Continuous-Time Estimator

In 2021, Shorten et al. proposed a TE estimator for spike train data. It uses the inter-spike interval to construct an instantaneous firing rate and later use this rate as the amplitude of the signal[45]. The paper then uses a kNN estimator (most notably the KSG estimator) to complete the calculation. In this way, the estimator can capture the exact timing of events and can represent both short- and long-term dependencies, which is important when virtually all excitatory synapses in the mammalian brain simultaneously express a number of different forms of synaptic plasticity such as paired-pulse depression (<20 ms) and facilitation (20-500 ms), depression and facilitation following trains of stimuli (seconds to minutes) and long-term potentiation (hours, days, or even longer)[49]. The continuous-time estimator addresses the issue in histogram-based estimators that they can represent either short- or long-term dependencies using small or large bin sizes.

Interpretation of the results

By definition, transfer entropy measures the amount of information transferred during a certain period. When it is normalized by time, we get the throughput of a directional

communication. With spike train data, this could tell us how fast one neuron x sends information to another neuron y, though essentially, it shows how much information we can get about y uniquely from the observation of x.

For the significance test, people usually use bootstrapping methods to get a p-value. As there is no no analytical form for the TE value over the null hypothesis that no causal relationship exists, researchers shuffle the data in different ways to generate spike trains that have some properties similar to the original data but without causal relationships and use these surrogate data for bootstrapping.

Apart from normalizing TE by time, there are also other ways to normalize the TE and get different interpretations. For example, Francisco et al. used a normalized directional phase TE on local-field potential recordings from the fronto-auditory network of vocalizing bats. They normalized the TE by $dTE_{X->Y} = TE_{X->Y}/(TE_{X->Y} + TE_{Y->X})$, and found predominant top-down (frontal-to-auditory cortex) information flow during spontaneous activity and pre-vocal periods[4]. It is a good idea to use the ratio as the result, since the absolute value of TE can vary for different processes, though this normalization cannot differentiate weak causality and strong bi-directional causality.

One might say TE reveals the causal relationship. While this is true when “causal” means Granger causal, it does not necessarily imply causality in the broader sense. Causality can exist without information transfer, but information transfer cannot exist without causality. Therefore, when we observe a significant TE value, we can get a qualitative result that concludes that one causes another, but we can not conclude that one does not cause another when TE value is not significant. For a quantitative result like the causal strength or functional connectivity strength, it needs further discussion.

In Karl J. Friston’s view, transfer entropy measures functional connectivity, since it only depends on the distributions from the data and does not model any interactions[5]. This means no conclusion about the mechanism change can be drawn from the analysis. One can only study the temporary relationship using TE. Using the analogy of server cluster again, TE can only tell us how much data has been transmitted during a certain period, but we cannot know whether the software that generates the data flow has changed or not. Some other researchers also argue that TE is not a measure of coupling strength and cannot be used to recover an estimate of a coupling parameter[50].

However, an empirical study showed that TE can be linearly correlated to coupling strength under certain conditions by using TE on synthetic data generated by neural mass models[41]. Due to the models they use, this result is only valid for the analysis of brain regions, when the populations work in the linear region, input noise is stable and the epochs lengths are longer than 10 s. It is no longer valid if other strong nonlinear effects (e.g. synchronization in nonlinear oscillators) become influential. This study sheds light on the interpretation of TE that it should be used with extreme caution to measure a coupling parameter (e.g. the weight of synapses).

Specifically for thalamocortical circuits, Adrià et al. interpret directional information (DI), a directed information-theoretic measure (similar to TE) together with other data such as firing

rate, stimulus and behavioral data[51]. They recorded spike trains in the ventral posterolateral nucleus (VPL) of the somatosensory thalamus and in primary somatosensory cortex (S1) while monkeys performed a vibrotactile detection task. In the analysis, they found the amount of VPL–S1 DI is dependent on the stimulus amplitude. They compared the vibrotactile detection task and the passive stimulation task to show that DI across VPL–S1 neuron pairs was sensitive to the task context. They found over 33% of bidirectional information occurs at zero delay, and the modulation of bidirectional information by the stimulus amplitude was mostly explained by zero-lag bidirectional information. By analyzing the behavioral performance subgroup (hit or miss) of spike train data, they discovered a correlation between the amount of feed-forward information during the first half of the stimulus window and the monkey's behavior.

Applications

Harris et al.[52] used TE to measure the amount of information transfer from relay neurons in the thalamus to spiny stellate (SS) cells in layer 4 (L4) of the primary visual cortex (V1) in rat brain slices. They stimulated thalamic axons making synapses onto whole-cell patch-clamped L4SS cells by replaying previously recorded spike trains (25 seconds, repeated 5 times), plotted energy efficiency, defined as the ratio of information transfer (in bits/sec) to EPSC energy consumption (in ATP/sec), against postsynaptic conductance and found that energy efficiency is maximised near the physiological conductance of the weak thalamocortical synapse.

Nigram et al.[44] applied TE to in vitro recordings from rodent somatosensory cortex and in vivo recordings from rodent orbitofrontal cortex during an odor discrimination task to quantify the information transfer within the cortex. The recording duration was approximately 1 hour per trial, capturing up to 500 neurons in each session. They defined the information transfer (IT) as the difference between TE from original and jittered data. Then they plotted the percentage of incoming and outgoing IT against the proportion of respective neurons, and found that 20% of the neurons contributed ~70% of the total information transfer.

Schroeder et al.[53] recorded spike trains from monkeys' (N=2) primary somatosensory cortex (S1, area 3b) and primary motor cortex (M1, area 4) before, during, and after ketamine-induced unconsciousness with stroking stimulation. Data were prepared by extracting 3-second windows of spike times per trial and concatenating them into 3- to 6-minute vectors per multiunit. They found the maximum TEs were significantly smaller under ketamine compared to the awake state in all directions within and between S1 and M1, and concluded that information transfer was disrupted between neuronal units in S1 and M1 during ketamine anesthesia.

Advantages and challenges

One of the key advantages of transfer entropy is its ability to quantitatively measure directional relationships, which is particularly important when the analysis requires more than just correlation. It is model-free in the way that no model about the relationship is needed for the measurement, which means no assumption about the data is required. Therefore it can detect any linear or nonlinear relationships as long as the conditional probabilities are estimated accurately.

Compared with Granger causality, TE addresses the common input problem, which happens when the assumption of Granger causality that all relevant information is included in the analysis is violated. Take spike train data as an example. Suppose we want to quantify the relationship between two brain areas, and already got recordings of two groups of neurons from these two brain areas. In the optimal setting, the application of Granger causality would also require recording all neurons connected to those in the groups—some of which may be located in other brain areas—in order to remain as faithful as possible to the original definition. Otherwise, Granger causality will discover a spurious relationship if there's a neuron that gives inputs to neurons in both brain areas. TE on the other hand, handles this by excluding the information in the past of the target process.

However, despite all advantages, certain issues limit the application of TE. The major issue is the computational complexity of the estimator. A naïve implementation may yield a complexity of $O(N^2)$, which can be optimized to $O(N\ log\ N)$ in low dimensional settings[54], which means the computational time grows faster than linearly but slower than quadratically as the number of samples N increases. Some estimators of TE, like histogram-based estimators, suffer from the curse of dimensionality, which requires exponential increase in computing time as the dimension of the variable increases. Due to no knowledge about the analytical distribution of TE value of null hypothesis, permutation tests (bootstrapping methods) are needed, which would cost at least 10,000 times (the default number of permutations in TRENTOOL[43] is 190,100) more computation to rigorously test a hypothesis. Also, the estimators may have high variance and require large amounts of data for reliable estimation[35]. Related to computational cost, we lack theoretical guidance for the choice of parameters when testing all physiologically plausible combinations of embedding lengths and delays are computationally intractable.

Finally, there is a problem of interpretation after we get the absolute value of TE. For the same two groups of neurons, TE may vary across different behaviors, even when no change of long-term effective neural coupling occurs. If a neuron is as complicated as a computer, we are trying to infer whether the running software is different by merely looking at the network throughput of them. It would probably be more realistic to only focus on the realtime (temporary) characteristics in this case.

Conclusions

In this review, I introduce the pipeline to measure the connectivity in spike train data from data acquisition using multi-unit extracellular recordings to data processing using dynamical causal modelling, Granger causality and transfer entropy.

Regarding the choice of method in the data processing, it mainly depends on what kind of results one wants from the data. There is a hierarchy of explanatory power and level of detail among these three methods. If one wants to study a real time relationship that temporarily reflects the activity of neurons, for example, how fast groups of neurons communicate during particular activities, transfer entropy should be the choice. If one wants to know a long-term structural change, that is, the change of functional connectivity, one should use Granger causality. If one wants to further investigate the structural change of the mechanism with

more details, like the change in the function that describes the interaction of the neurons, one should use dynamic causal modelling.

However, the explanatory power and level of detail will not come without cost. These three methods also have a hierarchy in the requirements of the model they use for estimation. Transfer entropy is model-free, meaning it does not require a predefined model of neuronal interactions for estimation. Consequently, it makes no assumptions about these interactions, making it universally applicable. Granger causality requires models for prediction to estimate. These models could be interpretable or not, which allow us to use powerful yet uninterpretable models like artificial neural networks. Dynamic causal modelling has the strictest requirements on models. It asks for an interpretable generative model. The models used in GC or DCM should have a satisfactory level of accuracy. Otherwise, the differences may be attributable to model misspecification or noise.

For the sake of robustness, TE is favored because it avoids structural model misspecification by not relying on predefined functional forms. In contrast, GC always assumes the causality can be revealed by the chosen predicting model. However, this assumption does not always hold true, even when a nonlinear model is employed, particularly if the model lacks certain critical terms. The distinction between linear and nonlinear GC illustrates this limitation—when certain types of relationships are not accounted for in the model, causality may go undetected. Similar issue also applies to DCM, where an incorrect or incomplete generative model can lead to misleading conclusions. Model misspecification is a critical limitation for parametric methods like GC and DCM, as their assumptions (e.g. linearity or predefined generative models) may not hold in practice. Universal function approximators, such as artificial neural networks, can mitigate structural misspecification by flexibly learning relationships directly from data. However, practical challenges, such as overfitting, may still introduce errors.

In addition to theoretical considerations, computational time becomes a critical factor in determining a method's applicability, particularly when dealing with the vast volumes of data obtained by neuroimaging techniques. Among three methods, Granger causality is the most efficient when using linear models like vector autoregressive (VAR) models, of which the computational time scales linearly with the number of samples (N). Although for better predicting performance, one might use nonlinear extensions of GC, which will cost more time. Dynamic causal modelling is more computationally demanding than GC, because it needs to solve an inverse problem over nonlinear dynamical models through repeated model simulations and parameter updates. Transfer entropy is the most computationally intensive of the three, owing both to the high-dimensional density estimation and to the additional cost of significance testing via surrogates. With optimization, its computational complexity for low dimensional embedding can be reduced from $O(N^2)$ down to $O(N\ log\ N)$.

When considering the choice for the spike train data obtained from multiple brain areas, one should firstly consider the connectome of the regions of interest. If the direct connection between the areas forms a directed acyclic graph, then GC with linear models should be the first choice since it is computationally efficient and only linear interactions need to be considered. When the graph is cyclic, nonlinear interactions need to be considered, though GC with linear models can still be performed as a pragmatic initial approach.

In the nonlinear case, when inferring functional or effective connectivity from data, TE is generally preferred over GC due to its higher accuracy. This is partly because GC assumes that all relevant processes influencing the source and target are included in the historical data (no unmeasured confounders). However, in neuroscience, only a limited subset of neurons can be recorded simultaneously in vivo, making GC less reliable. It struggles to distinguish whether neural signals are directly connected or merely correlated due to a shared, unobserved driver. Recent developments in partial information decomposition[55] offer an alternative approach to addressing this shared driver issue by disentangling unique and redundant information contributions.

DCM is most suitable when the underlying mechanism is relatively well understood, requiring only the estimation and comparison of specific parameters based on the data. Unfortunately, modeling statistical dependencies in multi-region spike train data is still in its infancy[56].

Concerning the stationarity assumption required by TE, we can reasonably assume that within a short time window (e.g., within each recording session), the underlying mechanism remains stable or changes only minimally, making this assumption a good approximation.

In the future, if we gain the capability to simultaneously record nearly all neurons relevant to both the source and target, GC with nonlinear models—particularly artificial neural network-based GC methods—would be a strong alternative. These models can make good use of large amounts of data and excel at approximating arbitrary nonlinear functions, making them highly effective for prediction tasks.

References

[1] K. J. Friston, “Functional and effective connectivity in neuroimaging: A synthesis,” *Hum. Brain Mapp.*, vol. 2, no. 1–2, pp. 56–78, Jan. 1994, doi: 10.1002/hbm.460020107.

[2] B. Zhang, C. Deng, C. Cai, and X. Li, “In Vivo Neural Interfaces—From Small- to Large-Scale Recording,” *Front. Nanotechnol.*, vol. 4, Jun. 2022, doi: 10.3389/fnano.2022.885411.

[3] A. S. Maida, “Chapter 2 - Cognitive Computing and Neural Networks: Reverse Engineering the Brain,” in *Handbook of Statistics*, vol. 35, V. N. Gudivada, V. V. Raghavan, V. Govindaraju, and C. R. Rao, Eds., in Cognitive Computing: Theory and Applications, vol. 35. , Elsevier, 2016, pp. 39–78. doi: 10.1016/bs.host.2016.07.011.

[4] F. García-Rosales *et al.*, “Echolocation-related reversal of information flow in a cortical vocalization network,” *Nat. Commun.*, vol. 13, no. 1, p. 3642, Jun. 2022, doi: 10.1038/s41467-022-31230-6.

[5] K. J. Friston, “Functional and effective connectivity: a review,” *Brain Connect.*, vol. 1, no. 1, pp. 13–36, 2011, doi: 10.1089/brain.2011.0008.

[6] N. Dijkstra, P. Zeidman, S. Ondobaka, M. a. J. van Gerven, and K. Friston, “Distinct Top-down and Bottom-up Brain Connectivity During Visual Perception and Imagery,” *Sci. Rep.*, vol. 7, no. 1, p. 5677, Jul. 2017, doi: 10.1038/s41598-017-05888-8.

[7] D. Beniaguev, I. Segev, and M. London, “Single cortical neurons as deep artificial neural networks,” *Neuron*, vol. 109, no. 17, pp. 2727-2739.e3, Sep. 2021, doi: 10.1016/j.neuron.2021.07.002.

[8] W. J. Murdoch, C. Singh, K. Kumbier, R. Abbasi-Asl, and B. Yu, “Definitions, methods, and applications in interpretable machine learning,” *Proc. Natl. Acad. Sci.*, vol. 116, no. 44, pp. 22071–22080, Oct. 2019, doi: 10.1073/pnas.1900654116.

[9] F. Briggs and W. M. Usrey, “Emerging views of corticothalamic function,” *Curr. Opin. Neurobiol.*, vol. 18, no. 4, pp. 403–407, Aug. 2008, doi: 10.1016/j.conb.2008.09.002.

[10] R. A. Mease, A. Sumser, B. Sakmann, and A. Groh, “Corticothalamic Spike Transfer via the L5B-POm Pathway in vivo,” *Cereb. Cortex*, vol. 26, no. 8, pp. 3461–3475, Aug. 2016, doi: 10.1093/cercor/bhw123.

[11] R. A. Mease and A. Groh, “Corticothalamic Layer 5 Pathways and Synapses,” in *The Cerebral Cortex and Thalamus*, F. Clasca, S. B. Hofer, W. M. Usrey, and S. M. Sherman, Eds., Oxford University Press, 2023, p. 0. doi: 10.1093/med/9780197676158.003.0015.

[12] P. Dayan and L. F. Abbott, *Theoretical neuroscience: computational and mathematical modeling of neural systems*. in Computational neuroscience. Cambridge, Mass.: MIT Press, 2001.

[13] A. R. C. Paiva, I. Park, and J. C. Príncipe, “Chapter 8 - Inner Products for Representation and Learning in the Spike Train Domain,” in *Statistical Signal Processing for Neuroscience and Neurotechnology*, K. G. Oweiss, Ed., Oxford: Academic Press, 2010, pp. 265–309. doi: 10.1016/B978-0-12-375027-3.00008-9.

[14] H. G. Rey, C. Pedreira, and R. Quian Quiroga, “Past, present and future of spike sorting techniques,” *Brain Res. Bull.*, vol. 119, pp. 106–117, Oct. 2015, doi: 10.1016/j.brainresbull.2015.04.007.

[15] N. A. Steinmetz *et al.*, “Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings,” *Science*, vol. 372, no. 6539, p. eabf4588, Apr. 2021, doi: 10.1126/science.abf4588.

[16] B. L. McNaughton, J. O’Keefe, and C. A. Barnes, “The stereotrode: a new technique for simultaneous isolation of several single units in the central nervous system from multiple unit records,” *J. Neurosci. Methods*, vol. 8, no. 4, pp. 391–397, Aug. 1983, doi: 10.1016/0165-0270(83)90097-3.

[17] R. H. Roth and J. B. Ding, “From Neurons to Cognition: Technologies for Precise Recording of Neural Activity Underlying Behavior,” *BME Front.*, vol. 2020, p. 7190517, Dec. 2020, doi: 10.34133/2020/7190517.

[18] R. Li *et al.*, “Design and implementation of high sampling rate and multichannel wireless recorder for EEG monitoring and SSVEP response detection,” *Front. Neurosci.*, vol. 17,

Jun. 2023, doi: 10.3389/fnins.2023.1193950.
[19] C. S. Nayak and A. C. Anilkumar, “EEG Normal Waveforms,” in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2024. Accessed: Nov. 06, 2024. [Online]. Available: http://www.ncbi.nlm.nih.gov/books/NBK539805/
[20] C. C. Ruff and S. A. Huettel, “Chapter 6 - Experimental Methods in Cognitive Neuroscience,” in *Neuroeconomics (Second Edition)*, P. W. Glimcher and E. Fehr, Eds., San Diego: Academic Press, 2014, pp. 77–108. doi: 10.1016/B978-0-12-416008-8.00006-1.
[21] J. M. Soares *et al.*, “A Hitchhiker’s Guide to Functional Magnetic Resonance Imaging,” *Front. Neurosci.*, vol. 10, Nov. 2016, doi: 10.3389/fnins.2016.00515.
[22] A. Guru, R. J. Post, Y.-Y. Ho, and M. R. Warden, “Making Sense of Optogenetics,” *Int. J. Neuropsychopharmacol.*, vol. 18, no. 11, p. pyv079, Jul. 2015, doi: 10.1093/ijnp/pyv079.
[23] K. Ziegler *et al.*, “Primary somatosensory cortex bidirectionally modulates sensory gain and nociceptive behavior in a layer-specific manner,” *Nat. Commun.*, vol. 14, no. 1, p. 2999, May 2023, doi: 10.1038/s41467-023-38798-7.
[24] K. J. Friston, L. Harrison, and W. Penny, “Dynamic causal modelling,” *NeuroImage*, vol. 19, no. 4, pp. 1273–1302, Aug. 2003, doi: 10.1016/S1053-8119(03)00202-7.
[25] G. E. P. Box and G. C. Tiao, *Bayesian Inference in Statistical Analysis*. John Wiley & Sons, 2011.
[26] L. Wasserman, “Bayesian Model Selection and Model Averaging,” *J. Math. Psychol.*, vol. 44, no. 1, pp. 92–107, Mar. 2000, doi: 10.1006/jmps.1999.1278.
[27] S. J. Kiebel, M. I. Garrido, R. Moran, C.-C. Chen, and K. J. Friston, “Dynamic causal modeling for EEG and MEG,” *Hum. Brain Mapp.*, vol. 30, no. 6, pp. 1866–1876, 2009, doi: 10.1002/hbm.20775.
[28] A. H. Williams, A. Degleris, Y. Wang, and S. W. Linderman, “Point process models for sequence detection in high-dimensional neural spike trains,” 2022.
[29] C. W. J. Granger, “Investigating Causal Relations by Econometric Models and Cross-spectral Methods,” *Econometrica*, vol. 37, no. 3, pp. 424–438, 1969, doi: 10.2307/1912791.
[30] B. Lindner, L. Auret, M. Bauer, and J. W. D. Groenewald, “Comparative analysis of Granger causality and transfer entropy to present a decision flow for the application of oscillation diagnosis,” *J. Process Control*, vol. 79, pp. 72–84, Jul. 2019, doi: 10.1016/j.jprocont.2019.04.005.
[31] S. L. Bressler and A. K. Seth, “Wiener–Granger Causality: A well established methodology,” *NeuroImage*, vol. 58, no. 2, pp. 323–329, Sep. 2011, doi: 10.1016/j.neuroimage.2010.02.059.
[32] A. Brovelli, M. Ding, A. Ledberg, Y. Chen, R. Nakamura, and S. L. Bressler, “Beta oscillations in a large-scale sensorimotor cortical network: Directional influences revealed by Granger causality,” *Proc. Natl. Acad. Sci.*, vol. 101, no. 26, pp. 9849–9854, Jun. 2004, doi: 10.1073/pnas.0308538101.
[33] C. W. J. Granger, “Economic processes involving feedback,” *Inf. Control*, vol. 6, no. 1, pp. 28–48, Mar. 1963, doi: 10.1016/S0019-9958(63)90092-5.
[34] J. Geweke, “Measurement of Linear Dependence and Feedback Between Multiple Time Series,” *J. Am. Stat. Assoc.*, vol. 77, no. 378, pp. 304–313, 1982, doi: 10.2307/2287238.
[35] A. Shojaie and E. B. Fox, “Granger Causality: A Review and Recent Advances,” *Annu. Rev. Stat. Its Appl.*, vol. 9, no. Volume 9, 2022, pp. 289–319, Mar. 2022, doi: 10.1146/annurev-statistics-040120-010930.
[36] S. Kim, D. Putrino, S. Ghosh, and E. N. Brown, “A Granger Causality Measure for Point Process Models of Ensemble Neural Spiking Activity,” *PLOS Comput. Biol.*, vol. 7, no. 3, p. e1001110, Mar. 2011, doi: 10.1371/journal.pcbi.1001110.
[37] S. Ghosh, D. Putrino, B. Burro, and A. Ring, “Patterns of spatio-temporal correlations in the neural activity of the cat motor cortex during trained forelimb movements,” *Somatosens. Mot. Res.*, vol. 26, no. 2, pp. 31–49, Jun. 2009, doi: 10.1080/08990220903098308.
[38] A. Casile, R. T. Faghih, and E. N. Brown, “Robust point-process Granger causality

analysis in presence of exogenous temporal modulations and trial-by-trial variability in spike trains," *PLoS Comput. Biol.*, vol. 17, no. 1, p. e1007675, Jan. 2021, doi: 10.1371/journal.pcbi.1007675.
[39] T. Schreiber, "Measuring Information Transfer," *Phys. Rev. Lett.*, vol. 85, no. 2, pp. 461–464, Jul. 2000, doi: 10.1103/PhysRevLett.85.461.
[40] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 4, pp. 623–656, Oct. 1948, doi: 10.1002/j.1538-7305.1948.tb00917.x.
[41] M. Ursino, G. Ricci, and E. Magosso, "Transfer Entropy as a Measure of Brain Connectivity: A Critical Analysis With the Help of Neural Mass Models," *Front. Comput. Neurosci.*, vol. 14, Jun. 2020, doi: 10.3389/fncom.2020.00045.
[42] R. Vicente, M. Wibral, M. Lindner, and G. Pipa, "Transfer entropy—a model-free measure of effective connectivity for the neurosciences," *J. Comput. Neurosci.*, vol. 30, no. 1, pp. 45–67, Feb. 2011, doi: 10.1007/s10827-010-0262-3.
[43] M. Lindner, R. Vicente, V. Priesemann, and M. Wibral, "TRENTOOL: A Matlab open source toolbox to analyse information flow in time series data with transfer entropy," *BMC Neurosci.*, vol. 12, no. 1, p. 119, Nov. 2011, doi: 10.1186/1471-2202-12-119.
[44] S. Nigam *et al.*, "Rich-Club Organization in Effective Connectivity among Cortical Neurons," *J. Neurosci.*, vol. 36, no. 3, pp. 670–684, Jan. 2016, doi: 10.1523/JNEUROSCI.2177-15.2016.
[45] D. P. Shorten, R. E. Spinney, and J. T. Lizier, "Estimating Transfer Entropy in Continuous Time Between Neural Spike Trains or Other Event-Based Data," *PLOS Comput. Biol.*, vol. 17, no. 4, p. e1008054, Apr. 2021, doi: 10.1371/journal.pcbi.1008054.
[46] A. Rozo *et al.*, "Benchmarking Transfer Entropy Methods for the Study of Linear and Nonlinear Cardio-Respiratory Interactions," *Entropy*, vol. 23, no. 8, p. 939, Jul. 2021, doi: 10.3390/e23080939.
[47] M. Staniek and K. Lehnertz, "Symbolic Transfer Entropy," *Phys. Rev. Lett.*, vol. 100, no. 15, p. 158101, Apr. 2008, doi: 10.1103/PhysRevLett.100.158101.
[48] A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," *Phys. Rev. E*, vol. 69, no. 6, p. 066138, Jun. 2004, doi: 10.1103/PhysRevE.69.066138.
[49] A. Citri and R. C. Malenka, "Synaptic Plasticity: Multiple Forms, Functions, and Mechanisms," *Neuropsychopharmacology*, vol. 33, no. 1, pp. 18–41, Jan. 2008, doi: 10.1038/sj.npp.1301559.
[50] M. Wibral, R. Vicente, and J. T. Lizier, Eds., *Directed Information Measures in Neuroscience*. in Understanding Complex Systems. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014. doi: 10.1007/978-3-642-54474-3.
[51] A. Tauste Campo *et al.*, "Feed-forward information and zero-lag synchronization in the sensory thalamocortical circuit are modulated during stimulus perception," *Proc. Natl. Acad. Sci.*, vol. 116, no. 15, pp. 7513–7522, Apr. 2019, doi: 10.1073/pnas.1819095116.
[52] J. J. Harris, E. Engl, D. Attwell, and R. B. Jolivet, "Energy-efficient information transfer at thalamocortical synapses," *PLOS Comput. Biol.*, vol. 15, no. 8, p. e1007226, Aug. 2019, doi: 10.1371/journal.pcbi.1007226.
[53] K. E. Schroeder *et al.*, "Disruption of corticocortical information transfer during ketamine anesthesia in the primate brain," *NeuroImage*, vol. 134, pp. 459–465, Jul. 2016, doi: 10.1016/j.neuroimage.2016.04.039.
[54] P. Ram and K. Sinha, "Revisiting kd-tree for Nearest Neighbor Search," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Anchorage AK USA: ACM, Jul. 2019, pp. 1378–1388. doi: 10.1145/3292500.3330875.
[55] S. Dutta and F. Hamman, "A Review of Partial Information Decomposition in Algorithmic Fairness and Explainability," *Entropy*, vol. 25, no. 5, Art. no. 5, May 2023, doi: 10.3390/e25050795.
[56] S. L. Keeley, D. M. Zoltowski, M. C. Aoi, and J. W. Pillow, "Modeling statistical dependencies in multi-region spike train data," *Curr. Opin. Neurobiol.*, vol. 65, pp. 194–202, Dec. 2020, doi: 10.1016/j.conb.2020.11.005.