THE UNIVERSITY OF CHICAGO

MODELING AND CHARACTERIZING POPULATION-SCALE NEURAL ACTIVITY WITH APPLICATIONS TO EPILEPSY

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES
AND THE PRITZKER SCHOOL OF MEDICINE
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

COMMITTEE ON COMPUTATIONAL NEUROSCIENCE

BY
GRAHAM AUSTIN SMITH

CHICAGO, ILLINOIS
DECEMBER 2022

To Grandma

You didn’t get to see this,

But you lived leaving no doubt

You wouldn’t be surprised.

No, no, nothing in the world can be imagined in advance, not the slightest thing. Everything comprises countless unique facets that are impossible to foresee. In imagination, we rush past them, unaware. Realities, though, are slow and indescribably detailed.

Rainer Maria Rilke, *The Notebooks of Malte Laurids Brigge*

“Stupidity may be encouraged. If a tactic does not work, I suggest employing methods to *make* it succeed before abandoning the idea outright. A fool uses a spear to simply jab; every tool can be utilized in many ways.”

Fehotep in *The Wandering Inn* (8.08 J) by pirateaba

Table of Contents

List of Figures

List of Tables

Acknowledgments

I thank my committee, for approving this dissertation, and for guiding me in its development:

My committee chair, Nicho Hatsopoulos, was my first principal investigator (PI) during my undergraduate research (also at UChicago) where, as my sister puts it, I "trained a monkey to control a robot with her mind." More precisely, I provided positive reinforcement (i.e. fruit juice) to a monkey performing a reach-to-grasp task using a free-standing robotic arm controlled by a microelectrode array implanted in her motor cortex. This remains the coolest experiment I've ever been a part of, and brain-computer interfaces remain my number one long-term interest. I also thank Zizou (the monkey) and Karthikeyan Balasubramanian (the postdoc) for being understanding and encouraging in my fledgling attempts at science.

Continuing to thank my committee, Nicolas Brunel, formerly of UChicago, now of Duke, has always provided incisive and relevant feedback on my work, and especially with close readings of my mathematics. He was my first rotation PI in graduate school, directing me in a pure theory project to model the dynamics of neural networks including spike-timing-dependent plasticity (commonly known as STDP). I didn't quite succeed, but working with him was so fascinating that it convinced me to switch to theory-focused work, a choice that clearly shaped the entirety of my graduate studies, this dissertation, and likely all my future work.

My final non-advisor committee member was Stephanie Palmer, who went far,

far beyond what is expected of a committee member in her support, both scientifically and personally. This was true in so many ways, but to choose two to be emblematic: she talked me through many problems in my middle-years on walks with her dogs around Hyde Park; she also invited me to TA for her at MBL in Woods Hole every year, which was a hybrid fantastic opportunity to lecture on a subject I'm passionate about (reproducible programming), get to know the incoming PhD students, and have a beautiful week-long vacation in Cape Cod at the end of the summer. She was an advisor in all but name.

So many other professors were helped me along, and to name just a few: Peggy Mason, who had the important role of telling me the perhaps obvious thing, "You need to take neuroscience classes if you want to do a neuroscience PhD," when I was still a math major undergrad; Vicky Prince, who was my Dean of Students for most of graduate school and all of my time on the Dean's Council, incredibly supportive, and also my MBL buddy and the only other person to attend qBio every year; Jason MacLean, who taught an incredible course on the neocortex where he single-handedly taught me how to read (and pick apart) academic papers, helped me figure out what graduate programs to apply to and why, and who then IDP-mentored me throughout graduate school when I ended up choosing UChicago.

Speaking of professors, before them there were so many teachers, every single one of whom I want to thank. I can scarcely fathom how lucky I've been to have had nothing but fantastic teachers, from preschool all the way through high school. I hesitate to name any because I want to name them all. To begin with, Mary Anderson, my first and fourth grade teacher. I don't remember much of elementary school, let alone first grade, but what I do remember couldn't be more important: first, she played match-maker for me and who would become one of my best friends; second, at the beginning of the year, I was one of two under-performing students

in math. By the end of the year I was “exceeding expectations” and now I have a PhD mostly based on math. I have no idea how she did that, but I will be forever grateful.

In fifth grade, Mark Olbin was the kind of teacher you only see in very good, very unbelievable movies. He made it impossible not to love every subject, with his various personas who brought levity, insanity, and mystery to learning. Above all the many things he taught me, I’ve carried with me one motto: “You say weird like it’s a bad thing.” Seven plus years studying computational neuroscience has been weird in the best way.

In middle school, I’m going to take a seeming left turn and specifically thank my English teacher, Alan Schultz, and my Latin teacher, Barbara Croft, who were a synergistic tag-team who laid an adamantine grammatical foundation that frustrates my attempts to read mediocre fantasy to this day. Earning a PhD requires so much critical reading and writing, both of which have been made much easier by the command of grammar that they instilled. (A PhD also requires questioning authority, which is why I now use the parentheticals that Mr. Schultz so abhorred.)

In high school, I’ll continue that left turn and again thank an English teacher, Kris Tully. When one lunch period I told her that I planned to major in math or computer science, she said, aghast, “What a waste.” That reflexive yet heartfelt comment has stuck with me for the past twelve years, a vote of absolute confidence in my writing abilities that I held through my aggressively quantitative undergrad and grad studies, and which I’m now taking to heart by taking a year off to write a fantasy novel.

After that nostalgic detour, we’ll jump back to my graduate program, where I thank Elena Rizzo for being an incredible program director throughout my entire time here, always quick to reply and address problems, and always happy to just

chat.

I also thank the Pritzker Endowment for the Neurosciences, from which I was awarded in my second year and which continued to fund me through the duration of my studies.

During my time in the computational neuroscience program, I've sat in Wim van Drongelen's lab, which has been incredibly supportive of my sometimes divergent research interests and always welcoming. Now, of course, I'm one of them, so that feels weird to say. They've become good friends along the way.

I thank especially Wim, initially for welcoming me into his lab, and then subsequently for stepping up as my co-advisor as Jack transitioned to retirement. From the beginning he was supportive of my non-epilepsy research and Julian proclivities, and ultimately critical when I moved to epilepsy applications. The triple correlation work in particular sprang from his insight in the summer of 2021 and I am incredibly grateful to him for inviting me to realize that insight. I'm not sure how many beers I owe him at this point, but it's a lot.

For my final academic acknowledgement, I thank my advisor, Jack Cowan. While my rotation with Nicolas was priming me to move more towards theory work, I was also in Jack's highly theoretical class, Mathematical Modeling of Large-Scale Brain Dynamics. I hadn't thought Jack would take me as a student—I'd done no theoretical neuroscience at that point—but midway through my second quarter he started class by saying, "Graham, I've been thinking about thesis projects for you." After that, he was my advisor, pulling me fully over the cliff into theory research. The initial project we chose—an ambitious attempt to derive an exact solution to the Wilson-Cowan equations—didn't quite materialize, but along the way Jack taught me to think big, while also keeping it simple, and always to argue with him. Thank you Jack. It was an honor and a pleasure to work with you.

Beyond academics, I thank my friends and family. They made the experience of this PhD a much happier one. CMAC (The University of Chicago Glee Club) was a particular source of joy, along with everyone in it. A special thanks to my roommate, Daniel Schwartz, who put up with me for seven years, and then let me sleep on his couch in Cambridge for three weeks during the final stretch of writing this document.

Last but certainly not least, my family:

Thank you, Hannah, for trying your best to keep me humble, and for flying several thousand miles to support me during my defense.

Thank you, Dad, for first teaching me math: for going way beyond what you needed to teach, and then backing way off when I burst into tears because I couldn't understand improper fractions. Thank you for listening as you sanded, to my ramblings: about dreaming, about lasers, even about BCIs. You made me feel like I could learn whatever I wanted, and that whatever I wanted to learn would be interesting.

Finally, thank you, Mom, for teaching me to use a spoon. (Among other things.) Thank you for bearing with me when, at 10:30 at night, abruptly full of energy, I would follow you around chattering as you got ready for bed. And thank you also for "a breathing body and a beating heart, strong legs, bones and teeth, and two clear eyes to read the world."[1] Those made things easier.

1 Billy Collins, *The Lanyard*.

Abstract

Continuing advances in neural recording technologies and ambitions for their use necessitate improvements in our ability to meaningfully model and characterize activity at the scale of neural populations. Here I contribute two theoretical advances in this field. First, I theoretically characterize critical properties of certain focal seizures, prove that the existing standard Wilson-Cowan model (WCM) cannot generate activity with these properties, and then extend the WCM to have those properties. This enables modeling average population activity in focal seizures with the WCM. Second, I introduce a novel approach to numerically characterize any population activity in both time and space. This is simply third-order correlation (triple correlation) which I prove uniquely characterizes any finite dataset. Moreover I show triple correlation can be summarized by partitioning its underlying motifs into fourteen equivalence classes that embody well-known computational properties (e.g. synchrony, feedback). These motif classes reflect underlying changes in the structure of spatiotemporal neural activity. Finally, I apply this latter theoretical advance, using these motif classes to automate detection of seizures in newborn infants.

Chapter 0

Introduction

A fantastical, yet concrete, goal in neuroscience is the development of a general brain-computer interface (BCI), a device that enables communication directly between a brain and a computer for both rapid and intuitive control (of the computer) and information flow (for the brain). The consequences of tightening this loop are hard to overstate, analogous to the advent of the computer itself. However, the a vast expanse lies between us and BCIs. They will require engineering technology that can interact with hundreds of millions or billions of neurons (and perhaps more synapses) and developing an understanding of those neurons sufficient to interrogate the resulting cascade of data. I have aimed my research at taking a small step down the latter path, by studying the behavior of neural tissue at the population scale.

The "population" is a term with subtly divergent meanings, one referring to a population of neurons discretely firing, and the other referring to the averaged activity of a population. The former usually interrogates hundreds or thousands of neurons, while the latter concerns itself with tens of thousands or millions. Discrete population activity results from spike-sorted recordings with spatiotemporal resolution on the micrometer and sub-millisecond scale (e.g. 30kHz Utah arrays, optoge-

netic recordings), while averaged population activity models electrodes that record more slowly from millimeters or centimeters of cortex (e.g. $<$1kHz Utah array LFP, EEG, fMRI). While it would be quite convenient if we could base a BCI on numbers that each summarize hundreds of thousands of neurons, I don't make that assumption, rather assuming that studying the behavior such averages will guide our understanding of the behaviors of the underlying populations. Regardless, studying the dynamics of such averages has produced decades of fruitful research. So here, we will consider data of both types, first population averages in chapter 1, then a technique that applies to both in chapter 2, and then finally applications of that technique to population average data in chapter 3.

Just as the concrete BCI informs my long-term aims, the smaller pursuit within this dissertation were informed by a concrete problem. I contend that any good understanding of population-scale activity—whether discrete or averaged—should provide insight into the premier population pathology: the seizure. In one sense, seizures are quite simple: populations become over-active. However, that description belies the heterogeneity within a disease that affects over 50 million people worldwide, over 20 million of whom cannot be treated by existing anti-epileptic drugs Beghi et al (2019). In chapter 1, I aim to model an epilepsy subclass, focal seizures, with the hope to provide theoretical insights and advances that can lead to tangible improvements in the lives of people suffering from this disease. In chapter 3, I aim to detect notoriously heterogeneous neonatal seizures in a completely new way, with the hope of starting down a path that will improve outcomes for newborns suffering from seizures.

In the first research chapter (chapter 1), I develop a theory of focal epilepsy with failure of inhibition, and do so by diving into recent work that adapted the Wilson-Cowan model (WCM), a population-averaging model developed by my advisor fifty

years ago. In so doing, I explain why failure of inhibition (FoI) may be necessary for the ignition of a certain type of focal seizure, and also show that with the possibility of failure of inhibition, a single system can move from population-scale functionality to seizure without any change in model parameters. Thus the model can theoretically predict the effects of anti-epilepsy treatments on seizures *relative to* their effect on population-scale computation/functionality.

In the second research chapter (chapter 2), I introduce the use of triple correlation as a means for characterizing spatiotemporal neural activity. I expand upon prior work in image analysis to prove that the triple correlation of any neural dataset uniquely characterizes that data (in fact, any finite dataset). Inspired by work in network-motif triplets, I then derive a set of fourteen motif classes into which all elements of triple correlation fall, and which have intuitive relationships to widely studied neural processing properties, such as feedback, feedforward, and synchrony. To prove the utility of this metric, I use it to detect structured activity in the midst of substantial noise across a multiple trials responding to a pseudo-stimulus. Having shown that these metrics measure real structure in the presence of noise, I additionally show that this structure varies across the neural response to a much more pharmacological intervention in real data.

Finally, in the third research chapter (chapter 3) I apply triple correlation to the problem of automating detection of seizures in newborn infants. These seizures are notoriously difficult to detect, both in clinical practice because infants cannot report seizures and often do not present overtly obvious seizures, and also in electrophysiology, because the EEGs of neonatal seizures are quite heterogeneous. In a detector constructed from the triple correlation signal, we find performance on-par with that of a detector constructed from a clinical-standard signal, amplitude-integrated EEG (aEEG). Moreover, unlike the one-note aEEG signal, triple correlation actually

reflects the heterogeneity of neonatal seizures, suggesting the possibility of categorizing such seizures for the purposes of divining etiology or treatment.

Through these chapters I hope to both attack epilepsy along two distinct avenues, and also to advance methodologies in two different realms: both at the population level, but differing in level of abstraction, mechanistic underpinnings, and even breadth of applicability. To introduce these chapters, I will begin by outlining a brief history of neural population-average models. Then I will describe the state of automated seizure detection and the concerns involved with such efforts.

0.1 A history of the Wilson-Cowan equations in modeling neural populations

Population models can largely be conceived as falling into one of two categories, which I will refer to as "averaged" and "discrete." The former describes the activity of entire populations of neurons with a single number, whereas the latter describes the population activity as a vector. This is not a formal distinction, and indeed one of the more common current approaches to population modeling involves projecting a population vector onto a lower dimensional subspace, which constitutes a muddy middle ground between the two categories as stated. This divide has a long history, perhaps most notably described in 1970 by the dichotomy "randomness in the small, design in the large," (Anninos et al, 1970; Harth et al, 1970) reflecting the understanding that connections between individual neurons (entries in a connectivity matrix "connecting" the population vector) are random, whereas large-scale connections between the populations-writ-large (e.g. the average connectivity, or

perhaps the eigenvalues of the connectivity matrix) reflect more conserved[1] computational structure. We will be focusing on the "averaged" model, as that is the subject of the Wilson-Cowan model (WCM). However, to arrive at the population average, we must begin with the raw vector, which was indeed developed earlier.

The first theoretical model of neural populations was developed by McCulloch and Pitts (1943) at the University of Chicago. By utilizing the fact of the neuron's "all-or-nothing" response, McCulloch and Pitts defined the activity of a neural "net" in terms of propositional logic. They were able to show that certain competing theories about neural interactions (as fundamental as disagreements about the nature of inhibition) were actually in a certain sense equivalent. In many ways, this was the birth of the field of computational neuroscience[2], though the field would not come to be called as such for decades yet.

A decade later, Beurle (1956) derived one of the first models of population activity by means of averaging. Beurle built on McCulloch and Pitts' work, again using the critical "all-or-nothing" feature of neural activity, but rather than treating the "population scale" as discrete description of nigh-infinite neurons, he asked if it would be possible to zoom out and describe them all at once, or at least many of them. This approach originated in work by Shimbel and Rapoport (1948), who were the first to attempt a statistical reduction of McCulloch and Pitts' work. However, Beurle was able to take a critical subsequent step due to recent physiology from Sholl (1953), which described the statistically typical pattern of dendritic arborization in cat visual cortex. Sholl found that the probability of "connexion" between two neurons decays exponentially with the distance between them. Consequently

1 "conserved" may be in many ways, whether across time, across brains, or across brain areas performing analogous computations

2 This distinction depends one of several rather silly argument, such as that because preceding modeling was of single cells, it could as well have been computational biology. I include the claim here for aesthetic purposes.

neurons near each other could be expected to receive similar inputs, on average, as their connections could be expected to follow similar distributions. Thus one could derive a local neural population's average activation as a function of the population's average input. Beurle did this, and demonstrated interesting behaviors supported by his model—traveling waves of various forms, such as plane and spherical waves—that are still relevant to modern research (Lu et al, 2011; Sato et al, 2012; Harris and Ermentrout, 2018; Rapela, 2018).

His equations are

$$\frac{dF}{dt} = R\Phi - F \tag{1}$$

$$\frac{dR}{dt} = -F \tag{2}$$

where F is the firing rate, R is the proportion of sensitive neurons, and Φ is the proportion of cells reaching threshold. Note that though there are no explicit spatial derivatives, Φ and F are related by a spatial integration.

When the spatial relationship is sufficiently simple that $\Phi = mF$, then Beurle's equations have an exact solution. This is relevant because the form of spatial connectivity found by Sholl (1953) is the same one that would theoretically give a constant ratio between Φ and F: an exponential decay in connectivity with space. This gives the exact solution of a firing rate F of cells at location x and time t by

$$F(x,t) = \operatorname{sech}^2(x - vt) \tag{3}$$

where v is the constant wave velocity[3]. This is, in fact, a traveling solitary wave, not unlike that seen in a focal seizure Meijer et al (2015), and also not unlike traveling

3 There are some additional, critical constants that are nevertheless irrelevant here. See the original paper for the full solution.

waves seen in various other computational processes (Lu et al, 2011; Sato et al, 2012; Harris and Ermentrout, 2018; Rapela, 2018).

Less than two decades after Beurle, Wilson and Cowan (1972) developed the first example of the form of population-averaged model that is still used today (in particular within this dissertation): the Wilson-Cowan model (WCM). The WCM shares the integro-differential form of Beurle's model, both derived via integration over McCulloch-Pitts neurons. However, the WCM diverges from Beurle's model in

1. including inhibitory neurons,

2. defining firing as *at least* rather than *precisely* threshold input.

The latter difference means that Beurle models the derivative of neural activity, rather than the activity itself (roughly). This has implications for the shape of traveling waves supported by his model (e.g. solitary waves versus wavefronts), but his core results about wave propagation remain[4]. Mathematically, the latter difference resulted in a sigmoidal firing response curve rather than Beurle's $sech$ curve. The original version of the WCM is given by

$$\begin{aligned}\tau_E \frac{dE}{dt} &= -\alpha_E E + (1 - E)\, S_E(A_{EE}E - A_{EI}I + P_E) \\ \tau_I \frac{dI}{dt} &= -\alpha_I I + (1 - I)\, S_I(A_{IE}E - A_{II}I + P_I)\end{aligned} \tag{4}$$

where E and I are excitatory and inhibitory activity, respectively; t is time; α_i is the rate of decay of population activation[5]; $S_j(x)$ is a sigmoid function describing the firing rate of population j in response to input current x; A_{ij} is the strength of connectivity from j to i; and P_j is an external stimulus to population j.

4 Indeed more kinds of wave propagation are possible with the addition of inhibition.

5 NB: Wilson and Cowan (1972) neglected to mention that α_E should be set to 1.1 for their results, an omission that Dr. Cowan regrets, chuckling, to this day.

The results discussed in the initial paper largely centered around the ability of the WCM to generate oscillations, both decaying and not, and descriptions of the phase spaces that gave rise to these behaviors. The following year, Wilson and Cowan (1973) described a spatially-extended version of (1.4) and expanded the ability of that spatially extended model to generate localized limit cycles and traveling waves, the latter much like Beurle (1956). Unlike Beurle, despite much effort on the part of myself and Dr. Cowan, the WCM resists an exact solution[6].

Subsequent works used the WCM to explain observed neuroscientific phenomena. Perhaps most relevant to chapter 1, Ermentrout and Cowan conducted a systematic analysis of the phase space of the WCM (Ermentrout and Cowan, 1979a,b, 1980). In these works, theories were developed from the WCM to explain temporal oscillations in neural activity (Ermentrout and Cowan, 1979b) as well as visual hallucinations (Ermentrout and Cowan, 1979a). Quite recently, Cowan et al (2016) provided a good overview of such applications, including some novel results fitting traveling waves observed in visual cortex (i.e. Muller et al, 2014; Benucci et al, 2007; Nauhaus et al, 2012, 2009), in response to both strong and weak stimuli. With chapter 1, I seek to expand the boundaries of what can be explained by the WCM.

0.2 A glimpse into the future of clinical neurophysiology: Can we use machines to interpret EEG?

6 With a polynomial approximation of the sigmoid, one can transform the WCM into an equation like the integral of the Korteweg-de Vries equation, which is essentially the Beurle equation. However, the inhibitory population muddles things and at the end of the day you have a couple higher-order derivative terms that you'd really rather not have. To be fair, those terms do disappear as the wave speed goes to zero, but that's only partially helpful.

Adapted from: **Smith G**, Lee S, van Drongelen W. A glimpse into the future of clinical neurophysiology: Can we use machines to interpret EEG? Clin Neurophysiol. 2018 Oct; 129(10):2186-2187.

Diagnosing epilepsy requires a clinical neurophysiologist trained to conduct a subtle and time-consuming review of EEG recordings. Historically, researchers have attempted to improve this diagnostic process by identifying epileptiform markers and by automating their detection (e.g. Gibbs and Gibbs, 1952; Kellaway et al, 1959; Gotman and Gloor, 1976; Gotman, 1982). One recently developed approach that can extract and classify signal features is deep learning, an application of artificial neural nets with the neural elements arranged in multiple layers. Indeed, Tjepkema-Cloostermans et al (2018) employ such Deep Neural Networks (DNNs) to detect interictal epileptiform discharges (IEDs) in EEG data. In the following, we introduce a brief history of neural networks, discuss the findings in Tjepkema-Cloostermans et al (2018), and speculate on future directions of deep learning applications in this field. We generalize these insights to automated detection more broadly to include the triple-correlation-based detector developed in chapter 3.

Application of artificial neural nets as in Tjepkema-Cloostermans et al (2018) is not the first time that artificial intelligence (AI) has been used to automate complex tasks. An early example, the perceptron, was developed in the 1950s (Rosenblatt, 1958). The perceptron passes a weighted sum of all its inputs through a thresholding non-linear function. This function behaves as a hard- or soft-switch, similar to the formalisms used to model individual neurons (McCulloch and Pitts, 1943) or populations (Wilson and Cowan, 1972). Optimizing the weights associated with each of its inputs trains the perceptron. With increased computational resources, this approach was significantly improved by extending it to networks with multiple units arranged in multiple layers with feedforward connections (e.g. Pao, 1989), so that optimization occured on the weights of inputs from each layer to the next. The weights across the entire network are optimized with a back-propagation procedure that is based on the chain-rule in calculus (cf. van Drongelen, 2018). One prob-

lem with this approach is that such networks required significant pre-processing of the raw data, a task usually performed by a trained engineer with domain expertise. Another major issue is that these networks can be trained to respond almost perfectly to a training set, but often demonstrate poor performance on similar but different test data—a problem called overtraining. Furthermore, a fundamental scientific problem with the artificial neural network in general is that, even if it performs perfectly, it usually does not at all create insight into the underlying mechanisms that make it work. These problems undoubtedly contributed to the limited interest in and success of neural net applications in neurophysiology in the 1980s and 1990s.

A significant change in the AI-landscape of the neural network application occurred fairly recently. Vast improvements in network optimization and performance were created due to qualitative and quantitative changes in network structure, in the network's training processes, and in its non-linear component (reviewed in LeCun et al, 2015). Because of the increased layers (i.e. depth) in the network as compared to the initial neural network applications, this latest approach in artificial network technology is often indicated by terms as 'deep learning network' or 'deep neural networks'. A defining quality of a deep learning network is that its early layers learn a feature representation from the training data. Deeper layers extract progressively more complex and high-level features, until the deepest layers finally classify inputs using these high-level features. In a toy example from computer vision, the first layer would extract edges from pixels, deeper layers would extract shapes from edges, and the deepest layers would classify round shapes containing pentagons and hexagons leading to the detection of a soccer ball.

In their pilot study, Tjepkema-Cloostermans et al (2018) evaluate the ability of two types of networks for the detection of IEDs: the convolutional neural networks

(CNNs) with feedforward connectivity and the long short-term memory (LSTM) structure that also includes recurrent connections. Importantly, the deep networks have eliminated or at least dramatically simplified the problem of pre-processing data. Therefore, the authors detect IEDs by feeding 2-second epochs of the EEG signal directly into the network's input layer. All epochs were first reviewed by a clinical neurophysiologist for the occurrence of IEDs, providing a 'gold standard' for the automated detector. Next, the network's connectivity is adjusted by a training process. The goal of the training is to provide a network output that emulates the clinical neurophysiologist by discriminating between epochs with and without IEDs. Tjepkema-Cloostermans et al (2018) used 41,381 epochs with 1,478 IEDs to train the networks, and tested their performance on a different, but comparable, data set of 8,775 epochs with 337 IEDs. Comparing the detector to the 'gold standard', it demonstrated 47.7% sensitivity (the % of true positives detected) and 98.0% specificity (the % of true negatives detected). The false alarms were 0.6/min across the entire data set, and only 0.03/min in the normal EEGs. Although these results are very promising, the relatively low sensitivity does create a risk that a pathological EEG, especially when few IEDs are present, is incorrectly classified as normal. Clearly, the network's sensitivity does not match performance of manually optimized detectors for epileptiform activity (e.g. Gotman and Gloor, 1976; Gotman, 1982, software by Persyst, Prescott, AZ), and that aspect does require improvement. However, the enormous advantage of the DNN is that no subjective adjustments are required, and the DNN's sensitivity to detect IEDs may improve with further training, increase of network size, a more optimal network structure, and if needed with a slight relaxation of the rather high specificity.

If the AI techniques continue to be successful, whereto might deep learning lead us in the future? Although there are many potential applications in medicine, we

will limit our speculations to a few examples in the field of epilepsy.

A portable EEG device equipped with machine learning algorithms to anticipate or detect seizure onset could alert the patient to get to a safe location or to take prophylactic measures to stop the seizure before it begins. In fact, an FDA-approved neurostimulator that has some of these capabilities already exists (Neuropace, Mountain View, CA, USA). Although such interventions are emerging as promising tools for epilepsy management, they still remain narrow in their applicability and limited in their effectiveness.

Beyond functioning as a simple detector for epileptiform activity, a DNN could serve as a powerful diagnostic tool for differentiating epilepsies. Practice standards provide guidelines for physicians when choosing drugs to treat specific types of seizures, but the reality is that many patients struggle through a gamut of medication combinations before achieving adequate seizure relief. Machine learning techniques have the potential to detect subtle epileptiform patterns and use these features to create new and more specific categorization of seizure disorders. With these categories, physicians could potentially offer more precise and evidence-based treatment plans better individualized for each patient. Such targeted approaches may also help to identify patients earlier who will not respond to drug regimens and curb the use of extreme and invasive interventions on patients who are unlikely to benefit from surgery.

With the advent of advanced algorithms that perform faster and more reliably than a human physician, a future of fully automated medicine is an enticing idea. The goal of replacing the neurophysiologist by a machine is a lofty one, and still far from realistic with the current state of technology. A more modest goal would be to use algorithms to make the review process more efficient, thereby creating more time for higher-level tasks that require a more human touch such as patient

counseling. In a medical system where facetime with physicians is becoming increasingly scarce, automation of purely technical tasks will give doctors more time to focus on patient care.

Chapter 1

Failure of Inhibition Bridges Focal Seizure and Population-Scale Computation

1.1 Introduction

Epilepsy—a disease affecting more than 0.6% of the global population (Beghi et al, 2019)—is defined by repeated seizures, which consist in runaway neural population activation. Despite that simple generalization, seizures are heterogeneous, broadly categorized into focal or generalized, and even more finely subcategorized due to a range of clinical and electrophysiological expressions. The concern of this paper is the focal seizure, which is defined by a spatially localized onset, theoretically the epileptogenic zone (EZ). The EZ has a surgical definition as the minimal area needed to be resected to rescue seizures (Jehi, 2018; Lüders et al, 2006; Andrews et al, 2020). However, one can never truly know if the resection was "minimal" and

This chapter is under review as a manuscript of the same name, co-authored with Jack D Cowan, Stephanie E Palmer, and Wim van Drongelen.

so in practice the EZ remains elusive, so researchers and clinicians pursue other, more practical, concepts more suited for their own goals. With the goal of treatment and preference for minimally invasive tools, clinicians investigate the seizure onset zone (SOZ), which is defined as the area recorded by the first electrode to detect seizure activity. With the goal of understanding and with much more invasive tools at their disposal, researchers have recently identified a "focal core," characterized by intense spiking activity and by decreased inhibitory spiking (Schevon et al, 2012). In this paper, we will focus on this latter concept of a focal core. We will use the addition of this decreased inhibitory spiking—modeled by failure of inhibition (FoI)—to show how the cortex underlying a focal core can support both functional computation and seizure, to explain why focal seizures are not easily rescued by long-range homeostatic mechanisms, and finally we will use that insight to argue that FoI may be a necessary feature of an EZ in some focal seizures.

A near universal criterion for the focal core is persistent hyper-excitation. Additionally, mounting evidence points to a key factor in the seizure core's hyper-excitation: a local decrease in inhibitory activity (Meijer et al, 2015). Most likely, this decrease in inhibition is actually a *failure* of inhibition (FoI), meaning that the inhibitory cells are receiving supra-threshold input currents, yet failing to fire. Researchers debate the underlying cause, but a leading hypothesis is depolarization block (DB), which occurs when sustained input to a neuron creates an ion gradient that does not allow sodium channels to reset after action potentials, thus blocking future action potentials even in the presence of supra-threshold input (Bianchi et al, 2012; Kim and Nykamp, 2017, 2014; Meijer et al, 2015). DB preferentially affects inhibitory neurons, as modeling shows that smaller neurons such as many inhibitory neurons have lower DB thresholds (cf. Fig. 1.1a, reproduced from Tryba et al, 2019). From this preferential effect, a seizure follows logically: if inhibitory

neurons are the brakes on neural activity and excitatory neurons are the accelerators, FoI cuts those brakes while leaving the accelerators running, and so excitatory activity grows beyond control. Note that this failure need not (and does not) occur in the larger area of seizing cortex (the penumbra): most seizing cortex oscillates due to local interactions between excitation and inhibition, driven by long-range input from the focal seizing core (Schevon et al, 2012). Herein we will focus on this seizing core exclusively, modeling its behavior in the face of failing inhibition.

To model the focal core, then, we need a population model that models FoI. However, nearly all existing population models make the simplifying assumption that neural populations respond monotonically to input current (e.g. Wilson and Cowan, 1972). For normal physiological activity, this is usually a valid assumption (though see Bianchi et al, 2012, for a physiological counter-example), but not in pathological high-input cases where phenomena like DB are possible. Two recent papers have proposed substitutions for the monotonic input-response function: one a Gaussian (Meijer et al, 2015), one a product of two sigmoids (Kim and Nykamp, 2017). Here for the first time we provide a derivation from first principles for a similar FoI population model, hence with population parameters directly rooted in neuron-level electrophysiology.

Given this model of the focal core, seizures are clear: they are activity states with high excitation and near-zero inhibition. A potentially epileptic system, then, must have a stable seizure state. Past works typically stopped here, establishing an epileptic binary: a model was epileptic if it has a stable seizure state and non-epileptic if it did not (e.g. Meijer et al, 2015; Kim and Nykamp, 2017). We argue to nuance this epileptic binary: First, we propose investigation of both seizure susceptibility and rescuability, to establish a gradient of "epilepticness." While such a gradient is obvious at the level of clinical observation, it is not *a priori* obvious

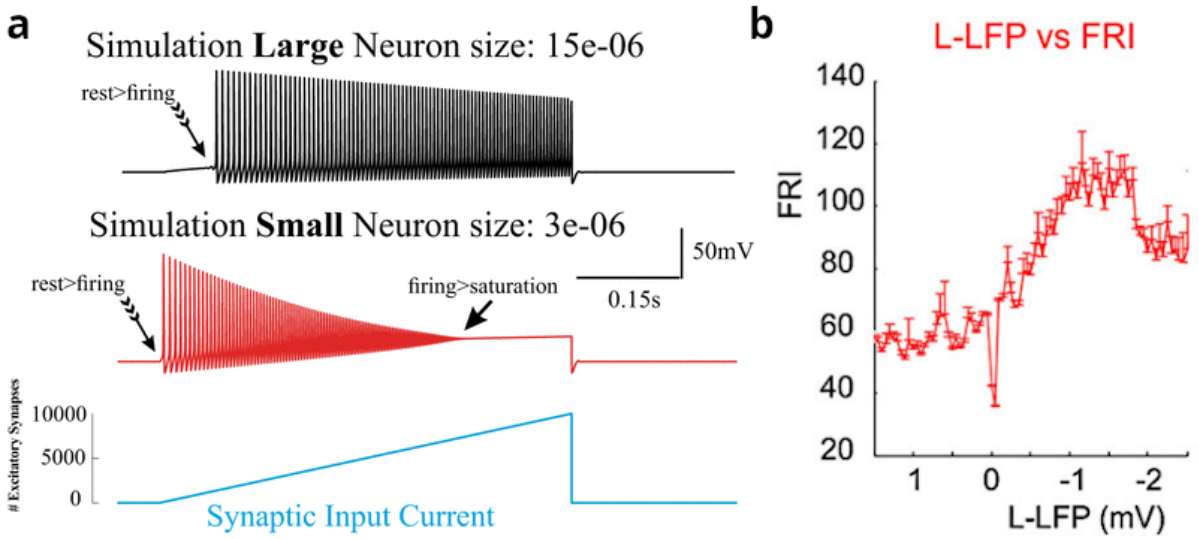


Figure 1.1: ***In silico* and *in vivo* failures to fire.** *Left*, a simulation of Hodkgin-Huxley model neurons with a putative excitatory neuron (i.e. large), *top black*, and a putative inhibitory neuron (i.e. small), *middle red*, both firing in response to the same input current, *bottom blue*. In response to increasing input, the excitatory neuron shows a monotonically increasing firing rate, with only minor to spike amplitude adaptation. In contrast, in response to the same input, the inhibitory neuron first increases firing rate, but finally fails to fire. Neuron sizes are diameters in meters. Left panel reprinted with permission from Tryba et al (2019) *Right*, experimentally recorded firing rate (FRI) plotted against low-frequency local field potential (L-LFP, an analogue of input current) shows that for sufficiently high input current, firing rate decreases. The data cannot distinguish between excitatory and inhibitory spikes, so the FRI includes both kinds of spikes. Right panel reprinted with permission from Meijer et al (2015).

that the gradient exists at the level of the focal core, as our model predicts. Second, we explicitly name qualitative difference in rescuability as the key explanation for the necessity of FoI in focal seizures, thus giving explicit logical justification for distinguishing between hyper-exited states with and without FoI. We argue that this reasoning strongly suggests that FoI would be present in at least one kind of EZ. Finally, we argue for the study of a subset of our models that are *tristable*, capable of both population-level functional computation (characterized by the existence of distinct physiological rest and activity states) and seizure. These tristable models allow for the possibility that epileptic cortical populations can function normally when not seizing without the need for any wholesale shift in neural parameters (compare to Saggio et al, 2020). Further, a tristable model can quantify seizure interventions (such as anti-epileptic drugs, AEDs) in terms of not only their potential to prevent and rescue seizure, but also in terms of their relative impact on cortical function, a critical clinical consideration.

1.2 Results

Our results come from a population model nearly identical to the WCM except in its inhibitory nonlinearity, just like two past models (Kim and Nykamp, 2017; Meijer et al, 2015). To begin, we will prove why such a modification is necessary. Then, we want to show that our modified model is consistent with experimental recordings of focal seizure core electrophysiology, and also that our model fits simple test data at least as well as prior FoI models.

1.2.1 Proof: WCM cannot model a functional focal core

Suppose we have an equation of the form of the WCM (1.4), which is to say a model with E and I populations, where the derivatives are proportional to the difference between a decay term and an activation term, and where the activation term is given by a non-negative monotonic function S whose input is a weighted difference between E and I. For simplicity, assume the external inputs P are zero, and all other parameters are equal to 1[1].

Suppose we have a fixed point (X_E, X_I), i.e. a point where $\frac{dE}{dt} = \frac{dI}{dt} = 0$. We want to show that there cannot exist a fixed point (Y_E, Y_I) such that $Y_E > X_E$ and $Y_I < X_I$.

Assume for contradiction that we have such points. Then we also have that arguments to the nonlinearity such that $Y_E - Y_I > X_E - X_I$. Then the firing function, being monotonic, preserves the inequality

$$S(Y_E - Y_I) \geqslant S(X_E - X_I). \tag{1.1}$$

Additionally $1 - Y_I > 1 - X_I$ (and both non-negative), so we have shown that the full activation terms are $(1 - Y_I)\, S(Y_E - Y_I) \geqslant (1 - X_I)\, S(X_E - X_I)$.

At a fixed point, the decay equals the activation, i.e. $(1 - Y_I)\, S(Y_E - Y_I) = Y_I$ and $(1 - X_I)\, S(X_E - X_I) = X_I$, so, combined with the above, implies $Y_I \geqslant X_I$, contradicting our initial assumption. Thus it is impossible to have two fixed points (X_E, X_I) and (Y_E, Y_I) such that $Y_E > X_E$ and $Y_I < X_I$ given a monotonic firing function.

If the firing function were not monotonic, then (1.1) would not hold, and consequently it would be possible to have two such fixed points.

See subsection 1.3.1 for our derivation of a model with such a firing function.

1 The argument holds for any P and as long as all other parameters are non-negative, which is necessary for the biophysical meaning of the equation. If the connectivity strengths were negative, then the populations not be excitatory and inhibitory.

1.2.2 Derived model consistent with prior work

Fig. 1.1 (right panel) displays a key hallmark of an FoI input-response curve: plateaus in response where small changes in input do not change the response, suggestive of saturation. The first, higher plateau indicates normal saturation: no more cells can be recruited. The second, lower plateau indicates saturation of the excitatory population combined with failure of the inhibitory population. Logically, these plateaus should exist: the first is a safety margin for physiological activity, and also has been noted by researchers for decades (e.g. Wilson and Cowan, 1972). The second follows if FoI truly preferentially affects inhibitory neurons: once the inhibition cuts out, all that remains are the already-maximally-firing excitatory neurons. The evidence for the preferential effect of FoI can be seen in the left panel of Fig. 1.1: Hodgkin-Huxley simulated neurons show a point in ramping input where inhibitory neurons fail to fire but excitatory cells continue (Meijer et al, 2015). Note also that the difference in plateau heights in the right figure is a 20% reduction, suggestive of the usually assumed inhibitory cell prevalence of 20%.

In Fig. 1.2 we see that our model exhibits such plateaus. This sharply contrasts another mechanism for "paradoxical" input-response, supersaturation, which does not plateau, but rather coincides with a decrease in excitatory activity until excitatory activity can decrease no further, at which point inhibitory activity increases (Sanzeni et al, 2020).

Another feature of Fig. 1.1b is the asymmetry of the left and right slopes. Unlike the plateaus, this is not a necessary feature (given our lack of understanding of the underlying distribution of failure thresholds). However, it is what we empirically observed and we had no *a priori* reason *not* to expect asymmetry. So, in comparison to a model with a necessarily-symmetric Gaussian nonlinearity, we consider it a mark in favor of our model that our model does fit asymmetric input-response

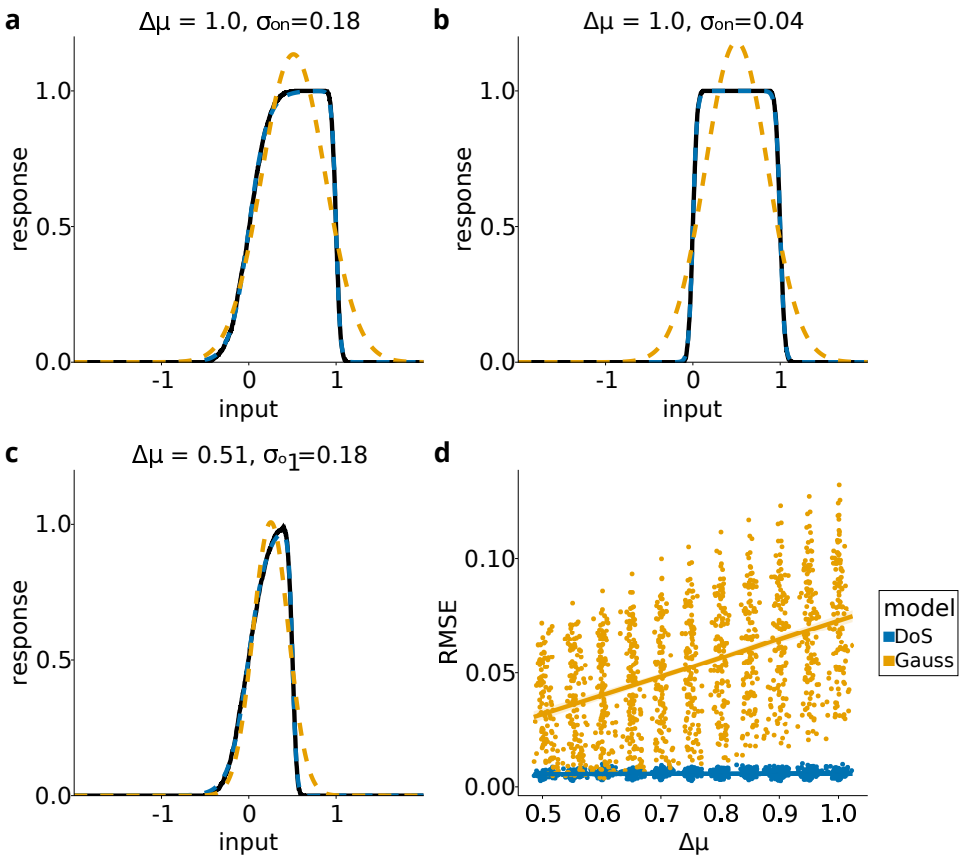


Figure 1.2: **Comparing model fits of synthetic data.** Sigmoids fit data qualitatively better than a Gaussian when either 1) $\Delta\mu = \mu_{\text{fail}} - \mu_{\text{fire}}$ is large relative to either σ_{fire} or σ_{fail} (i.e. the curve plateaus), or 2) when $\sigma_{\text{fail}} \neq \sigma_{\text{fire}}$ (i.e. the curve is asymmetric). **A,B,C**, Input response curves (solid black) generated by sampling neurons' firing and failing thresholds from "fire-then-fail"-constrained Gaussian distributions. Each input-response curve was generated analogous to Figure 1.5: 1000 binary neurons were created by drawing firing and failing thresholds from normal distributions, $\theta_{\text{fire}} \sim \mathcal{N}(\mu_{\text{fire}}, \sigma_{\text{fire}})$ and $\theta_{\text{fail}} \sim \mathcal{N}(\mu_{\text{fail}}, \sigma_{\text{fail}})$. Then any neurons with $\theta_{\text{fail}} \leqslant \theta_{\text{fire}}$ were discarded and resampled to enforce the biophysical constraint $\theta_{\text{fire}} < \theta_{\text{fail}}$. The dashed lines indicate the best fit by either difference-of-sigmoids (DoS, dashed blue) or Gaussian (Gauss, dashed orange) models. Note that DoS model overlaps both the ground truth and product-of-sigmoids model (latter not shown). Titles indicate sample statistics of the "fire-then-fail"-constrained Gaussian distributions. **D** Mean squared error (MSE) of fits to input-response curves as a function of difference $\Delta\mu$ between sample means of the underlying on and off threshold distributions, and superimposed linear fits of MSE vs $\Delta\mu$. Shading indicates 95% confidence interval (barely wider than line). Statistics not indicated by subplot titles, **A,B,C:** $\sigma_{\text{fail}} = 0.04$. **A:** $\text{RMSE}_{\text{Gauss}} = 0.10$, $\text{RMSE}_{\text{DoS}} = 0.005$. **B:** $\text{RMSE}_{\text{Gauss}} = 0.13$, $\text{RMSE}_{\text{DoS}} = 0.003$. **C:** $\text{RMSE}_{\text{Gauss}} = 0.06$, $\text{RMSE}_{\text{DoS}} = 0.007$. **D:** $\mu_{\text{fire}} = 0$, $\mu_{\text{fail}} \in [0.5, 1.0]$, and both $\sigma_* \in [0.04, 0.24]$.

curves.

Generally Fig. 1.2 illustrates that our model fits potentially asymmetric and potentially plateauing input-response functions better than a model with a Gaussian nonlinearity and just as well as one with a product-of-sigmoids nonlinearity.

1.2.3 Seizures in phase space

At its most basic, a seizure is heightened neural activity that is not immediately controlled, with pathological consequences. Within a neural population model, unitless and lacking clinical consequence, it can be hard to distinguish between heightened activity states that are computational versus epileptic. In general, this would need careful fitting of model parameters to biological analogues and resulting electrophysiology. However, in our particular case of focal seizures with failure of inhibition, we can be more categorical: we can find states that a) have minimal inhibition compared to excitation and b) are qualitatively more difficult to rescue.

A nuance of focal seizures is that they involve a *propagation* of heightened neural activity—a traveling wave. Traveling waves are not possible in point-models like the one described herein, as traveling requires spatial extent. However, following Harris and Ermentrout (Harris and Ermentrout, 2018) we can imagine our point-model as space-clamped; meaning, our single point represents activity that is constant across some spatial extent. Since a traveling wave is a transition from one space-clamped state to another, we can imagine a transition between fixed points in our model as a traveling wave. An additional nuance is that traveling waves in focal seizures tend to be solitary waves (leaving refractory cortex in their wake) rather than wavefronts (leaving still-pathologically-heightened activity in their wake) (Schevon et al, 2012). We note that regardless of the mechanisms for this refraction, the seizure will not cease until the wave itself is controlled, be it

solitary or wavefront. So we are concerned with the state that constitutes the peak of the wave, which we call the seizure state.

We quantify states in two ways: first, simply by excitatory activity (E). Second, by a metric we call the "seizure index" (SI). Since we are concerned only with focal seizures, we want our seizure index to capture the contrast between excitatory activity and inhibitory activity. Since in addition to high contrast, seizures have high excitatory activity, we further weight this contrast by the maximal population activity.

$$SI = \frac{E - I}{E + I} \max(E, I) \tag{1.2}$$

This index varies between -1 and 1, where $SI = -1$ corresponds to complete failure of excitation with strong inhibition (impossible in our model without external stimulation); $SI = 0$ corresponds to balanced excitation and inhibition; and $SI = 1$ corresponds to complete failure of inhibition with strong excitation.

For focal seizures in particular, literature shows that the core activity consists in heightened excitatory activity and depressed inhibitory activity (plus penumbral oscillations which here we disregard) (Meijer et al, 2015). So a seizure state is a high-SI state, which occur most ideally along the E-axis of the phase space. For our search, consider Fig. 1.3. We can discount the stable point near (on) the origin: this is a rest state with both low-E and low-SI, so is certainly not a seizure state. The FoI model (right panel) has a high-E, high-SI fixed point on the E-axis, at the intersection of the final arms of both the E- and I-nullclines. We will refer to this as a final-arm seizure state. For this same parameterization, the purely-firing model (left panel) has no other stable fixed point on or near the E-axis. In general, purely-firing systems only have another stable fixed point on the E-axis if the I-nullcline (dashed red) shifts so that the third arm of the E-nullcline (solid blue) intersects on the first arm of the I-nullcline (before rising). This is a high-SI state, so purely on

the basis of electrophysiology, we might consider such a state to be seizure. We will refer to it as a first-arm seizure state. We will argue in the next section that first-arm and final-arm seizure states are fundamentally different, and that first-arm seizure states may not even be seizures, even though the states in isolation may appear identical.

In sum, we will refer to stable states near the origin as rest states; stable (or limit cycle) states on the second (rising) arm of the I-nullcline as active states (rising dashed red); stable zero-E states on the first arm of the FoI or purely-firing I-nullcline as first-arm seizure states (horizontal dashed red); stable states on the final arm of the FoI I-nullcline as final-arm seizure states (horizontal dashed red); and stable states between active and final-arm seizure states as either high-active states or quasi-seizure states.

1.2.4 Epileptic systems

An epileptic system is one that can transition to a seizure state. Traditionally this includes slow parameter transformations of the entire phase space (Saggio et al, 2020), but we will only concern ourselves with transitions *within* a phase space, and so only call a system epileptic if it has a seizure state in its phase space (e.g. Fig. 1.3, right panel). We do not claim this applies to all or even most epilepsies. This is possible because we restrict ourselves to focal epilepsy and our FoI model of focal epilepsy permits both computation and seizure within the same system. Thus, within a single fixed system, we can model epileptic systems (foci) that function normally until seizing in response to electrophysiological stimulus. As with seizure states, we refer to two broad categories of epileptic systems: first-arm and final-arm.

The first-arm epileptic system can arise from either FoI or purely-firing models.

It can only be bistable, and specifically has a rest state in addition to the eponymous first-arm seizure state. In these systems, the E-support of the E-nullcline is a subset of the E-"support"[2] of the first arm of the I-nullcline. Consequently the E-nullcline cannot intersect with the I-nullcline on any arm but the first. Thus the fixed points of a first-arm epileptic system are independent of whether the underlying model includes FoI, since the first arm is unaffected by FoI.

In contrast, a final-arm epileptic system must include FoI, because the final arm of the I-nullcline only returns to the E-axis given FoI (as in Fig. 1.4, right panel). Final-arm epileptic systems have a rest state, a final-arm seizure state, and up to two active stable states, and so they are bistable, tristable, or tetrastable. Since a bistable final-arm epileptic system has a rest state, it may be functional at a sub-population scale. A tristable system could add to the bistable system either an active state or a high-active state. The active state has somewhat balanced E and I activity and so transitions between it and the rest state could be involved in population-scale computation (e.g. Sanzeni et al, 2020, inhibition stabilized networks). The high-active state has as much E-activity as the seizure state, but with inhibition still mostly or entirely intact, and so could correspond to a herald spike, which have comparable amplitude to the seizure itself, but have brief duration (de Curtis and Avanzini, 2001). A tetrastable system has both active and high-active states.

These final-arm epileptic systems depend strongly on FoI. In all cases, a final-arm seizure state cannot exist without FoI, and moreover, the number of stable states may change based on the inclusion of FoI, as any high-active and final-arm seizure states will collapse into a single fixed point in the absence of FoI (cf. Fig. 1.4). A bistable FoI-system will still be a bistable system without FoI, but the seizure state will become a high-active state. A tristable FoI-system with a high-active state will

2 "support" is in quotations here because the first arm of the I-nullcline is zero, and hence is not supported in the usual sense.

become a bistable system, seeing its high-active and seizure states collapse into a single high-active state. A tristable FoI-system with an active state without FoI will become a tristable purely-firing-system with both an active and a high-active state. A tetrastable FoI-system will also become a tristable system without FoI, where its high-active and seizure states collapse into a single high-active state. In all cases the rest state remains unchanged with or without FoI.

1.2.5 Focal seizures are rooted in failure of inhibition

The qualitative difference in rescuability between first-arm and final-arm seizure states makes a strong argument that intractable focal epilepsies may be rooted in FoI. While high-SI states congruent with the high-E, low-I electrophysiology do exist in first-arm epileptic systems even in the absence of FoI, systems in these seemingly-epileptic states can be returned to rest by straightforward external drive to the inhibitory population (say by a homeostatic mechanism). The same is not true of final-arm seizure states.

Not all perturbations are created biophysically equal. We assume our model encompasses the local network, with the standard assumption that non-local connections are exclusively excitatory[3]—therefore, only excitatory perturbations are biophysically reasonable [4] Without FoI, excitatory perturbations corresponds to perturbations up and to the right, and this works perfectly to rescue either the high-active or the first-arm seizure state: excite the inhibitory population (upward

3 Experiments suggest up to 0.5% of long-range projections are inhibitory (Tamamaki and Tomioka, 2010)

4 A caveat: other perturbations are allowed if and only if we assume there was already a non-zero external input. In that case, we can effect a negative perturbation by decreasing that external input. However, we are attempting to model the seizure core; if it is only seizing due to an external stimulus, and stops when that external stimulus stops, then this it is not self-sustaining, and so is not the core of interest. This does not exclude the possibility of other, more distributed seizures, merely sets them aside for the purposes of this paper.

perturbation), and the increased I will cause local inhibition leading to E decreasing (up-and-to-the-left result), and so the seizure is rescued. However, *with* FoI, we cannot always perturb upward. Wherever inhibition is failing, we can only perturb rightward and *downward*, because exciting the inhibitory population will cause inhibitory activity to *decrease*. In particular, the final-arm seizure state has undergone a complete failure of inhibition, so I cannot be perturbed. Thus, *the final-arm seizure state cannot be rescued.*[5]

1.2.6 Tristable systems model functional, yet epileptic, cortex

Models often omit that epileptic cortex should probably function when not seizing. They may incidentally account for sub-population-scale computation by virtue of having a non-seizing stable state, but population-scale computations are known (Harris and Ermentrout, 2018; Muller et al, 2016; Rapela, 2018; Sato et al, 2012), and these computations are thus far unaccounted for in seizure modeling. To account for these, we are interested in tristable (or tetrastable) models, so that state change can serve a function other than seizure. In exploring the parameter space of our FoI model, we find many tristable systems with rest, active, and seizure states, as well as tetrastable systems with a high-active state.

These systems can model the joint effect of epilepsy treatment on both the seizure state and the active state. Consider Fig. 1.4, which depicts an epileptic system before and after a treatment by a hypothetical anti-epileptic drug (AED). The hypothetical AED acts to reduce the efficacy with which the excitatory population drives the inhibitory population, a counterintuitive treatment that only makes sense in light of FoI (Liou et al, 2020). The primary effect of this treatment is to shift the

5 Another interesting caveat occurs in the ultra-high excitation case when excitatory cells might fail, thus "rescuing" the seizure via complete failure of all neural activity.

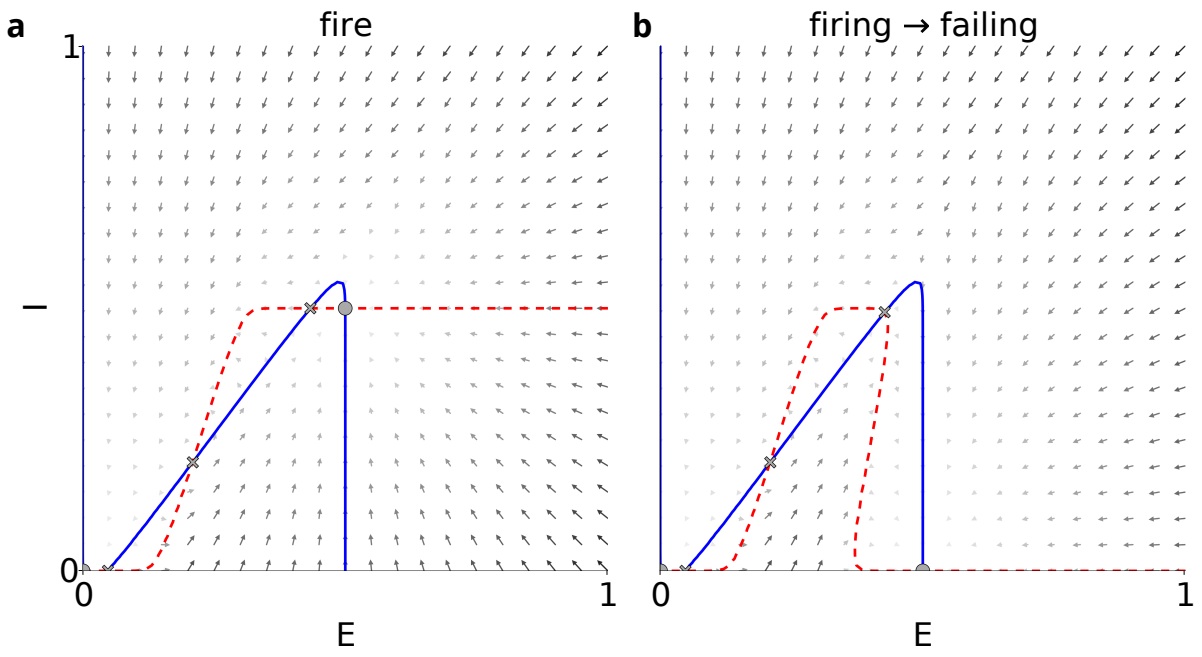


Figure 1.3: **FoI adds seizure state to otherwise identical models.** These example phase spaces characterize models differing only in the inhibitory population's nonlinearity, either purely firing, *left*, or firing and then failing, *right*. While the standard WCM model has a highly excited stable state (*left panel, right-most circle*), only with the addition of FoI do we find a stable state with high excitation and low inhibition. Moreover, in the standard WCM, due to the monotonicity of the inhibitory nullcline (*red dashes*), no high-E/low-I stable state can coexist with any stable state with even moderate inhibitory activity (such as exists in an inhibition stabilized network in Sanzeni et al, 2020). Excitatory nullcline, *solid blue*, inhibitory nullcline, *dashed red*. Stable fixed points, *circles*; unstable fixed points, *crosses*. Both panels: $A_{EE} = 17$, $A_{EI} = 9$, $A_{IE} = 19$, $A_{II} = 4$, $\theta_{E\text{fire}} = 1.5$, $\theta_{I\text{fire}} = 4$. Right panel: $\theta_{I\text{fail}} = 8$

inhibitory nullcline rightward (the nullcline's shape also changes slightly). This increases the seizure threshold substantially, but also increases the average amplitude of the limit cycle (as well as changing the frequency slightly). Additionally, perhaps the seizure state's being closer to the separatrix would make it easier to rescue, by whatever means the brain may use. In concrete effect, though, the system is substantially less susceptible to seizure due to perturbations from rest, slightly less susceptible to seizures due to perturbations from its active state, its active state operates at a higher activity level and reduced frequency, and the system may be

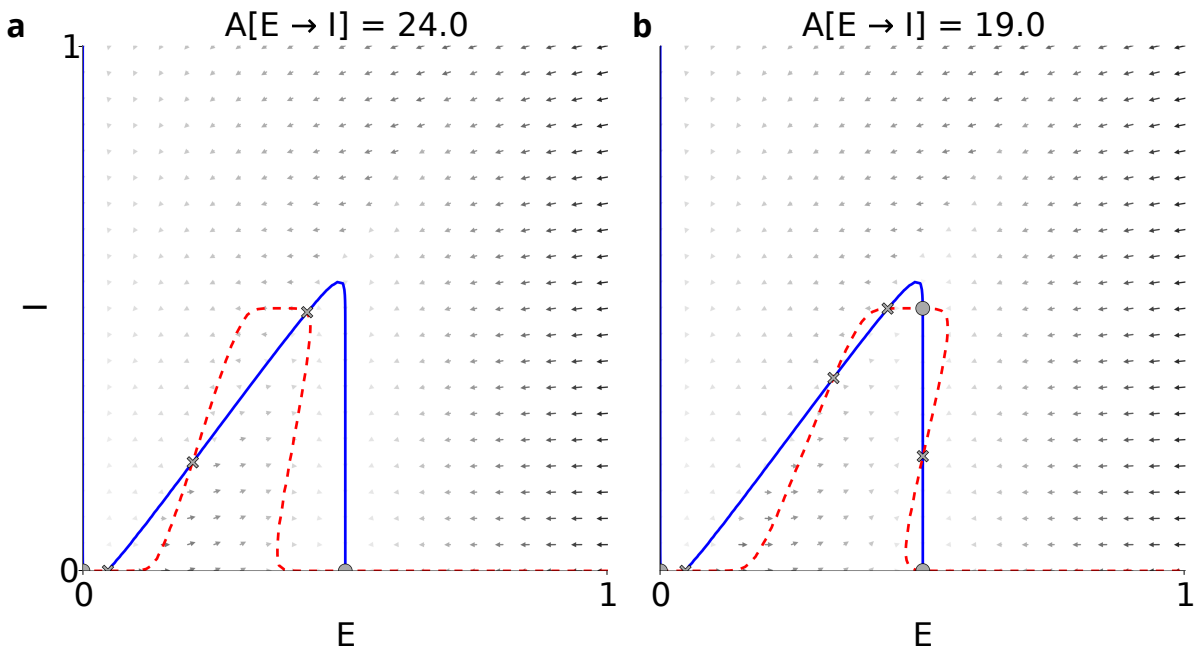


Figure 1.4: **Excitation of inhibitory cells (A_{IE}) nonlinearly modulates seizure susceptibility and rescuability.** By decreasing A_{IE} from the left to right panel, we remake the seizure profile of the model. The models share three fixed-points of interest: rest (low-E/low-I), active (moderate-E/moderate-I), and seizure (high-E/zero-I), and the FoI model has additionally a stable high-E/high-I fixed-point (one might consider this high-E state a candidate for herald spikes, cf. de Curtis and Avanzini, 2001). The higher A_{IE} seizure state is further from the physiological active regime, meaning rescuing the seizure will require more external activity modulation. Excitatory nullcline, *solid blue*, inhibitory nullcline, *dashed red*. Fixed points indicated stable, *circles*, or unstable, *crosses*. Both panels: $A_{EE} = 19$, $A_{EI} = 13$, $A_{II} = 6, \theta_{E,\text{fire}} = 1.5, \theta_{I,\text{fire}} = 4, \theta_{I,\text{fail}} = 8$.

disposed to inter-ictal spikes.

How could we then improve on this treatment? By decreasing A_{IE} a bit further, we could eliminate the seizure state entirely, but we also risk eliminating the active state, and thus compromising cortical function. We would rather either a) move the right-hand slope of the I-nullcline further right, independent of the left-hand slope, or b) move the final arm of the E-nullcline further right, independently of the middle arm. Both of these changes are possible. The right-hand slope of the I-nullcline depends on the failure threshold of FoI, so any AED that could increase

this independently of the firing threshold would theoretically treat epilepsy with far less effect on cortical function. The E-nullcline final arm can be targeted through the excitability of the E population, though not truly independently of the middle arm.

1.3 Methods

1.3.1 Deriving populations from single cells

Like the preceding works (Meijer et al, 2015; Kim and Nykamp, 2017), our model will be a modified WCM. In keeping with WCM (Wilson and Cowan, 1972), our derivation begins with a single neuron modelled as a binary switch, but with a modification for the inhibitory neurons.

Single cell model

Wilson and Cowan (1972) used a binary switch as their single cell model. To incorporate failure of inhibition, we similarly use a binary-state neuron, but with two switches: the original onset (firing), and a new offset (failing). Since we are only concerned with failure of inhibition here, we keep the original single-switch for the excitatory model neuron, so that the excitatory population dynamics are precisely as in the WCM. What follows is our derivation for the inhibitory population only.

Wilson and Cowan let D_{fire} be the distribution of firing thresholds. To obtain a sigmoid, they assumed D_{fire} is unimodal. Analogously, we let D_{fail} be the distribution of thresholds at which firing fails, also unimodal, with the additional biophysical fire-then-fail assumption that for a given neuron, i, $\theta_{\text{fire}}^{(i)} < \theta_{\text{fail}}^{(i)}$. If a cell would

fail before firing, then we contend more has gone wrong than we are interested in modeling, and likely at the cellular scale. The first row of figure 1.5 illustrates this construction.

Population averaging

Next, we average over these distributions, but first take a step back to understand the meaning of their average. In their original paper, Wilson and Cowan wrote: $S(x) = \int_0^x D_{\text{fire}}(\theta)\, d\theta$. This finds the proportion of neurons firing when the population recieves input x by finding the proportion of neurons with threshold less than x. In our case, then, we need to find the proportion of neurons with $\theta_{\text{fire}} < x < \theta_{\text{fail}}$. To start, we can count the number of neurons with $\theta_{\text{fire}}^{(i)} < x$, most simply $\sum_{i=1}^{N_{\text{neuron}}} \mathbf{1}[\theta_{\text{fire}}^{(i)} < x]/N_{\text{neuron}}$, where $\mathbf{1}$ is the indicator function. However, rather than counting by neuron, we can count by threshold, namely $\int_0^x D_{\text{fire}}(\theta)\, d\theta$, which we know has sigmoid shape. Similarly the number of neurons with $\theta_{\text{fail}} < x$ is $\int_0^x D_{\text{fail}}(\theta)\, d\theta$. Our biophysical "fire-then-fail" assumption that $\forall i : \theta_{\text{fire}}^{(i)} < \theta_{\text{fail}}^{(i)}$, gives us that the neurons in the second quantity were included in the first quantity, so we can simply subtract them. Thus

$$\begin{aligned} S(x) &= \int_0^x D_{\text{fire}}(\theta)\, d\theta - \int_0^x D_{\text{fail}}(\theta)\, d\theta \\ &= S_{\text{on}} - S_{\text{off}}\,. \end{aligned} \tag{1.3}$$

Again given our fire-then-fail assumption $\forall i : \theta_{\text{fire}}^{(i)} < \theta_{\text{fail}}^{(i)}$, we know S is a function taking values between 0 and 1 that starts asymptotically at zero, before approaching some maximum as a sigmoid-like function, and then returning to zero asymptotically. Note that in the case that the distributions do not overlap (i.e. $D_{\text{fire}} D_{\text{fail}} \approx 0$ everywhere) the maximum the function approaches is 1. For simplicity, we assume

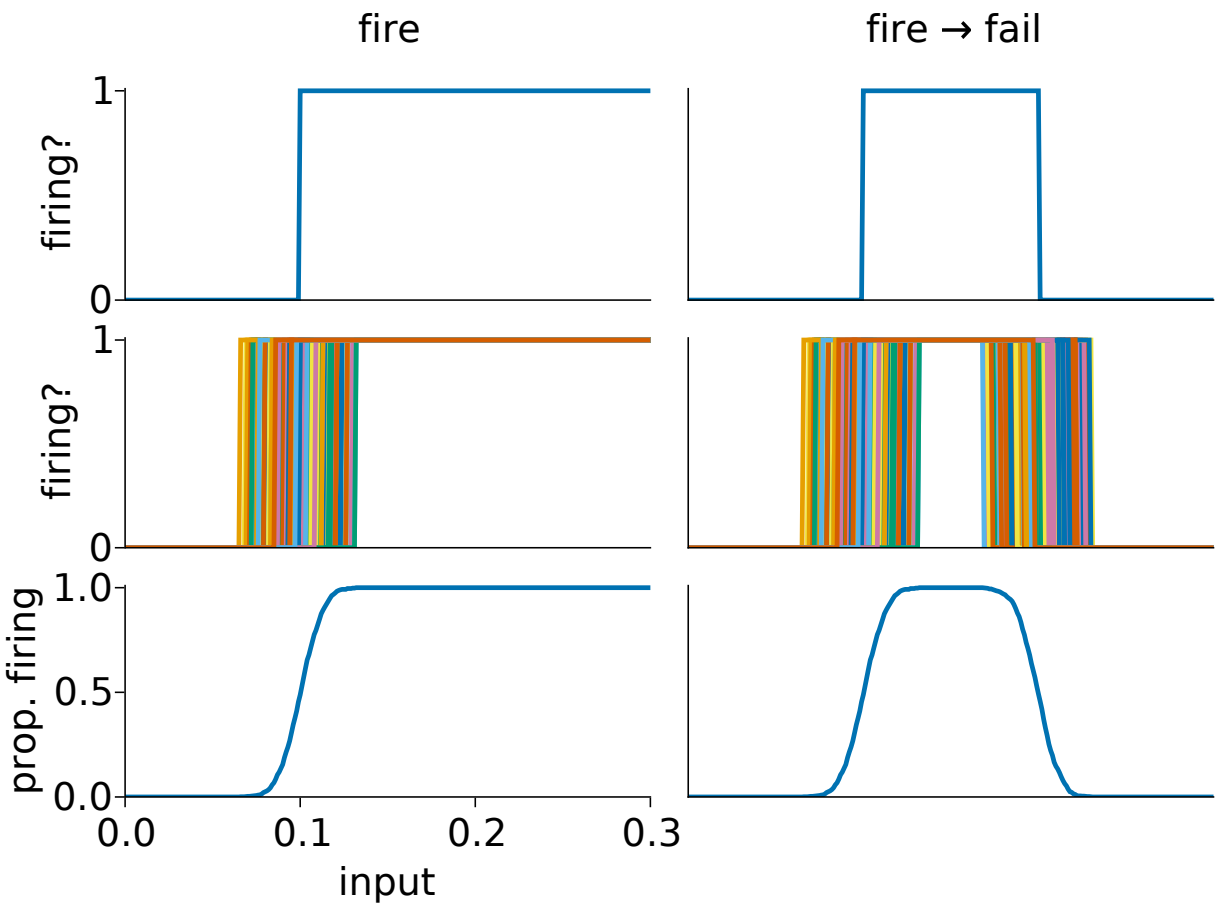


Figure 1.5: **Sketch of nonlinearity derivation for a purely firing population, *left*, and for a population that also fails, *right*.** *Top*, example binary switch neuron, either only firing as in the WCM *left* or with failure at high input *right*. *Middle*, superimposition of many neurons with their thresholds drawn from Gaussian distributions. *Bottom*, firing response curves curves. The proportion of neurons firing as function of input current, i.e. the average of the middle row. When the neurons only fire, the firing response is a sigmoid as in WCM (*left*). With failure, the firing response is the difference between two sigmoids (*right*). This is as seen in the second figure of Meijer et al (2015), differing in that we take the average and then approximate the sigmoidal sides by logistic functions, rather than approximating the whole shape by a Gaussian.

this to be the case, which equates to assuming that few cells have a firing threshold greater than the least failure threshold.

To follow the original choice by Wilson and Cowan (1972), and for computational efficiency, we approximate S_{on} and S_{off} as logistic functions, so that S is a difference-of-sigmoids (DoS) with four parameters. Throughout, we will use the logistic function as our sigmoid, but as in Wilson and Cowan (1972), the results will hold similarly for any choice of sigmoid. We show in Fig. 1.2 that this choice of sigmoid fits theoretical firing responses well, when the firing and failing thresholds are considered to be sampled from two normal distributions constrained by the fire-then-fail assumption.

Thus we define the input-response functions of the excitatory and inhibitory populations as

$$\begin{aligned} S_E(u) &= (1 + \exp(-a_E(u - \theta_E)))^{-1} \\ S_I(u) &= (1 + \exp(-a_I(u - \theta_I)))^{-1} \\ &\quad - (1 + \exp(-a_{fail}(u - \theta_{fail})))^{-1} \end{aligned}$$

Our definition of S_I constitutes the only difference between our model and the original WCM. S_E and the rest of the model are as in the original WCM (see Wilson and Cowan, 1972, for derivation details).

$$\begin{aligned} \tau_E \frac{dE}{dt} &= -E + (1 - E)\, S_E(A_{EE}E - A_{EI}I + P_E) \\ \tau_I \frac{dI}{dt} &= -I + (1 - I)\, S_I(A_{IE}E - A_{II}I + P_I) \end{aligned} \tag{1.4}$$

As in WCM, the dependent variables E and I describe the proportion of neurons firing in the excitatory and inhibitory populations respectively, and the dependent variable t is time. The parameters and their values are described in Table 1.1.

symbol	description	value (E, I)
τ	time constant	$(7.8, 34.32)$
$A_{i,j}$	connectivity strength, $j \to i$	varied
P	stimulus	0
a	firing onset sigmoid slope	5
θ	firing onset mean threshold	$(1.5, 4)$
a_{fail}	failure sigmoid slope	5
θ_{fail}	failure threshold	8

Table 1.1: **Description of model parameters and example values.** The time constants were taken from (Sanzeni et al, 2020). The sigmoid slopes were chosen both to have high gain (Harris and Ermentrout, 2018) within the range of connectivity strengths tested, and to avoid potential problems of interpretation due to overlapping distributions of on and off thresholds. The onset thresholds were chosen in line with (Harris and Ermentrout, 2018; Cowan et al, 2016), though in contrast with (Meijer et al, 2015). The offset threshold was chosen to be sufficiently larger than the onset thresholds such that the underlying threshold distributions would not overlap.

1.3.2 Synthetic data and model fitting

All simulations and analyses were carried out in Julia v1.6.2 (Bezanson et al, 2012). Our code can be found at the linked Github repository and its associated registry (github.com/grahamas/FailureOfInhibition2022).

We created $N_{\text{neuron}} = 1000$ firing-and-failing neurons by sampling two thresholds, one firing, $\theta^{(i)}_{\text{fire}}$, and one failing, $\theta^{(i)}_{\text{fail}}$, for each neuron i. We drew each threshold from a normal distribution $\theta^{(i)}_{\text{fire}} \sim \mathcal{N}(\mu_{\text{fire}}, \sigma_{\text{fire}})$ and $\theta^{(i)}_{\text{fail}} \sim \mathcal{N}(\mu_{\text{fail}}, \sigma_{\text{fail}})$, under the biophysical fire-then-fail constraint that for a given neuron i we must have $\theta^{(i)}_{\text{fire}} < \theta^{(i)}_{\text{fail}}$. For each of these neurons, call its firing function $f^{(i)} : \mathbb{R} \to \{0, 1\}$, so $f^{(i)}(x) = 1$ when $\theta^{(i)}_{\text{fire}} \leqslant x < \theta^{(i)}_{\text{fail}}$, and 0 otherwise. Then we can define the input-response function of a population of such neurons as $R : \mathbb{R} \to [0, 1]$ as $R(x) = \sum_{i=1}^{N_{\text{neuron}}} f^{(i)}(x)/N_{\text{neuron}}$.

For model fitting, we define three models: Gauss, the Gaussian model used in Meijer et al (2015), Kim, the product-of-sigmoids model used in Kim and Nykamp

(2017), and DoS, our difference-of-sigmoids model.

$$\begin{aligned}
\text{Gauss}(x; \mu, \sigma, A) &= Ae^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \\
S(x; \theta, a) &= \frac{1}{1 + e^{-a(x-\theta)}} \\
\text{Kim}(x; \theta_{\text{fire}}, a_{\text{fire}}, \theta_{\text{fail}}, a_{\text{fail}}) &= S(x; \theta_{\text{fire}}, a_{\text{fire}})(1 - S(x; \theta_{\text{fail}}, a_{\text{fail}})) \\
\text{DoS}(x; \theta_{\text{fire}}, a_{\text{fire}}, \theta_{\text{fail}}, a_{\text{fail}}) &= S(x; \theta_{\text{fire}}, a_{\text{fire}}) - S(x; \theta_{\text{fail}}, a_{\text{fail}})
\end{aligned}$$

Using a root-mean-squared error (RMSE), these functions were fit to the population input-response function $R(x)$ with discrete values x taken from an evenly spaced range of 2000 numbers between -2 and 2 (inclusive).

For both DoS and Kim, the initial guesses for θ_{fire} and θ_{fail} were the first and last values above half-maximum, respectively. For a_{fire} and a_{fail}, the initial values were the reciprocals of the distances between the θ initial conditions and the first (a_{fire}) and last (a_{fail}) values close to the maximum (greater than 0.99max).

For the Gaussian, the initial A was simply the maximum, the initial μ was halfway between the initial θ_{fire} and θ_{fail}, and the inital σ was the difference between the initial guesses for θ_{fire} and θ_{fail} divided by $2\sqrt{2\log(2)}$.

The synthetic data were fitted by each of these three models using the L-BFGS algorithm with forward autodifferentiation implemented in the `Optim.jl` package (Mogensen and Riseth, 2018).

1.3.3 Phase space analysis

Plotted nullclines were calculated as zero-level contours of the derivative fields using the marching squares algorithm implemented in the `JuliaGeometry/Contour` package, with the $E \times I = [0, 1] \times [0, 1]$ space sampled at a resolution of 0.01.

Fixed points were identified as intersections of said nullclines, calculated pre-

cisely through iterative application of the above nullcline calculation on progressively smaller windows around putative intersections. The nullclines intersections were found as zeros of the Manhattan norm of the gradient, with error tolerance `sqrt(eps()) = 1.5e-8`. Putative intersections were assumed to be separated by a Euclidean distance of at least `1e-4` in $\mathrm{E} \times \mathrm{I}$ space.

1.4 Discussion

Our model can make nuanced predictions about the relative effects of putative epilepsy treatments on both seizures and normal function. We accomplished this first by considering seizure susceptibility and rescuability, rather than the mere existence of a seizure state, since a system may have the potential to seize without actually seizing. Such a system would be tristable given physiological states that are visible at the population level (e.g. rest and active states as in inhibition stabilized networks). Then, analogous to the susceptibility and rescuability of the seizure state, these physiological states can be characterized by their excitability and stability, i.e how easily the system can transition between rest and active states. Any epilepsy treatment must improve either seizure susceptibility or rescuability (preferably both), while also keeping the system's excitability and stability within functional limits. Our model can predict these effects.

We also argue that FoI is key to focal seizures, to the point that it may be a key feature of EZs: the fact that a small region's runaway activity unilaterally drives the seizure necessitates that inhibition is failing in some sense. Mathematically, this follows from accepting the possibility of a functional system that has the potential to seize *focally*. If we discard the possibility of a seizure state sustained by external inputs (i.e. not focal, not epileptogenic), then the seizure must correspond to a fixed

point in the system's phase space, one which certainly has high levels of excitatory activity, and which experiments indicate has suppressed levels of inhibitory activity (Tryba et al, 2019). Given that the system is functional, we assume the existence of an active state, which has both less excitatory activity and more inhibitory activity than a putative seizure state. Under these assumptions, the seizure state can only exist if the inhibitory nullcline decreases after the active state (which must be on or after the increasing arm of the I-nullcline). Such a decrease is equivalent to FoI. Thus, given that focal seizures are experimentally characterized by hyperexcitation and depressed inhibition, focal seizures in otherwise functional cortex must involve FoI.

The presence of depressed inhibitory activity in focal seizures has been the subject of some debate. However, evidence suggests, and we were unable to find evidence to the contrary, that inhibition fails specifically in the focal core. Of those papers we found that seemed to contradict failure of inhibition, all were of recordings potentially within the penumbra (e.g. Truccolo et al, 2011). We emphasize that our model only applies to the focal core. An important future direction for this research is in the penumbral population activity: how do penumbral populations behave when driven by an ictogenic core whose inhibition is failing? Does that activity correspond to experimental recordings? How is the functional excitability of penumbral populations affected by AEDs?

Our results are agnostic to the underlying cause of FoI. Our derivation assumes a cause analogous to depolarization block (DB), that FoI occurs because inhibitory cells fail to fire above a certain input threshold, but this assumption is without reference to the underlying mechanics of FoI. This is reflected in the simplicity of our first principles: binary switch neurons with only the most minimal fire-then-fail assumption on the relationship between firing and failing thresholds. A potentially

important future direction will be to substitute empirically-derived relationships between these thresholds, which must surely exist and must reflect the exact mechanism underlying FoI.

Chapter 2

Third-order motifs are sufficient to fully and uniquely characterize spatiotemporal neural network activity

2.1 Introduction

Meaningfully capturing the complexity of neural networks is a daunting task: a range of different approaches are currently in use to quantify network behavior (e.g. Abeles and Gerstein, 1988; Dechery and MacLean, 2018; Schneidman et al, 2006; Jiang et al, 2017; Sporns and Kötter, 2004; Jovanović and Rotter, 2016; Milo et al, 2002; Buzsáki and Draguhn, 2004). Intuitively, we assume that completely characterizing this complexity should require correlations of such high-order as to be incomprehensible, perhaps even unenumerable (Yu et al, 2011). However here we show that *third-order* correlations fully characterize even the most complex neu-

This chapter is under consideration as a manuscript of the same name, co-authored with Sarita S Deshpande and Wim van Drongelen. It was also published as a pre-print on *bioRχiv* under the title "Unique Characterization of Spatiotemporal Neural Network Activity" (2021), before being updated to the current version in 2022. Sarita and I contributed equally in research and authorship.

ral recordings.

This full characterization arises from the unique correspondence between a dataset and its triple correlation, per the triple correlation uniqueness (TCU) theorem. Introduced decades ago in optical sciences (Bartelt et al, 1984), this theorem states that any finite image has unique triple correlation. The TCU theorem has since languished[1], but we hope to bring triple correlation to the forefront of data analysis with one simple observation: any finite dataset can be interpreted as an "image." In neuroscience, this encompasses any completed recording of neural activity: functional magnetic resonance imaging, local field potential, single-unit electrophysiology, multi-electrode array electrophysiology, voltage-sensitive dye imaging, etc. Beyond these, the TCU theorem applies to any finite dataset in any field.

Here, we bring the TCU theorem into the heart of neuroscience with a focus on analysis of spike rasters as the representation of neuronal population activity. We propose that the activity patterns of a neural network can be ideally characterized by its triple correlation, c_3. We begin by defining c_3, which comprises triplet motifs, the relationships among three events. We briefly summarize the proof of the uniqueness of c_3 and explore the implications of this uniqueness for neural data. We then classify the triplet motifs by event sequencing and neuron recruitment to derive a natural summary of c_3 in fourteen motif classes. This summary falls out naturally and maps onto network properties of both computational and biological interest, including synchrony, feedback, feedforward, convergence, and divergence. We illustrate the utility of that simple summary with some straightforward simulations and apply our analysis to experimental recordings. In sum, we present

1 The theory was invoked in arguments about the Julesz conjecture Julesz (1962); Julesz et al (1973); Yellott and Iverson (1992); Yellott (1993); Victor (1994), which stated that infinite textures with identical third-order correlations are indistinguishable to a human observer. As noted by Victor (1994), these arguments were made somewhat convoluted by the fact that the Julesz conjecture concerns textures rather than images. A mathematical texture is a statistical ensemble of images (i.e. an infinite collection of images), whereas an image is finite and bounded.

a novel, theory-based approach to quantifying network activity.

2.2 Methods

We worked with a multi-neuron raster of spikes (Fig. 2.1A) to outline our approach, though the analysis works the same for any finite dataset (see section 2.5).

2.2.1 Defining triple correlation

Unlike pairwise correlation (a function of one lag between two spikes), triple correlation characterizes three-way interactions as a function of **two** lags among three spikes (Fig. 2.1B). We will call $\mathcal{C}_3$ the triple correlation transform, which transforms a spatiotemporal raster into its triple correlation. So for a spatiotemporal raster, $r(n,t)$ (n, space; t, time), the triple correlation c_3 is a function of four variables: $\mathcal{C}_3[r] = c_3(n_1,t_1,n_2,t_2) = \langle r(n,t)r(n+n_1,t+t_1)r(n+n_2,t+t_2)\rangle_{n,t}$. The operation $\langle ...\rangle_{n,t}$ computes the average over all bins (n,t) in the raster. The argument ... is 1 when (n,t) is part of a spiking triplet with lags (n_1,t_1,n_2,t_2), and 0 otherwise. In practice, we only calculate the triple correlation up to a certain maximum spatiotemporal lag which we determine based on experimental and computational considerations on a per-experiment basis (noted in each figure). In most cases calculating the full c_3 would be a needless computational expense as usually the triple correlations with longer lags become more likely to be dominated by chance. See subsection 2.5.4 “Computing the triple correlation” for implementation details.

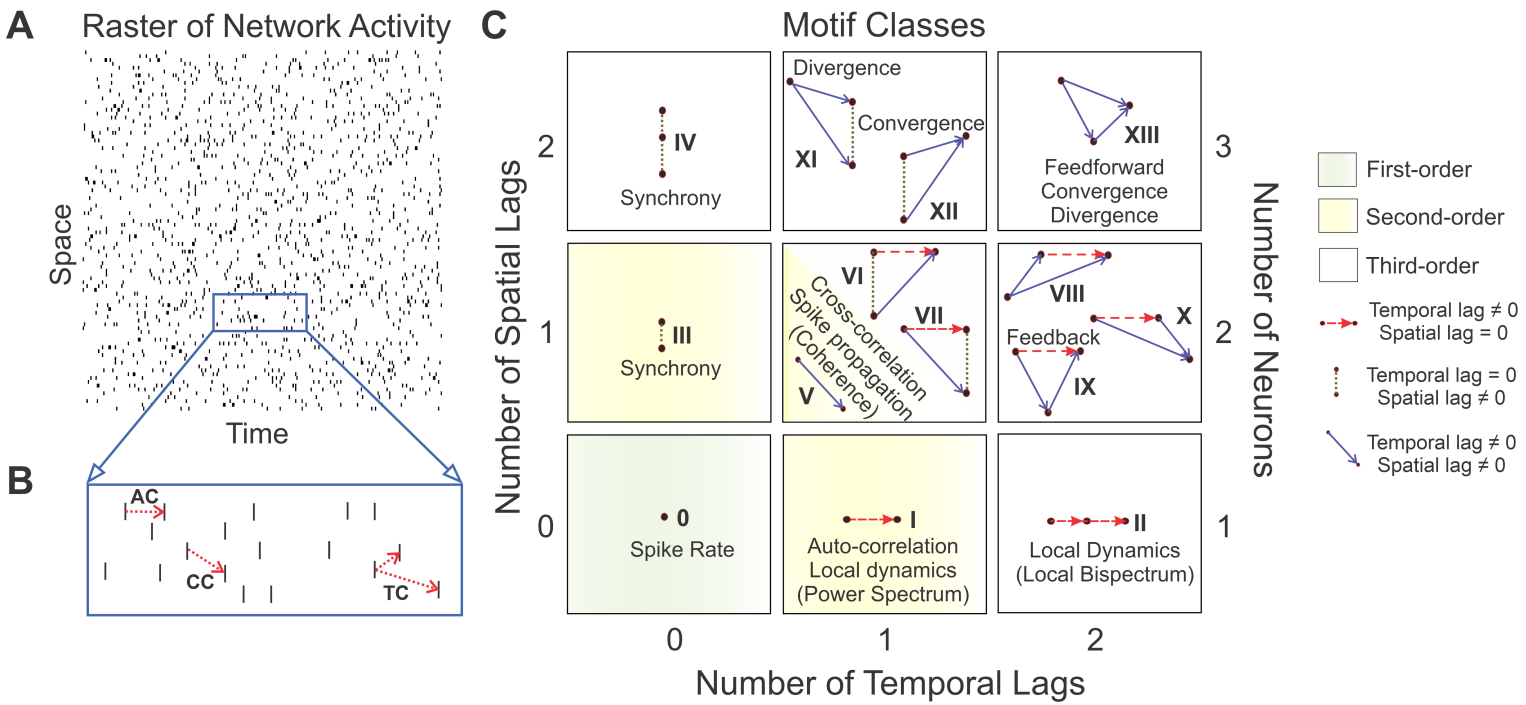


Figure 2.1: **Application of the fourteen motif classes to spike rasters.** For a given raster (Panel A), second-order correlations can relate activity within a neuron (auto-correlation, AC; Panel B) or between neurons (cross-correlation, CC; Panel B). Triple correlation (TC; Panel B) relates three bins, separated by up to two temporal and two spatial lags. Three particular spike bins constitute a triplet motif (e.g. as shown in Panel B). We classify these motifs into fourteen motif classes by the motif's spike sequence (Panel C; see section 2.5 for a complete derivation). Dot: a single spike bin. Horizontal red dashed arrow: intraneuronal spike bins, i.e. space lag $= 0$, e.g. I and II. Vertical stippled line: synchronous spike bins, i.e. time lag $= 0$, e.g. III and IV. Solid blue arrow: interneuronal spike bins (e.g. V). These 14 motif classes can also embody well-known neuronal processing properties (such as synchrony, feedback, convergence, divergence, and feedforward) in both the time (listed adjacent to the motif class) and frequency domains (listed within parentheses). First- (0) and second-order (I, III, and V) motif classes are highlighted in green and yellow, respectively. The second-order motif classes I, III, and V are constituent motif classes that comprise the third-order motifs. The remaining ones are third-order motif classes.

2.2.2 Overview of uniqueness proof

To prove uniqueness, we apply the Triple Correlation Uniqueness (TCU) theorem, originally applied in optical sciences (Bartelt et al, 1984; Yellott and Iverson, 1992; Yellott, 1993). It states that if we have two images x and y with equal triple correlations ($\mathcal{C}_3[x] = \mathcal{C}_3[y]$), then the images themselves must be equal ($x = y$). The TCU theorem applies to any finite bounded dataset, including continuous signals and multi-dimensional data—a fact that can be understood by interpreting any dataset as an image—so we apply it to a spike raster. We begin with two rasters x and y whose triple correlations are equal. This implies that the triple correlations' Fourier transforms (bispectra) are equal, which we can write as equality between the product of three characteristic functions. Characteristic functions are the Fourier transform of probability distributions and are well-known in statistics to have certain nice properties (Feller, 1950). Using those properties, we perform simple algebraic manipulations of that equality to derive that the Fourier transform of x equals the Fourier transform of y times an exponential function, $\mathcal{F}[x](\sigma) = \mathcal{F}[y](\sigma)e^{j a \sigma}$, where σ is frequency, j is the imaginary number, and a is some constant. This means that raster x equals raster y translated by some constant a. Thus, rasters with equal triple correlations are themselves equal up to translation. See section 2.5 for the full proof, both in the two dimensions of a typical raster and in the fully general N dimensions. As a consequence of this result, spiking activity can in principle be recovered in its entirety from triple correlation (see "Reconstruction Algorithms for Finite Images" in Yellott and Iverson, 1992).

2.2.3 Summarizing triple correlation with motif classes

In order to better interpret triple correlation, which is inherently high-dimensional, we summarized it along lines meaningful to the underlying neural dynamics, combining the motifs into qualitatively distinct motif classes, M_i. To do this, we asked what differences between three-spike motifs constitute fundamental differences. First, we grouped motifs according to whether their lags are zero or non-zero, producing $2^4 = 16$ groups for the four spatio-temporal lags . Next we distinguished within these groups according to the signs of the lags, which expanded those 16 groupings into 169 lag-sign motifs (proven in subsection 2.5.6 "The number of lag-sign motifs"; enumerated in Table A1). These 169 lag-sign motifs enumerate all possible shapes of the motifs where the identity of each node matters, and where space is ordered. However, for our purposes many of these shapes are functionally the same. For example, we only distinguished between zero and non-zero spatial lags: zero meaning "intra-neuronal" and non-zero meaning "inter-neuronal." Further, node identity is irrelevant. Using these considerations, we grouped together these 169 lag-sign motifs according to those invariant under reflecting over the horizontal axis and node identity permutation. As a result, we found fourteen motif classes (Fig. 2.1C; proven in subsection 2.5.7 "The number of motif classes") .

2.2.4 Controlling for expected contributions

When using motif classes to quantify network patterns, it is essential to distinguish between the occurrence of these motif classes due to underlying network behavior versus due to chance. Based on simple combinatorics, we expect contributions for each motif class to differ by orders of magnitude. For example, motif classes I and III each have only one varying lag, whereas motif class XIII has four. So, if λ is

the maximum lag between spikes in a motif, the expected contributions of motif classes I and III are $O(\lambda)$, while the expected contribution of motif class XIII is $O(\lambda^4)$. Intuitively, we also expect the motif class contributions to vary depending on the number of spikes required in the motif: while there are more motifs in motif class XIII, the chances of any one of them occurring is lower than motif class I because a motif needs three spikes for motif class XIII but only two for motif class I. So given a fixed spike rate p, we calculated the theoretically expected motif class contributions under the simplifying assumption that every neuron is a Poisson process with rate p per bin: $E[M_i \mid p, \lambda] = NT\#(M_i \mid \lambda)p^{n_i}$ where NT is the size of the raster, $\#(M_i \mid \lambda)$ is the number of motifs given maximum lags λ, and n_i is the number of events in a motif in motif class i. (See subsection 2.5.8 "Expected contributions per motif class" for case-by-case calculation)

We note that the higher-order motifs are composed of lower-order motifs, e.g. looking at motif class X in Fig. 2.1C, it is constructed of two kinds of arrows: those of motif classes I and V. We say that motif classes I and V are constituent motif classes of motif class X (see also Fig. 2.4 for more explicitly deconstructed examples). We want to control also for this dependency. By computing the expected relationship between the contribution of a higher-order motif class and the contributions of its lower-order constituent motif classes, we derived $E_c[M_i]$ which is the expected motif-class contributions given a spike rate and controlling for lower-order constituent motif-class contributions (constituent-controlled expectation). For lower order motifs, this is the same (e.g. $E_c[M_I] = E[M_I]$), but for higher order motifs this reflects the expected contribution in a raster of Poisson processes with the observed contributions of its constituent motifs, i.e. $E_c[M_X] \approx E[M_X \mid M_I, M_V]$. We report $M_i/E_c[M_i] - 1$ so that positive values indicate higher contribution than expected, negative indicate less, and zero indicates contributions in line with those expected

due to noise and the observed lower-order constituent motif-class contributions.

2.3 Results

Our principal goal is to bring triple correlation into general usage, with the triple correlation uniqueness theorem providing theoretical foundation for its use. Since triple correlation is entirely untested in the field, we approached it as we would any new tool: we showed that it works in the simplest possible case, before adding noise and testing it across the whole gamut of possible patterns. After these checks, we applied triple correlation to open-source real-world data.

2.3.1 Application of triple correlation summary

For a simple example to test this approach, we simulated a network spike raster with synchronous, periodic firing at a frequency of $f = 0.08$ arbitrary units (Fig. 2.2A). From this raster, we determined contributions of all motif classes by summing the raster's triple correlation across all motifs in each motif class (M is a vector of M_i for all $i \in 0 : \mathrm{XIII}$, where $M_i = \sum c_3(m)$ for motifs m in motif class i). We also simulated rate-matched pure-noise rasters (Fig. 2.2B) as surrogates and calculated their motif-class contributions ($\hat{\mu}$ estimating $\mu = \{E[M_i]\}_{i\in 0:\mathrm{XIII}}$). These surrogates were simulated Poisson processes with identical firing rates for every neuron. In addition we calculated a theoretical value, the constituent-controlled expectation ($\mu_c = \{E_c[M_i]\}_{i\in 0:\mathrm{XIII}}$; see Methods). We plotted all three quantities for each motif class (Fig. 2.2C). These quantities on their own are difficult to read: any difference between motif classes is completely overshadowed by the inevitable combinatorial differences in the number of triplet motifs per motif class. To account for these expected combinatorics and effectively report the contribution of each motif class,

we calculated the constituent-controlled ratio ($M/\mu_c - 1$; Fig. 2.2D), which highlights the deviation of each motif-class contribution from that expected due to noise and lower-order constituent motif class contributions. To demonstrate that the contributions m_i from surrogate rasters do not differ much from the constituent-controlled expectation $E_c[M_i]$, we plotted the estimated-to-constituent-controlled ratio ($\hat{\mu}/\mu_c - 1$) for all $n = 100$ surrogate rasters (Fig. 2.2E).

We see that our summary of the triple correlation reflects the simple underlying structure: motif-classes III and IV, which correspond purely to neural synchrony, are highest above the expected contributions. Motif-classes VI, VII, XI, and XII, which each includes a synchronous component, are also above chance expectations (Fig. 2.2D). Thus as a single facet of our analysis, we are able to detect not only second-order synchrony (motif class III), which is considered important in neuroscience research (Roscoe et al, 1985; Kreuz et al, 2011, 2013), but also third-order synchrony (motif class IV). We recognize that this is an ideal case, in which the synchronicity is also visibly apparent from the raster itself. We will now proceed to show that this particular detection works in the face of noise, and that all the other facets of our analysis (i.e. the other motif classes) also work.

2.3.2 Noise is no obstacle

We explored the effects of increasing noise on detecting network structure using faster synchronous firing with added uniform noise input ($f = 0.12$ AU, SNR = 0 dB; Fig. 2.3A). Similar to the results of those with slower synchronous firing (Fig. 2.2D), motif classes III, IV, VI, VII, XI, and XII (those with elements of synchrony) are found more often than chance (Fig. 2.3B). We increased noise (SNR = -9dB; Fig. 2.3C), resulting in lower magnitude signals across all motif classes, yet still detecting synchronous signals (Fig. 2.3D). In this case, the underlying network structure is still

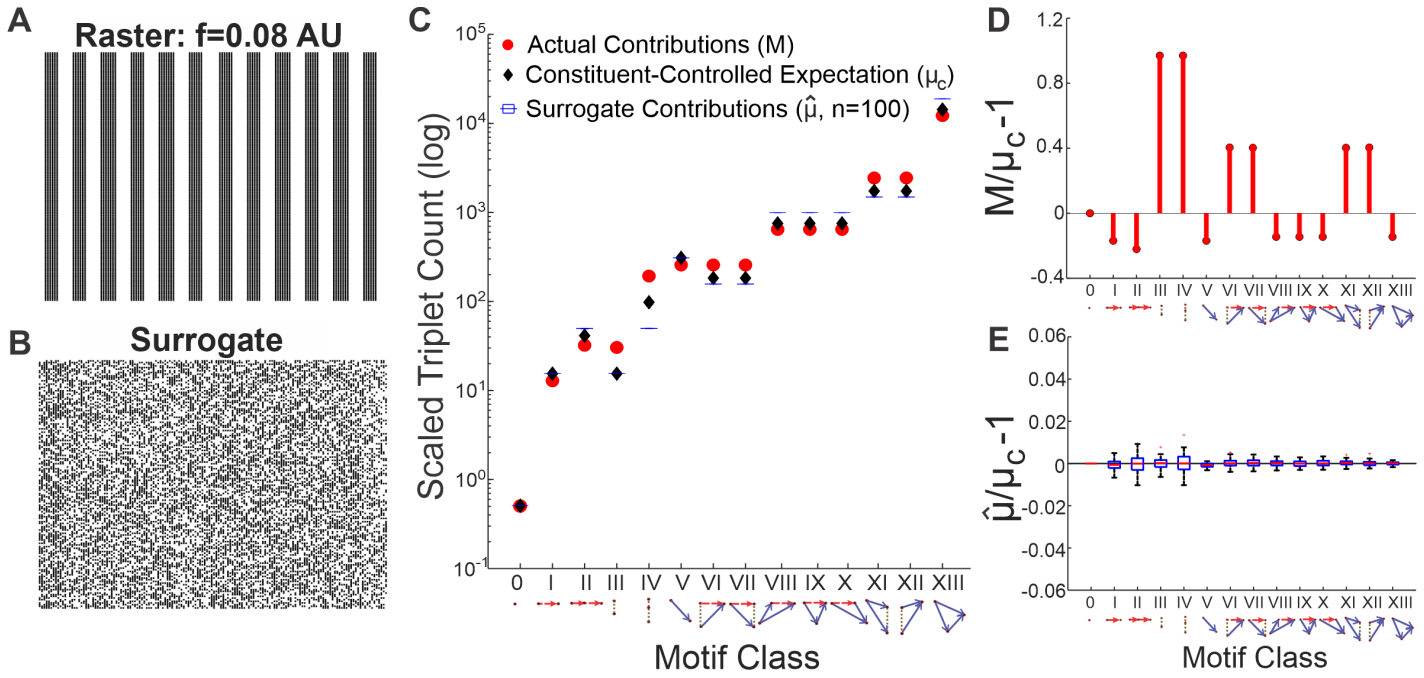


Figure 2.2: **Motif-class contributions to the raster's triple correlation.** A) 150×150 spike raster generated by thresholding an 0.08 AU frequency sine wave. B) Surrogate raster generated by randomly shuffling the periodic raster in panel A. C) Various motif-class summary metrics calculated from a triple correlation using lags up to 20 bins in time and space: the actual contributions per motif class (M, red circles); the constituent-controlled theoretically expected contributions (μ_c, black diamond) conditioning on spike rate and controlling for the observed contributions of lower-order motifs; surrogate contributions ($\hat{\mu_c}$, blue boxplots visible as horizontal lines due to relatively small variance) the average motif-class contributions across $n = 100$ shuffled surrogate rasters. D) The constituent-controlled ratios ($M/\mu_c - 1$) per motif class. $M/\mu_c - 1$ for purely synchronous motif classes III and IV are highest; motif classes VI, VII, XI, and XII (which all consist an element of synchrony) also show positive $M/\mu_c - 1$ values. Note that motif class 0 always has zero signal because motif class 0 is the spike rate, which is controlled for by both N and T. E) The estimated-to-constituent-controlled ratios ($\hat{\mu}/\mu_c - 1$) for 100 noise simulations fluctuate around 0 for all motif classes and are shown as box-and-whisker plots. The centerline is the median, the bottom and top edges of the box are the first and third quartiles, and the whiskers extend to the minimum and maximum values.

overtly present in the raster. We then increased the noise such that the synchronous network structure in the raster is not overtly present (SNR = -17dB;Fig. 2.3E), and found that while all signals move even closer to 0, those with synchrony are still elevated (Fig. 2.3F). Thus, triple correlation reflects underlying structure in the face of added noise. Finally, we show that in an extremely noisy raster (SNR = -40dB; Fig. 2.3G), the synchronous signals are no longer detected as all motif classes now show even lower magnitude signals, approaching 0 (Fig. 2.3H). In this overwhelmingly noisy simulation, the motif-class spectrum is dominated by chance.

2.3.3 Detecting various spiking sequences

Next we tested that our method correctly detects various isolated spike sequences (triplets). Our approach correctly detected third-order motifs by the appropriate motif class. We illustrate six simulations, each including only one repeated triplet across the raster. In addition to motif class 0, all motif classes are composed of one or more constituent motif classes (I, III, and/or V). Of the simulations depicted, the first three show triplets whose motif classes are purely composed of a single constituent motif class (Fig. 2.4A-C), and the other three show motif classes composed of a mix of two different constituent motifs (Fig. 2.4D-F). Fig. 2.4A-C show that the second- and third- order motifs for local dynamics (Fig. 2.4A; I & II), synchrony (Fig. 2.4B; III & IV), and feedforward (Fig. 2.4C; V & XIII) are correctly detected. Note that motif class V (simple second-order cross-correlation, which we describe as “spike propagation”) is in fact a constituent part of every third-order motif other than those in motif classes II and IV , and so was detected at lower levels in every subsequent test. Fig. 2.4D-F shows the other three cases: feedback (IX; composed of I and V), divergence (XI; composed of III and V), and convergence (XII; composed of III and V).

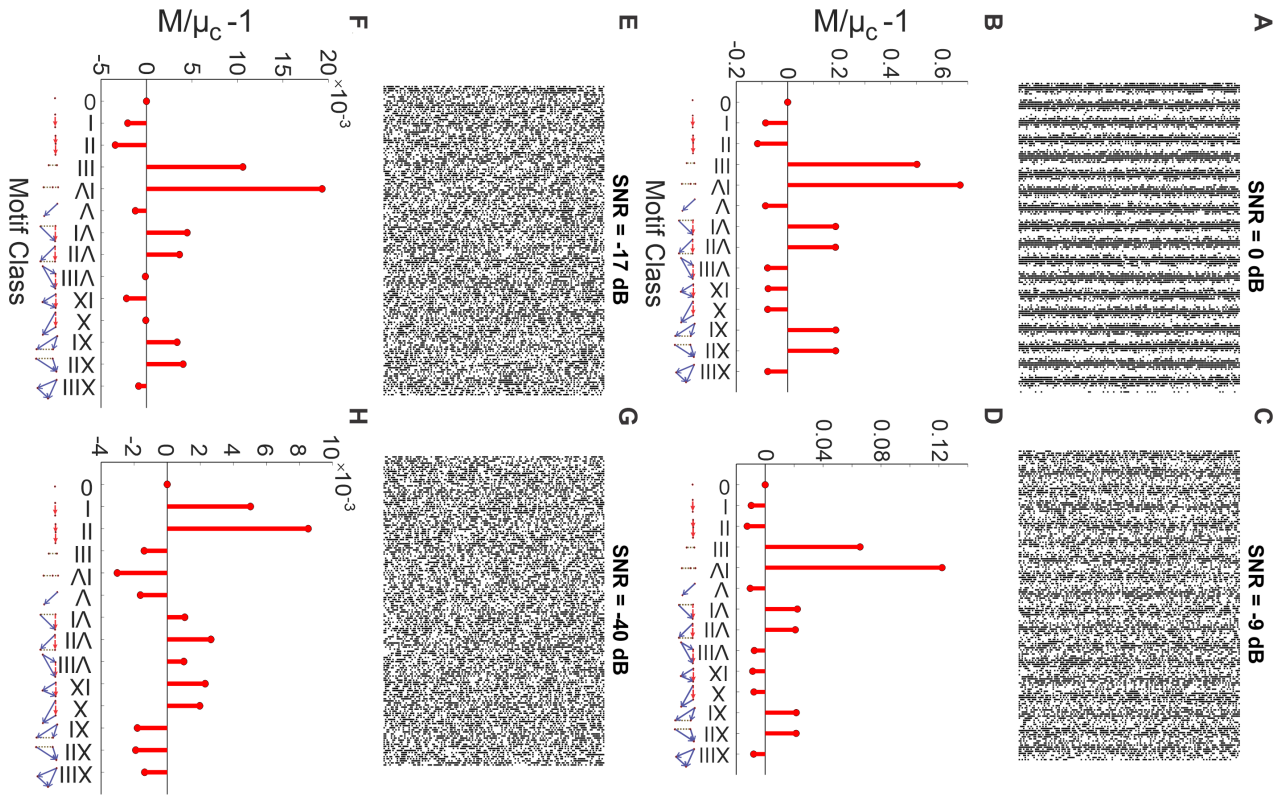


Figure 2.3: **Detecting synchrony amidst increasing noise.** A) 150×150 spike raster plot generated by thresholding a 0.12 AU frequency sine wave with added noise. The noise consists of uniform noise scaled to give the desired signal-to-noise ratio (SNR = 0 dB). B) The motif-class contributions (M) of the above raster relative to chance ($M/\mu_c - 1$). These were calculated from a triple correlation using lags up to 14 bins in time and space. These values show increased contributions of motif classes with synchronous elements, as expected. C) The periodic signal is embedded in more noise (SNR = -9 dB), albeit still visible from the raster. D) Same as panel B with lower magnitude signals. E) The synchronous structure is embedded in more noise (SNR = -17 dB), but now not overtly apparent to the naked eye. F) Same as panels B and D with lower magnitude signals. All motif-class contributions are closer to 0, but motif classes with synchronous elements are still detected. G) The synchronous structure is now embedded in extreme noise (SNR = -40 dB). H) Same as panels B, D, and F, but now other motif-class signals are also detected (and not just those with synchronous elements) due to chance. All motif classes have even lower magnitude signals, approaching pure noise.

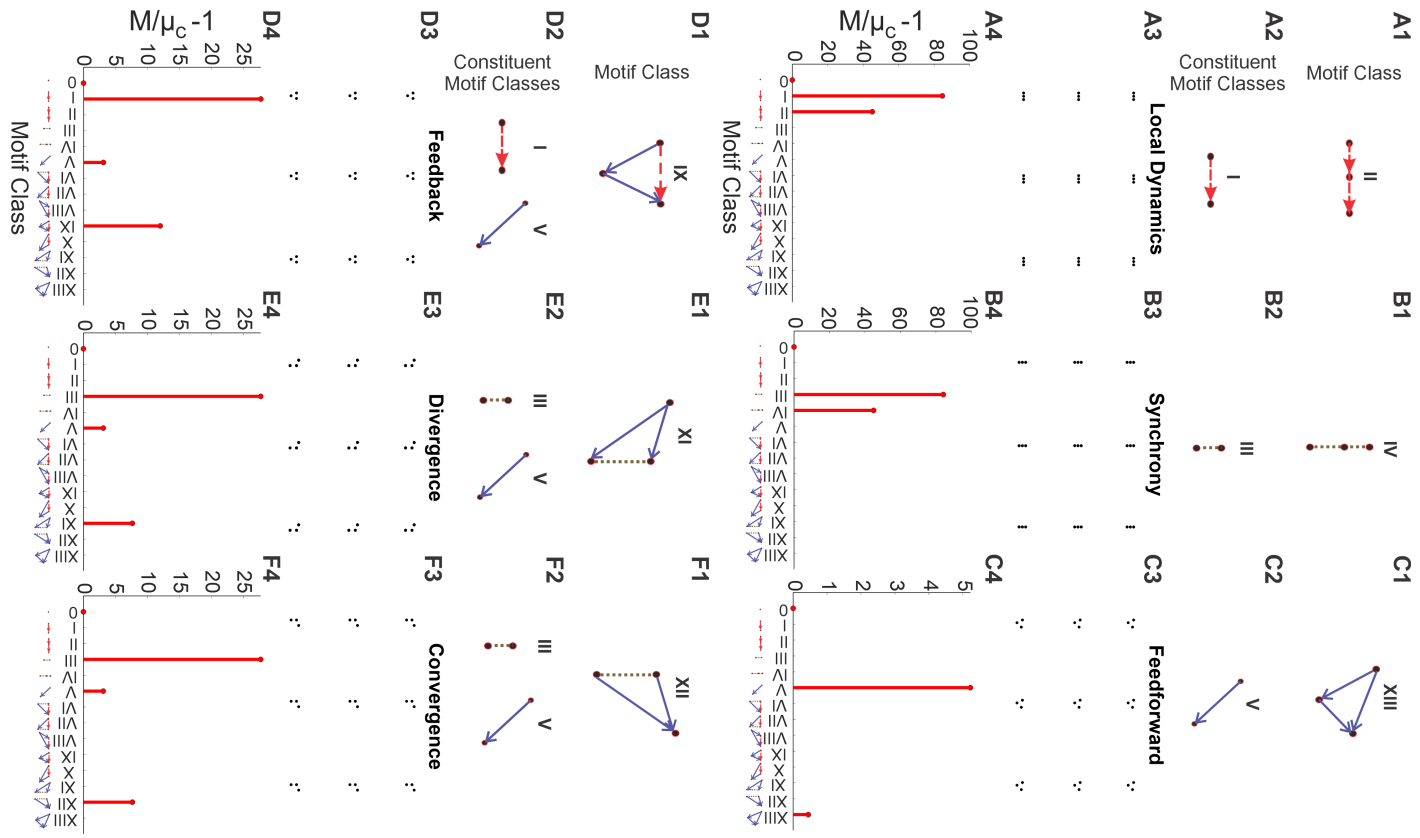


Figure 2.4: **Detecting motif classes in individual patterns.** We simulated 150 × 150 rasters, each consisting of a single repeated triplet. *Row 1*: the motif class of the repeated triplet. *Row 2*: the motif class' constituent motif classes. *Row 3*: the simulated raster. *Row 4*: the motif-class contributions ($M/\mu_c - 1$) calculated from the row 3 raster's triple correlation, using lags of up to 14 bins (which is less than the separation between motifs in the raster). In all cases, the highest-order motif with a non-zero contribution is the motif class, and the remaining non-zero contributions are the constituent motif classes. Note the changing y-axis scale: motif classes I and III are far less likely to occur from chance than motif class V.

In every case, including those not depicted, the tested triplets were reflected by the correct motif-class contribution (along with the triplets' constituent parts). For example, a feedback triplet elevates the contribution from motif-class IX, along with contributions from motif-classes I and V, which together constitute motif-class IX (Fig. 2.4D). Note that the particular summary used in this approach does not depend on quantitative details. Because of this, our detector works purely on 1) the qualitative spike sequencing that defines the motif-class, and 2) the presence of that motif class anywhere within the raster.

2.3.4 Application to experimental data

We further tested our approach using open-source, publicly available network data recorded from rat cortical cultures (details of this dataset are described in Hyvärinen et al (2019) and obtained from Kapucu et al (2022)). Briefly, rat cortical neurons were cultured in microelectrode array (MEA) well plates consisting of 16 (4x4 array) electrodes. Data were collected at 22 days *in vitro*. The datasets consisted of already-detected spike time data from baseline wells ($n = 35$ wells) and from wells treated with the following pharmacological agents ($n = 35$ wells):

1. Control (n=7 wells)
2. γ-aminobutyric acid (GABA, 10 μM, n=7 wells)
3. 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX, 50 μM, n=7 wells), an α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA)/kainate receptor antagonist
4. D-(-)-2-amino-5-phosphonopentanoic acid (D-AP5, 50 μM, n=7 wells), which is an N-methyl-D-aspartate (NMDA) receptor antagonist

5. Gabazine (30 μM, n=7 wells), a $GABA_A$ receptor antagonist

The raw data were sampled at 12,500 samples/s/channel, and the spikes were detected by the researchers as described in Kapucu et al (2022). We downsampled the spike rasters to 500 samples/s/channel. Where multiple spikes were occasionally binned together, we counted them as a single spike. Representative 1-minute snippets of each raster per condition are depicted in Figure 2.6. We computed the triple correlation for each well of the MEA plate with temporal lags of -50ms:50ms and spatial lags of -1:2 electrodes[2] in each of the two dimensions of the 4x4 array. We first show that the results from the triple correlation approach to quantify spike rate (contributions of motif class 0, Fig. 2.5A) concur with the results reported in Hyvärinen et al (2019) (their Fig. 7C). Then for each pair of baseline and treatment wells, we calculated the ratio between $M/\mu_c - 1$ for the treatment well over the baseline well, and reported these values with a boxplot for each treatment (Fig. 2.5B-D). In these box-and-whisker plots, the center line represents the median (50th percentile), the bottom and top edges of the box represent the first (25th percentile) and third (75th percentile) quartiles, the whiskers represent the minimum and maximum range of data not considered outliers, and individual data points outside of the box-and-whisker plot represent outliers.

The drugs used in this study provide specific levers to control synaptic activity: GABA increases and gabazine decreases inhibitory synaptic function; CNQX blocks faster excitatory synapses (AMPA/kainate); and D-AP5 blocks slower excitatory synapses (NMDA). These have broadly straightforward effects on the network's firing rate: potentiating inhibitory synapses or blocking excitatory synapses decreases the firing rate, while blocking inhibitory synapses increases the firing rate (Fig. 2.5A; Hyvärinen et al (2019)) Modulating the inhibitory synapses has a similar

2 The spatial lags are asymmetric in order to cover the MEA without overlap, given periodic boundary conditions.

gross effect: excess inhibition suppresses contributions throughout, while decreasing inhibition leads to an increase in structured firing across the motif-class spectrum (Fig. 2.5B). When antagonizing particular excitatory receptors—with either faster or slower post-synaptic potentials—the triple correlation reflects changes in network behavior beyond simple firing rate modulation. On top of the gross reduction in firing, antagonizing faster excitatory AMPA/kainate receptors results in a decrease in synchrony and (most) higher-order motifs involving synchrony (Fig. 2.5C). On the other hand, antagonizing slower excitatory NMDA synapses, creating a network in which excitation is governed by faster synapses, results in a marked increase across the motif-class spectrum, despite the overall reduction in activity. The particularly high prevalence of third-order synchrony (motif class IV) reflects the fact that the remaining firing is overwhelmingly synchronous, well above the expectation governed by chance (Fig. 2.5D, Fig. 2.6). These results agree with the connectivity graphs comparing faster and slower correlations in AMPA and NMDA networks (Suresh et al, 2016, their Fig. 9).

2.4 Discussion

The essence of this study is the introduction of a new analytical tool that fully characterizes network activity using triplet motifs. Our mathematically unique statistical analysis not only encompasses typical first- and second-order analyses (e.g. auto-correlation, cross-correlation; Fig. 2.1), but also includes third-order correlations that are reflective of nonlinear network behaviors. In the preceding section, we demonstrated the robustness of our approach with networks of escalating complexity: in a simple synchronous test case (Fig. 2.2), in the face of substantial noise (Fig. 2.3), and in an experimental dataset (Fig. 2.5). We also validated our sim-

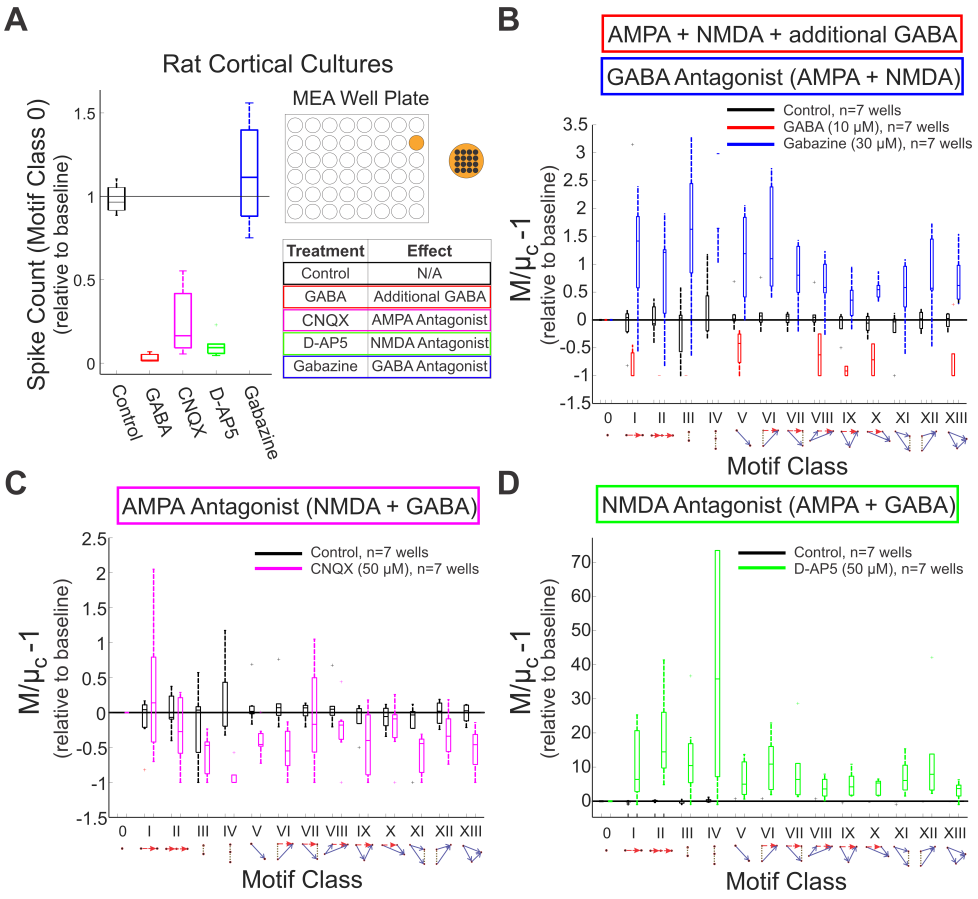


Figure 2.5: **Application of the triple correlation approach to experimental data.** We determined motif-class contributions in an open-source dataset of rat cortical cultures (Kapucu et al, 2022). Networks ($n = 70$ wells; $n = 35$ baseline and $n = 35$ treated) were cultured on microelectrode array (MEA) well plates—each well configured with 4x4 electrodes. The treated wells were exposed to the following experimental conditions ($n = 7$ wells per condition): control (black), GABA (red), gabazine (blue), CNQX (magenta), and D-AP5 (green). Each treatment well was matched to an untreated baseline well. For each motif class, we depict the distribution of ratios between the treatment and baseline wells. In these box-and-whisker plots, the centerline is the median, the bottom and top edges of the box are the first and third quartiles, and the whiskers extend to the minimum and maximum values. A) The normalized spike counts relative to baseline (contributions of motif class 0) are shown for each experimental condition. The table shows the effect of each of the experimental conditions. B) $M/\mu_c - 1$ values are shown for control-, GABA-, and gabazine-treated cultures. Note that some motif classes (IV, VI, VII, XI, XII) do not show values for GABA treatment due to lack of spiking. C) Results for control- and CNQX-treated cultures. D) The motif-class spectra for control and D-AP5-treated cultures indicate increased network structure for the latter, while its level of activity was reduced (Panel A).

ple summary metric in response to all computationally relevant motif classes (Fig. 2.4). The beauty of this approach lies in its flexible application to a multitude of finite data sets, including spike raster plots, fMRI images, local field potential (LFP), electroencephalogram (EEG), and data analyses in many other disciplines. Thus, we are excited to introduce this new analytical approach and eagerly anticipate its use as a tool to uncover new insights into network behavior.

Historically, neural network activity has been an important topic of investigation (McCulloch and Pitts, 1943). Hebb's foundational idea of cell assemblies, which aimed to link physiology and function (Hebb, 1949), have often been investigated by searching for repeated patterns of spiking. Typically, these cell assemblies are investigated in the context of searching for multispike pattern activation that may not be time-locked to any external stimulus or action, as "would be the case for internal processes like recalling a memory or planning a movement" (Abeles and Gerstein, 1988). The field has expanded considerably over the decades, fostering research in precise zero-phase lag synchronization (Abeles, 1991; Singer, 1999; Russo and Durstewitz, 2017), temporally-coded sequential patterns (Skaggs and McNaughton, 1996; Buzsáki and Draguhn, 2004), and synfire chain-like structures (Abeles, 1982; Hertz and Prügel-Bennett, 1996; Diesmann et al, 1999; Ikegaya et al, 2004). Our approach builds on this active area of research by integrating across multiple patterns within spatiotemporal motifs. Much as multispike patterns are putatively reflective of cell assemblies, these spatiotemporal motifs are potentially reflective of underlying structure, and thus, changes in structure could be indicative of network state shift, for example: the transition between normal and seizure states (Jirsa et al, 2014), preparation-to-movement in monkeys (Riehle et al, 1997), sleep/wake states (Saper et al, 2010), or neuromodulation in response to pharmacological agents (Iorio et al, 2013).

When applying our approach to investigate a particular hypothesis, there are two critical questions that should be asked: 1) what choice of spatiotemporal lags is relevant for the dataset (e.g. are there *a priori* synaptically relevant temporal lags?); and 2) depending on the null hypothesis, what choice of method for random chance firing is best suited for the dataset? The former presents an exciting and unexplored frontier that allows researchers to tailor this approach to a variety of hypotheses. The latter can be informed by prior literature either on null hypothesis distributions (to calculate theoretical expectations) or surrogate generation (to estimate the same). In order to present a foundational concept, we used Poisson processes to model our neurons, both in theory and in surrogate datasets. For future applications, when testing specific hypotheses using experimental data, a more nuanced surrogate (e.g. Stella et al (2022); Baker and Lemon (2000); Grün (2009)) should be used: e.g. a stimulated neuron will not obey a simple Poisson process. Thus, the particular surrogate would depend on the hypothesis being tested and must match the null hypothesis.

Among signal processing tools, a prominent example of another unique transform of neural data is the Fourier transform, with amplitudes and phases of frequencies as fundamental units (van Drongelen, 2018). Much of medical imaging relies on the uniqueness of the Fourier transform, including computerized tomography (CT) and magnetic resonance imaging (MRI), which uses the Fourier transform to generate images (van Drongelen, 2018). In EEG and LFP research, application of Fourier transform has progressed our understanding based in the frequency spectrum (van Drongelen, 2018). Since triple correlation also constitutes a unique characterization of neural data, we present our approach as a more complex yet still useful tool, as it is not only nonlinear and higher-order, but also comprises fundamental units, triplet motifs, that are still intuitively informative. Analogous to

the typical summary of the Fourier transform's frequency spectrum into frequency bands (alpha, beta, gamma, etc.), we summarize the triple correlation's motifs into a spectrum of fourteen motif classes. However, unlike the EEG frequency bands, which depend on clinically defined ranges, the motif-class spectrum arises directly from theory and is purely derived from possible spike sequences. It is a natural summary, and the classes themselves reflect their parsimony: they distinguish fundamental properties of computation, such as synchrony (motif classes III and IV), feedback (motif class IX), etc. (Fig. 2.1C). Furthermore, the constituent motif classes (I, III, and V) capture the well-known second-order correlations in analyzing neural spike data. Thus, while there are many alternative summaries, ours is both natural and useful in quantifying activity patterns (as exemplified in Figs. 2-5). Curiously, these theoretically defined fundamental units agree with previously data-driven work that has pointed to the primacy of third-order network motifs (Bojanek et al, 2020; Dechery and MacLean, 2018; Schneidman et al, 2006; Jiang et al, 2017).

Not only is triple correlation unique in the time domain, but its own Fourier transform, the bispectrum, is unique in the frequency domain. The bispectrum of the spike raster includes the well-known power spectrum and cross-spectrum (and coherence), the Fourier transform of autocorrelation (I) and cross-correlation (V) respectively for binary data (e.g. van Drongelen, 2018). The bispectrum's uniqueness gives theoretical weight to the growing consensus among researchers that insights about neural signals are encoded in the bispectrum (Bou Assi et al, 2018; Gagliano et al, 2019). Specifically, this underscores the importance of the fundamental elements of the bispectrum, which are the relationships between frequencies' amplitudes and phases. In some cases, as in our simulated rasters (Figs. 2-4), the spatial frequencies may not be relevant due to the arbitrary ordering of neuron rows in the rasters. In other cases, where the spatial dimension has a real ordering

(e.g. our MEA dataset; Fig. 2.5) the spatial frequency bands can be an untapped source of insight.

Both triplet motifs and inter-frequency analyses are already important topics of investigation in neuroscience (e.g. Jovanović and Rotter, 2016; Milo et al, 2002; Bojanek et al, 2020; Dechery and MacLean, 2018; Sporns and Kötter, 2004; Jansen et al, 2021). Here, we have provided a direct avenue for these investigations and proven the fundamental importance of these topics to any spatiotemporal neural data. Furthermore, the success of using the bispectrum as an input to artificial neural network seizure classifiers (Bou Assi et al, 2018) suggests that the equivalently-unique and also-meaningful triple correlation might also prove a good feature space on which to train machine learning algorithms. While we applied triple correlation to simulated and experimental spiking activity, our methodology can extend beyond spike rasters to even higher dimensionality rasters and to continuous-valued signals, such as multi-electrode LFP data or EEG recordings (see proof section "Proof in Online Materials"). Ultimately, just as frequency bands have been considered fundamental components of brain activity with Fourier transform, here we propose triplet motifs as new fundamental building blocks, one step more complex, that hold promise as an innovative approach to analyzing spatiotemporal neural activity across the breadth of recording modalities.

Data availability

All data used is publicly available from Kapucu et al (2022).

Code availability

The code used to generate all the results and figures in this paper is available on Github in the repository grahamas/RasterUniqueCharacterizationCode.

2.5 Supplemental Methods

2.5.1 Uniqueness of Triple-Correlation for Network Spiking Activity

We represent neuronal activity as a typical two-dimensional raster, $x(n,t)$, where n is neuron location and t is time. We assume that the raster is a binary (black and white, cf. Fig. 2.1A), meaning $x(n,t) = 1$ if neuron n fires at time t, and $x(n,t) = 0$ otherwise. We note that the reasoning in the proofs below works just as well for any raster taking bounded values (Yellott and Iverson, 1992) analogous to a greyscale image, such as would be the case with local field potential recordings or the electroencephalogram. Indeed the reasoning below applies to any finite bounded dataset.

The triple correlation of $x(n,t)$ is

$$c_3(\nu_1,\tau_1,\nu_2,\tau_2) = \iint x(n,t)x(n+\nu_1,t+\tau_1)x(n+\nu_2,t+\tau_2)\,dn\,dt \tag{2.1}$$

for spatial lags ν_1,ν_2 and temporal lags τ_1,τ_2. This is the triple correlation function that the TCU theorem shows uniquely characterizes the spike raster.

We define $\mathcal{F}[c_3]$ as the Fourier transform of c_3, which means

$$\mathcal{F}[c_3](\sigma_1,\omega_1,\sigma_2,\omega_2) = \iiiint c_3(\nu_1,\tau_1,\nu_2,\tau_2)e^{-j\sigma_1\nu_1}e^{-j\omega_1\tau_1}e^{-j\sigma_2\nu_2}e^{-j\omega_2\tau_2}\,d\nu_1\,d\tau_1\,d\nu_2\,d\tau_2 \tag{2.2}$$

We can rewrite this integral in terms of the Fourier transform of x by substituting

(2.1) into (2.2) and rearranging the integration order to find

$$\begin{aligned}\mathcal{F}[c_3](\sigma_1,\omega_1,\sigma_2,\omega_2) = \\ \iiiint x(n,t)x(n+\nu_2,t+\tau_2)\left[\iint x(n+\nu_1,t+\tau_1)e^{-j\sigma_1\nu_1}e^{-j\omega_1\tau_1}\,d\nu_1\,d\tau_1\right] \\ e^{-j\sigma_2\nu_2}e^{-j\omega_2\tau_2}\,d\nu_2\,d\tau_2\,dn\,dt.\end{aligned} \tag{2.3}$$

The part in between the [...] in (2.3) can be rewritten as $e^{j\omega_1 t}e^{j\sigma_1 n}X(\sigma_1,\omega_1)$, with $X(\sigma_1,\omega_1)$ denoting the two dimensional Fourier transform of $x(\nu_1,\tau_1)$. Similarly, the part of Equation (2.3) for ν_2 and τ_2 and their double integral can be arranged to evaluate to $e^{j\omega_2 t}e^{j\sigma_2 n}X(\sigma_2,\omega_2)$. Substitution of these results in (2.3) results in:

$$\mathcal{F}[c_3](\sigma_1,\omega_1,\sigma_2,\omega_2) = X(\sigma_1,\omega_1)X(\sigma_2,\omega_2)\iint x(n,t)e^{j(\omega_1+\omega_2)t}e^{j(\sigma_1+\sigma_2)n}\,dn\,dt \tag{2.4}$$

The double integral evaluates to $X(-\sigma_1-\sigma_2,-\omega_1-\omega_2) = X^*(\sigma_1+\sigma_2,\omega_1+\omega_2)$, where the asterisk denotes a complex conjugate. Thus we see that the Fourier transform of c_3 is the bispectrum:

$$\mathcal{F}[c_3(\nu_1,\tau_1,\nu_2,\tau_2)] = X(\sigma_1,\omega_1)X(\sigma_2,\omega_2)X^*(\sigma_1+\sigma_2,\omega_1+\omega_2). \tag{2.5}$$

This relationship between triple correlation and the bispectrum is the third-order equivalent of the Wiener-Khinchin-Einstein theorem.

Since images have a finite support and all of the above integration limits are implicitly at $(-\infty,\infty)$, we can modify the results for finite support by multiplying the spatiotemporal domain data by a two-dimensional boxcar window, w, limited between $(-\Sigma,\Sigma,-\Omega,\Omega)$ (or any other window with that support): $x_b = xw$. The frequency domain results are then characterized by the Fourier transforms convolved (denoted by $\circledast$) with the boxcar's Fourier transform (W) (or that of the window ap-

plied) denoted by: $X_B = X \circledast W$. By using this notation, the results in equations (2.1)–(2.5) would be adapted by adding the b and B subscripts.

Yellott and Iverson (1992) show in a constructive proof that a finite image (in our case a spike raster of a finite size) can be uniquely reconstructed from its third-order correlation. These authors also show and discuss how this does not hold for images of infinite size. Here we do not further discuss this aspect because the size of a spike raster (or a snapshot of any modality of neural activity) is always finite.

One critically important message of this paper is that the time domain's triple correlation and the corresponding bispectrum in the frequency domain uniquely determine the firing pattern of a network. Yellott (1993) presents this as the TCU theorem. If we apply Yellott's TCU theorem to a spike raster, we get the following.

Theorem 1. If $x(n, t)$ is a raster with bounded support and another raster $y(n, t)$ has the same triple correlation function as that of x, then $y(n, t) = x(n + a, t + b)$ for a pair of constants a, b.

Theorem 2. If $x(\boldsymbol{s})$ is a raster with bounded support in N spatial dimensions and another raster $y(\boldsymbol{s})$ has the same triple correlation function as that of x, then $y(\boldsymbol{s}) = x(\boldsymbol{s} + \boldsymbol{a})$ for a vector of constants $\boldsymbol{a}$.

Here we present proofs in line with the proof in Yellott and Iverson (1992) for two- and N-dimensional spatiotemporal data.

2.5.2 Proof of Theorem 1 (two dimensions)

Given the equality of third-order correlation functions, we can use Equation (2.5) to find that

$$X(\sigma_1, \omega_1)X(\sigma_2, \omega_2)X(-\sigma_1 - \sigma_2, -\omega_1 - \omega_2) = \tag{2.6}$$
$$Y(\sigma_1, \omega_1)Y(\sigma_2, \omega_2)Y(-\sigma_1 - \sigma_2, -\omega_1 - \omega_2)$$

To borrow some convenient results from probability theory (see any introductory text, e.g. Feller, 1950), we note that X and Y can be considered characteristic functions since we can consider x and y probability distributions: as finite images, x and y are bounded and nonnegative, and without loss of generality we can normalize them such that their integral is 1. Characteristic functions (the Fourier transforms of probability distributions) have two properties that are convenient for our purposes, the first of which is that they are non-zero in a region around the origin, thus allowing us the following division (for further considerations of rigor in the one dimensional case, cf. Moszner, 1980):

$$\frac{X(\sigma_1,\omega_1)X(\sigma_2,\omega_2)}{Y(\sigma_1,\omega_1)Y(\sigma_2,\omega_2)} = \frac{Y(-\sigma_1-\sigma_2,-\omega_1-\omega_2)}{X(-\sigma_1-\sigma_2,-\omega_1-\omega_2)}. \tag{2.7}$$

Next we rewrite both complex functions in terms of their amplitude and phase, i.e. $X(\sigma,\omega) = |X(\sigma,\omega)|\, e^{j\phi(X(\sigma,\omega))}$, where ϕ gives the phase. So then we can rewrite the right-hand side of (2.7)

$$\frac{X(\sigma_1,\omega_1)X(\sigma_2,\omega_2)}{Y(\sigma_1,\omega_1)Y(\sigma_2,\omega_2)} = \frac{|Y(-\sigma_1-\sigma_2,-\omega_1-\omega_2)|\, e^{j\phi(Y(-\sigma_1-\sigma_2,-\omega_1-\omega_2))}}{|X(-\sigma_1-\sigma_2,-\omega_1-\omega_2)|\, e^{j\phi(X(-\sigma_1-\sigma_2,-\omega_1-\omega_2))}}. \tag{2.8}$$

Here we use the second convenient property of characteristic functions, namely that they are Hermitian, i.e. $X(-\sigma,-\omega) = X^*(\sigma,\omega)$. By setting $\sigma_2 = 0, \omega_2 = 0$ in (2.8), we can use this Hermitian property to derive the fact that $|X(\sigma_1,\omega_1)|^2 = |Y(\sigma_1,\omega_1)|^2$ for any σ_1,ω_1. In particular, $|X(-\sigma_1-\sigma_2,-\omega_1-\omega_2)| = |Y(-\sigma_1-\sigma_2,-\omega_1-\omega_2)|$ so we can flip those terms.

$$= \frac{|X(-\sigma_1-\sigma_2,-\omega_1-\omega_2)|\, e^{j\phi(Y(-\sigma_1-\sigma_2,-\omega_1-\omega_2))}}{|Y(-\sigma_1-\sigma_2,-\omega_1-\omega_2)|\, e^{j\phi(X(-\sigma_1-\sigma_2,-\omega_1-\omega_2))}} \tag{2.9}$$

We can also rewrite those same terms thanks to the same Hermitian property.

$$= \frac{|X^*(\sigma_1+\sigma_2,\omega_1+\omega_2)|\, e^{j\phi(Y^*(\sigma_1+\sigma_2,\omega_1+\omega_2))}}{|Y^*(\sigma_1+\sigma_2,\omega_1+\omega_2)|\, e^{j\phi(X^*(\sigma_1+\sigma_2,\omega_1+\omega_2))}} \tag{2.10}$$

Simple complex properties are $|X| = |X^*|$ and $\phi(X) = -\phi(X^*)$, which give us

$$= \frac{|X(\sigma_1+\sigma_2,\omega_1+\omega_2)|\, e^{j\phi(X(\sigma_1+\sigma_2,\omega_1+\omega_2))}}{|Y(\sigma_1+\sigma_2,\omega_1+\omega_2)|\, e^{j\phi(Y(\sigma_1+\sigma_2,\omega_1+\omega_2))}} \tag{2.11}$$

Next, we return from the amplitude and phase notation to the complex functions themselves:

$$= \frac{X(\sigma_1+\sigma_2,\omega_1+\omega_2)}{Y(\sigma_1+\sigma_2,\omega_1+\omega_2)}. \tag{2.12}$$

Now we define $H(\sigma,\omega) = \frac{X(\sigma,\omega)}{Y(\sigma,\omega)}$ so we can rewrite (2.12) as

$$H(\sigma_1,\omega_1)H(\sigma_2,\omega_2) = H(\sigma_1+\sigma_2,\omega_1+\omega_2). \tag{2.13}$$

Therefore by a basic result of complex analysis $H(\sigma,\omega) = e^{j(a\sigma+b\omega)}$, and thus

$$Y(\sigma,\omega) = X(\sigma,\omega)e^{j(a\sigma+b\omega)}. \tag{2.14}$$

This holds in a region near the origin, and another property of characteristic functions is that, if their probability distribution has finite support, the probability distribution is uniquely determined by the value of the characteristic function in a region of the origin. Thus, in the spatiotemporal domain,

$$y(n,t) = x(n+a,t+b) \text{ with constants } a,b. \tag{2.15}$$

2.5.3 Proof of Theorem 2 (N dimensions)

Here we extend the preceding proof to N dimensions. This would be necessary for analysing many clinical and experimental data sources, e.g. the plane of a two-dimensional micro-electrode array produces $2+1$ dimensional data (as in our example presented in Fig. 2.5). Even higher dimensions may be useful in the case where long-range connections can be represented as a hidden dimension (e.g. orientation tuning in V1). The proof is generic enough that in practice the result applies to all finite datasets.

Note that the number of motifs (i.e. the size of the triple correlation matrix) scales with exponent $2N + 2$.

We define the $N + 1$-dimensional spatiotemporal vector $\boldsymbol{s} = (x_1, x_2, \ldots, x_N, t)$. Then we notate the triple correlation of an N dimensional raster as $c_T(\boldsymbol{s}_1, \boldsymbol{s}_2)$. We notate the $N + 1$-dimensional variable's Fourier transform as

$$\mathcal{F}[\boldsymbol{s}] = \boldsymbol{\Omega} = (\sigma_1, \sigma_2, \ldots, \sigma_N, \omega), \tag{2.16}$$

and the corresponding bispectrum as $C_T(\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2)$.

Since the properties of characteristic functions that are key to this proof still hold in N dimensions (see any introductory probability theory text, e.g. Feller, 1950), the proof proceeds identically to that in two dimensions, but with vector notation.

Given the equality of third-order correlation functions, we can use Equation (2.5) to find that

$$X(\boldsymbol{\Omega}_1)X(\boldsymbol{\Omega}_2)X(-\boldsymbol{\Omega}_1 - \boldsymbol{\Omega}_2) = Y(\boldsymbol{\Omega}_1)Y(\boldsymbol{\Omega}_2)Y(-\boldsymbol{\Omega}_1 - \boldsymbol{\Omega}_2) \tag{2.17}$$

To borrow some convenient results from probability theory, we note that X and Y can be considered characteristic functions since we can consider x and y probability distributions: as finite images, x and y are bounded and nonnegative, and without

loss of generality we can normalize them such that their integral is 1. Characteristic functions (the Fourier transforms of probability distributions) have two properties that are convenient for our purposes, the first of which is that they are non-zero in a region around the origin, thus allowing us the following division (for further considerations of rigor in the one dimensional case, cf. Moszner, 1980):

$$\frac{X(\boldsymbol{\Omega}_1)X(\boldsymbol{\Omega}_2)}{Y(\boldsymbol{\Omega}_1)Y(\boldsymbol{\Omega}_2)} = \frac{Y(-\boldsymbol{\Omega}_1 - \boldsymbol{\Omega}_2)}{X(-\boldsymbol{\Omega}_1 - \boldsymbol{\Omega}_2)}. \tag{2.18}$$

Next we rewrite both complex functions in terms of their amplitude and phase, i.e. $X(\boldsymbol{\Omega}) = |X(\boldsymbol{\Omega})|\, e^{j\phi(X(\boldsymbol{\Omega}))}$, where ϕ gives the phase. So then we can rewrite the right-hand side of (2.18)

$$\frac{X(\boldsymbol{\Omega}_1)X(\boldsymbol{\Omega}_2)}{Y(\boldsymbol{\Omega}_1)Y(\boldsymbol{\Omega}_2)} = \frac{|Y(-\boldsymbol{\Omega}_1 - \boldsymbol{\Omega}_2)|\, e^{j\phi(Y(-\boldsymbol{\Omega}_1 - \boldsymbol{\Omega}_2))}}{|X(-\boldsymbol{\Omega}_1 - \boldsymbol{\Omega}_2)|\, e^{j\phi(X(-\boldsymbol{\Omega}_1 - \boldsymbol{\Omega}_2))}}. \tag{2.19}$$

Here we use the second convenient property of characteristic functions, namely that they are Hermitian, i.e. $X(-\boldsymbol{\Omega}) = X^*(\boldsymbol{\Omega})$. By setting $\boldsymbol{\Omega}_2 = \mathbf{0}$ in (2.19), we can use this Hermitian property to derive the fact that $|X(\boldsymbol{\Omega}_1)|^2 = |Y(\boldsymbol{\Omega}_1)|^2$ for any $\boldsymbol{\Omega}_1$. In particular, $|X(-\boldsymbol{\Omega}_1 - \boldsymbol{\Omega}_2)| = |Y(-\boldsymbol{\Omega}_1 - \boldsymbol{\Omega}_2)|$ so we can flip those terms.

$$= \frac{|X(-\boldsymbol{\Omega}_1 - \boldsymbol{\Omega}_2)|\, e^{j\phi(Y(-\boldsymbol{\Omega}_1 - \boldsymbol{\Omega}_2))}}{|Y(-\boldsymbol{\Omega}_1 - \boldsymbol{\Omega}_2)|\, e^{j\phi(X(-\boldsymbol{\Omega}_1 - \boldsymbol{\Omega}_2))}} \tag{2.20}$$

We can also rewrite those same terms thanks to the same Hermitian property.

$$= \frac{|X^*(\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2)|\, e^{j\phi(Y^*(\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2))}}{|Y^*(\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2)|\, e^{j\phi(X^*(\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2))}} \tag{2.21}$$

Simple complex properties are $|X| = |X^*|$ and $\phi(X) = -\phi(X^*)$, which give us

$$= \frac{|X(\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2)|\, e^{j\phi(X(\boldsymbol{\Omega}_1+\boldsymbol{\Omega}_2))}}{|Y(\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2)|\, e^{j\phi(Y(\boldsymbol{\Omega}_1+\boldsymbol{\Omega}_2))}} \tag{2.22}$$

Next, we return from the amplitude and phase notation to the complex functions themselves:

$$= \frac{X(\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2)}{Y(\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2)}. \tag{2.23}$$

Now we define $H(\boldsymbol{\Omega}) = \frac{X(\boldsymbol{\Omega})}{Y(\boldsymbol{\Omega})}$ so we can rewrite (2.23) as

$$H(\boldsymbol{\Omega}_1)H(\boldsymbol{\Omega}_2) = H(\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2). \tag{2.24}$$

Therefore by a basic result of complex analysis $H(\boldsymbol{\Omega}) = e^{j\boldsymbol{k}\cdot\boldsymbol{\Omega}}$, and thus

$$Y(\boldsymbol{\Omega}) = X(\boldsymbol{\Omega})e^{j\boldsymbol{k}\cdot\boldsymbol{\Omega}}. \tag{2.25}$$

This holds in a region near the origin, and another property of characteristic functions is that, if their probability distribution has finite support, the probability distribution is uniquely determined by the value of the characteristic function in a region of the origin. Thus, in the spatiotemporal domain,

$$y(\boldsymbol{s}) = x(\boldsymbol{s} + \boldsymbol{a}) \text{ with vector of constants } \boldsymbol{a}. \tag{2.26}$$

2.5.4 Computing the triple correlation

For computational applications presented here, we consider third-order correlation analysis applied to a discrete raster with neuron rows n, $1:N$, and time columns t,

$1:T$, where each pixel $r(n,t)$ is filled with a 0 (white for no activity) or 1 (black for spike). The complete triple correlation function, while formally requisite, incurs substantial computational cost while adding increasingly noisy information. We limit our computation to lag windows $-W_t : W_t$ and $-W_n : W_n$ chosen such that two local synaptic connections could not take longer than W_t, and two neurons are unlikely to be connected further than W_n. With this restricted lag window, we do not zero pad our raster (as in Yellott and Iverson, 1992) but instead use boundary conditions periodic in space (since our spatial ordering was already arbitrary[3]) and restrict our calculation to a subset of time such that no motifs extend beyond time zero or the raster's duration. In this case, discrete equivalent cd_3 of the triple correlation expression c_3 (Equation 2.1) or the short form reported in the main text, $c_3(n_1,t_1,n_2,t_2) = \langle r(n,t)r(n+n_1,t+t_1)r(n+n_2,t+t_2)\rangle_{n,t}$, is

$$cd_3(n_1,t_1,n_2,t_2) = \tag{2.27}$$

$$\frac{1}{\#(W_n,W_t,N,T)} \sum_{t=1+W_t}^{T-W_t} \sum_{n=1}^{N} \tag{2.28}$$

$$r(n,t)r((n+n_1-1)\%N+1,t+t_1)r((n+n_2-1)\%N+1,t+t_2).$$

We scale by the number of spike bins in the summation

$$\#(W_n,W_t,N,T) = (T-2W_t)(N). \tag{2.29}$$

Equations (2.27) and (2.29) were used in all our simulations.

3 In the MEA case, we also use periodic spatial boundary conditions even though the ordering is not arbitrary. Zeropadding introduces problems as spikes near the center of the array are more "valuable" than spikes near the boundaries (they contribute to more motifs). An alternative would be to introduce noise past the boundaries, but that creates its own problems in the choice of noise. Since our analyses already assumed a lack of spatial information (in our choice of motif class) we decided to simply discard the spatial information for the purposes of these analyses, and therefore to use the spatial periodic boundary conditions.

Note that in the text we report triple correlations "using lags up to X in time and space." This corresponds to $W_n = X/2$ and $W_t = X/2$. When X is odd, we use the more general lag windows $-\mathrm{floor}(X/2) : \mathrm{ceil}(X/2)$, with corresponding changes to the summation bounds and scaling instead by $(T - X)(N)$. We report the maximum lag as X because, e.g., $cd_3(-10, t_1, 10, t_2)$ has a maximum temporal separation between spikes of 20, despite the fact that technically the lag arguments are at most 10.

2.5.5 Summarizing triple correlation into motif classes

We summarized the triple correlation as fourteen motif-classes (Fig. 2.1). The main text discusses the reasoning for this particular choice, and here we describe the details.

Let $[n_1, t_1|n_2, t_2]$ denote a motif, i.e. an argument to triple correlation, which would have a value $c_3(n_1, t_1, n_2, t_2)$. We want to group together these motifs according to what we can qualitatively distinguish. We only hold two qualitative distinctions:

1. temporal: we distinguish between before, simultaneous, and after (temporal coordinate less than, equal to, or greater)

2. spatial: we distinguish between the same and different (spatial coordinate equal and spatial coordinate not equal)

We will interpret the motif as a three node graph: one base node implicitly (0,0), and two other nodes, (n_1, t_1) and (n_2, t_2). Note that neither of our qualitative distinctions involve node identity. So, for example, we would not distinguish between $[0, 1|0, 2]$ and $[0, 2|0, 1]$, even though one the first node follows the second, and in the

other the reverse is true. In both, the motif includes nodes at the same spatial coordinate in sequence, so we do not distinguish between them.

By grouping motifs that are indistinguishable under these criteria, we develop motif classes. We enumerate these motif classes in Fig. 2.1. From that enumeration, it is clear that no motif can belong to two motif classes. To show that this enumeration is complete (there are no more motif classes and all motifs fall into one of these classes), we count the number of motif classes below. To begin, we count "lag-sign" motifs, which is to say, if we only care about the sign of the lags, how many motifs are there. This first step avoids some of the small complications in our definition, i.e. the different treatments of space and time, as well as the subtle ways they interact when we discard node identity. Having counted lag-sign motifs, we then reduce these lag-sign motifs by discounting spatial direction and node identity to finally arrive at the number of motif classes.

2.5.6 The number of lag-sign motifs

Given that three node combinations determine c_3, we want to know how many ways there are to order them in time and space, with the caveat that we may also place one node on top of another in either time, or space, or both (to account for zero lag). We have enumerated all such possibilities in Table A1. Here, we count them to prove our enumeration is complete. To do so, we divide the cases according to how many different temporal and spatial bins the nodes have between them. For example, if all the nodes are completely distinct in time and space, then there are both three spatial bins and three temporal bins between the three nodes. On the other hand, if we overlap all three nodes in both time and space, then there is only one spatial bin and one temporal bin between them. All the cases are labeled in Table 2.1. Note that the cases are symmetric, so we will calculate the values as in

the upper left half of the table, i.e. where the number of spatial bins is greater than or equal to the number of time bins. The arrangement of this table is the same as in Fig. 2.1C. As described above, we call this arrangement of nodes separated by lags a motif. When we only note the sign of the lags, we call it a lag-sign motif.

A. In the case where no nodes share a spatial bin (neuron), there are $3!$ ways to order them in space. Similarly if the nodes share no time bins, then there are $3!$ ways to order the time bins. So, if no spatial bins nor time bins are shared, then there are $3!3! = 36$ orderings.

B. Given that the nodes only have two distinct spatial bins between them, there are three ways to choose which nodes share the same spatial bin, and then two ways to choose whether the remaining node has a spatial bin before or after the identical nodes' spatial bin, for $3 \cdot 2 = 6$ spatial orderings. As above, if no time bins are identical, then there are $3!$ temporal orderings. So there are $3 \cdot 2 \cdot 3! = 36$ orderings.

C. If nodes share the same spatial bin, then there is only one way to order that neuron. As above, if no time bins are identical then there are $3!$ temporal orderings, giving a total of $1 \cdot 3! = 6$ orderings. The same is true in the symmetric case when the nodes share the same time bin, but all have different neurons.

D. We already established that for the case where two spatial bins are the same, there are $3 \cdot 2 = 6$ orderings. The same is true when the nodes only have two distinct time bins. So there are $3 \cdot 2 \cdot 3 \cdot 2 = 36$ orderings.

E. We already established that for the case where two spatial bins are the same, there are $3 \cdot 2 = 6$ orderings, and for the case where all time bins are the same, there is only one ordering. So there are $3 \cdot 2 \cdot 1 = 6$ orderings.

F. In the case where all spatial bins and all time bins are the same, there is only one ordering.

We sum up the table: $A + 2B + 2C + D + 2E + F = 36 + 2 \cdot 36 + 2 \cdot 6 + 36 + 2 \cdot 6 + 1 = 169$ lag-sign motifs.

2.5.7 The number of motif classes

To complete our summary, we reduce the 169 lag-sign motifs counted in the previous section by appealing to two symmetries: one in space, and one in permutation. The spatial one is straightforward: since we have not assigned meaning to space, reordering the spatial bins does not change our interpretation. The permutation symmetry is more technical: the way triple correlation is calculated, the ordering of the nodes matters. In second-order correlation (call it ρ), this is even symmetry: $\rho(x) = \rho(-x)$. More generally, a motif has some polygonal structure (or possibly a point or a line; treat these as straightforward special cases of the following argument). Each vertex of the polygon has a node identity: in the triple correlation case, one is the base node, one is the first node, given by the base node plus (n_1, t_1), and one is the second node, given by the base node plus (n_2, t_2). But we don't actually care which is the base, which is the first, and which is the second node: triple correlation is invariant under permutation of these nodes. We reduce the number of lag-sign motifs by accounting for these symmetries. Again working from Table 2.1:

A. There are $3!$ ways to permute node identity, and $3!$ ways to permute neuron identity, independently, giving a product of 36 permutations. So all 36 lag-sign motifs in category A reduce to 1 motif class (Fig. 2.1C, XIII).

B(2T). For category B with two time bins, there are $3!$ node permutations, and only 3 spatial permutations: two spatial bins share a time bin, and the third spatial

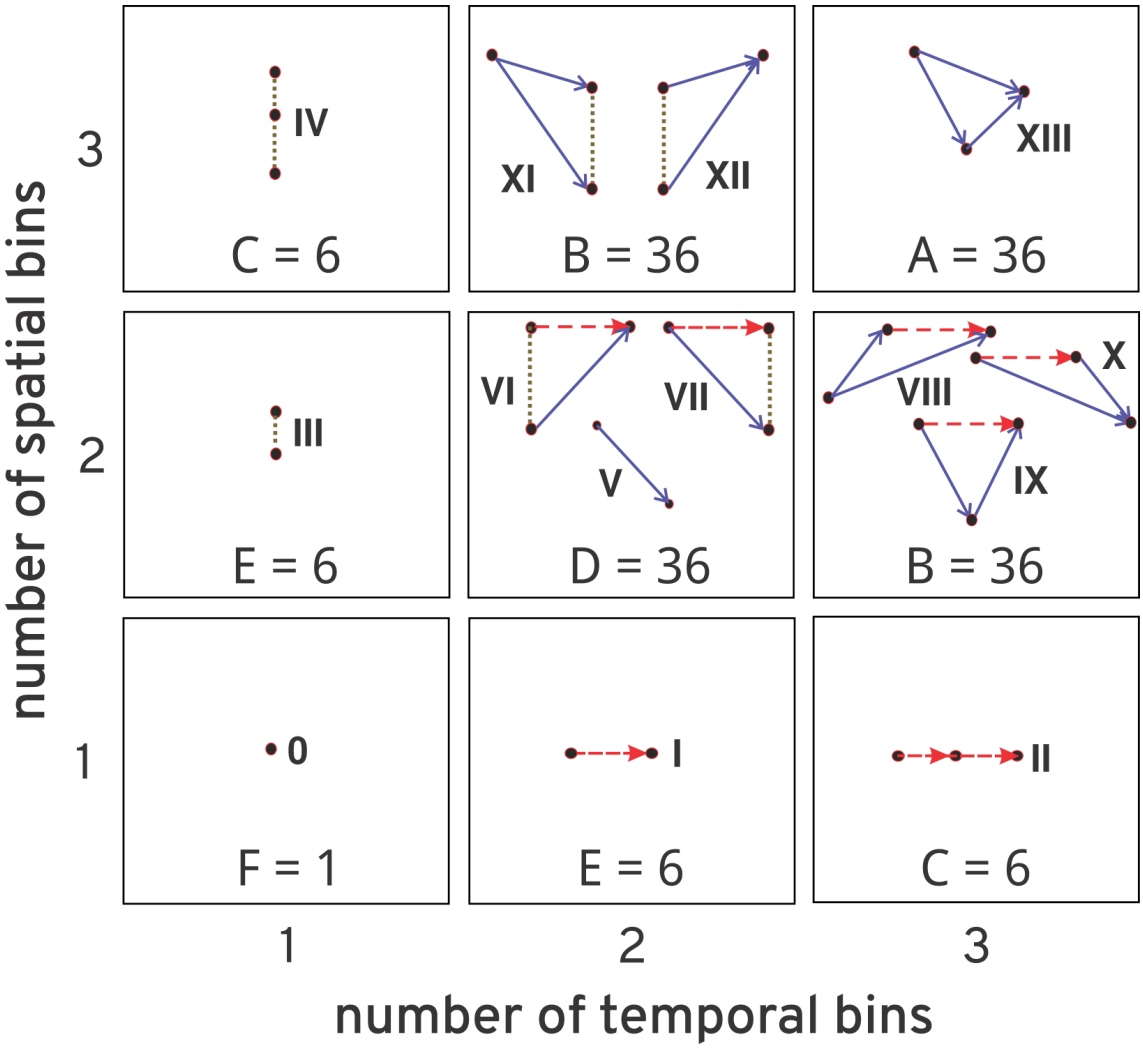


Table 2.1: **Counting the number of lag-sign motifs per motif class.** The number of possible orderings in space and time given two possible time lags and two possible spatial lags, which corresponds to up to three possible bins in both time and space. Note that in case of a spike raster of single unit activity, each spatial bin corresponds to one neuron. This table corresponds to Fig. 2.1C in the main text.

bin may appear above, between, or below these two spatial bins (due to the fact that we are reducing from lag-sign motifs, not from lag-motifs). The product is 18 total permutations, so the 36 lag-sign motifs in category B(2T) reduce to 2 motif classes (Fig. 2.1C, XI and XII).

B(2S). For category B with two spatial bins, there are only two spatial permutations: one spatial bin has two time bins, and the remaining spatial bin can either be above or below. There are still $3!$ node permutations, leading to a total of 12 permutations. Therefore the 36 lag-sign motifs in category B(2S) reduce to 3 motif classes (Fig. 2.1C, VIII, IX, and X).

D. By the same argument as in B(2S), there are two spatial permutations, and again $3!$ node permutations, leading to a total of 12 permutations. Therefore the 36 lag-sign motifs in category D reduce to 3 motif classes (Fig. 2.1C, V, VI, and VII).

C(1T). When there is only one time bin, spatial permutation is the same as node permutation, so there are only the $3! = 6$ permutations. Therefore the 6 lag-sign motifs in category C(1T) reduce to 1 motif class (Fig. 2.1C, IV).

E(1T). As in C(1T), all 6 reduce to 1 motif class (Fig. 2.1C, III).

C(1S). When there is only one spatial bin, the lag-sign motifs have no spatial permutations, only node permutations, so the 6 lag-sign motifs in category C(1S) reduce to 1 motif class (Fig. 2.1C, II).

E(1S). As in C(1S), all 6 reduce to 1 motif class (Fig. 2.1C, I)

F. In the case where all lags are zero, triple correlation has no symmetric entries but the zero entry, so there are no permutations, and the 1 lag-sign motif (which also contains only one lag-motif) remains 1 motif class (Fig. 2.1C, 0).

The above sum to 14 motif classes, as depicted in Fig. 2.1.

2.5.8 Expected contributions per motif class

Because it is important to asses how these motif-class contributions be explained by chance (or not), we simulate 100 activity matched noise rasters for each raster under investigation, and compare the raster's motif-class spectrum with that of its noise matched rasters. In the following, we derive the theoretical expectations for rasters governed by noise.

Let r be a raster with N neurons and T time bins. In a binary raster, a spike is indicated by $r(n,t) = 1$, else $r(n,t) = 0$. We calculate the triple correlation c_3 over the whole raster with periodic boundary conditions (i.e. $r(n+\eta, t+\tau) = r((n+\eta) \mod N, (t+\tau) \mod T)$, with an additional 1 offset). Define $\pi(\boldsymbol{x}, \boldsymbol{\lambda}) = r(\boldsymbol{x})r(\boldsymbol{x}+\boldsymbol{\lambda}_1)r(\boldsymbol{x}+\boldsymbol{\lambda}_2)$ and for shorthand $\pi(\boldsymbol{\lambda}) = \pi(0, \boldsymbol{\lambda})$. In general with lags $\boldsymbol{\lambda} = (\eta_1, \tau_1, \eta_2, \tau_2)$

$$\begin{aligned} c_3(\boldsymbol{\lambda}) &= c_3(\eta_1, \tau_1, \eta_2, \tau_2) \\ &= \sum_{(n,t)\in I(r)} r(n,t)r(n+\eta_1, t+\tau_1)r(n+\eta_2, t+\tau_2) \\ &= \sum_{\boldsymbol{x}\in I(r)} \pi(\boldsymbol{x}, \boldsymbol{\lambda}) \end{aligned}$$

where $I(r)$ is the set of all indices of r, in this case $I(r) = (1:N) \times (1:T)$. We are interested in the motif class contributions, which are the sums of triple correlations for lags that constitute each of the fourteen motifs. So, if $\boldsymbol{\lambda} \in \Delta(M_i)$ denotes that the lags $\boldsymbol{\lambda}$ constitute a triplet in motif class i, then we define the motif class contribution as

$$M_i(r) = \sum_{\boldsymbol{\lambda}\in\Delta(M_i)} c_3(\boldsymbol{\lambda})$$

The expectation then is given by

$$\begin{aligned}
\langle M_i(r) \rangle &= \left\langle \sum_{\boldsymbol{\lambda} \in \Delta(M_i)} c_3(\boldsymbol{\lambda}) \right\rangle \\
&= \left\langle \sum_{\boldsymbol{\lambda} \in \Delta(M_i)} \sum_{x \in I(r)} \pi(x, \boldsymbol{\lambda}) \right\rangle \\
&= \sum_{\boldsymbol{\lambda} \in \Delta(M_i)} \sum_{x \in I(r)} \langle \pi(x, \boldsymbol{\lambda}) \rangle \\
&= \sum_{x \in I(r)} \sum_{\boldsymbol{\lambda} \in \Delta(M_i)} \langle \pi(x, \boldsymbol{\lambda}) \rangle \\
&= NT \sum_{\boldsymbol{\lambda} \in \Delta(M_i)} \langle \pi(\boldsymbol{\lambda}) \rangle
\end{aligned}$$

In fact, $\langle \pi(\boldsymbol{\lambda}) \rangle$ is constant for all $\boldsymbol{\lambda} \in \Delta(M_i)$ for a given i, so

$$\langle M_i(r) \rangle = NT \#(\Delta(M_i)) \langle \pi(\lambda_i) \rangle . \tag{2.30}$$

We can easily calculate $\langle \pi(\boldsymbol{\lambda}) \rangle$ if we assume independent bins. See below for the calculation in each of the cases of independent Bernoullis (spiking data) and independent standard normals (LFP, EEG).

$\#(\Delta(M_i))$ is the more challenging problem, which we calculate case-by-case below in the section "The number of triplet motifs."

Independent spiking with constant probability (Bernoulli)

For a simple spiking raster, we assume all bins have independent probability p of spiking, i.e. they are all independent Bernoullis.[4] This lets us substitute into equa-

4 This also means that all the neurons are simulated Poisson processes with identical rates.

tion 2.30 the following probabilities

$$\begin{aligned}
\langle \pi(\boldsymbol{\lambda}_0) \rangle &= p \\
\langle \pi(\boldsymbol{\lambda}_i) \rangle &= p^2 \qquad \text{for } i = I, III, V \\
\langle \pi(\boldsymbol{\lambda}_i) \rangle &= p^3 \qquad \text{for } i = II, IV, VI, VII, VIII, IX, X, XI, XII, XIII
\end{aligned}$$

where $\boldsymbol{\lambda}_i$ is a representative triplet of motif class i. This representative is sufficient because the expectation of the product of bins purely depends on the number of distinct bins multiplied. Within a motif class, all triplets have the same number of distinct bins, with motif class 0 having only one spike, motif classes $I, III,$ and IV having two spikes, and the remaining motif classes having three spikes. Knowing these values, we have the expectations of each motif class.

Independent standard normals (EEG)

For a standard normal, we have that $\langle \pi(\boldsymbol{\lambda}_j) \rangle = 0$ for all j. This is because the expectation of a product of standard normals raised to some powers is zero if any of those powers is odd. Formally, $\langle \prod_i r_i^{\alpha_i} \rangle = 0$ iff α_i is odd for any i. The expectation $\langle r(0,0) r(\boldsymbol{\lambda}_1) r(\boldsymbol{\lambda}_2) \rangle$ must be zero because it always has a standard normal raised to either the first or third power. Thus $\langle M_i(r) \rangle = 0$ when r is a matrix of independent standard

The number of triplet motifs

The quantity $\#(\Delta(M_i))$ simply counts the number of motifs in a given motif class. This is independent of the data's distribution, but dependent on the data's shape. Let Λ_t be the number of time lags considered, and let $\Lambda_t^{\pm}$, Λ_t^{+}, and Λ_t^{-} be the number of time lags that are nonzero, positive, and negative respectively. Similarly define

Λ_n, $\Lambda_n^{\pm}$, Λ_n^{+}, and Λ_n^{-}.

Our first step for each motif class will be to define $\Delta(M_i)$, for which purpose we will define a notation to indicate the sign and equality of each part of the two spatiotemporal lags (the third lag implicitly zero). For example, in the case of motif class XIII, we write $\Delta(M_{XIII}) = [x^{\pm}, t^{\pm} \mid y^{\pm}, s^{\pm}]$. This indicates that all four lags are nonzero ($\pm$ superscript) and none are equal ($x \neq y$ and $t \neq s$). In contrast, we could write $[0, t^{+} \mid 0, t^{+}]$ to indicate both spatial lags are zero, and the temporal lags are both positive and equal. We add a coefficient "2" to indicate that we want to additionally include set of lags symmetric with the notated lag, i.e. flipping which lag is first and which is second, e.g. $2[0, 0 \mid 0, t^{\pm}] = [0, 0 \mid 0, t^{\pm}] + [0, t^{\pm} \mid 0, 0]$.

I Motif class I consists in lags $\Delta(M_I) = 2[0, 0 \mid 0, t^{\pm}] + [0, t^{\pm} \mid 0, rt^{\pm}]$. We count these as follows: there is only one way to choose the $(0, 0)$ lag, and $\Lambda_t^{\pm}$ ways to choose a $(0, t^{\pm})$ lag. Therefore there are $1 \cdot \Lambda_t^{\pm}$ lags in $[0, 0 \mid 0, t^{\pm}]$. By symmetry there are the same number in $[0, t^{\pm} \mid 0, 0]$, so the coefficient 2 works as advertised, giving $2\Lambda_t^{\pm}$ lags in $2[0, t^{\pm} \mid 0, 0]$. For the second addend, there are $\Lambda_t^{\pm}$ ways to choose the $(0, t^{\pm})$ lag. Since the lags are equal there is only one choice for the second lag, leading to a $\Lambda_t^{\pm}$ lags in $[0t^{\pm} \mid 0t^{\pm}]$. In total then, there are $3\Lambda_t^{\pm}$ lags in $\Delta(M_I) = 2[0, 0 \mid 0, t^{\pm}] + [0t^{\pm} \mid 0t^{\pm}]$. We write this

$$\begin{aligned} \#(\Delta(M_I)) &= \#(2[0, 0 \mid 0, t^{\pm}] + [0t^{\pm} \mid 0t^{\pm}]) \\ &= 2(1 \cdot \Lambda_t^{\pm}) + (\Lambda_t^{\pm} \cdot 1) \\ &= 3\Lambda_t^{\pm} \end{aligned}$$

II For motif class II, we have $\Delta(M_{II}) = [0, t^{\pm} \mid 0, s^{\pm}]$. This time for our second lag, we have one fewer choice because we must have that the temporal lags do not

equal. So we find

$$\begin{aligned}\#(\Delta(M_{II})) &= \#([0, t^{\pm} \mid 0, s^{\pm}]) \\ &= \Lambda_t^{\pm}(\Lambda_t^{\pm} - 1)\end{aligned}$$

III For motif class III, the situation is the same as motif class I, but with n rather than t, so $\Delta(M_{III}) = 2[0, 0 \mid x^{\pm}, 0] + [x^{\pm}, 0 \mid x^{\pm}, 0]$, giving

$$\begin{aligned}\#(\Delta(M_{III})) &= \#(2[0, 0 \mid x^{\pm}, 0] + [x^{\pm}, 0 \mid x^{\pm}, 0]) \\ &= 2(1 \cdot \Lambda_n^{\pm}) + (\Lambda_n^{\pm} \cdot 1) \\ &= 3\Lambda_n^{\pm}\end{aligned}$$

.

IV Similarly motif class IV has $\#(\Delta(M_{III})) = \Lambda_n^{\pm}(\Lambda_n^{\pm} - 1)$.

V

$$\begin{aligned}\Delta(M_V) &= 2[0, 0 \mid x^{\pm}, t^{\pm}] + [x^{\pm}, t^{\pm} \mid x^{\pm}, t^{\pm}] \\ \#(\Delta(M_V)) &= 2\Lambda_n^{\pm}\Lambda_t^{\pm} + \Lambda_n^{\pm}\Lambda_t^{\pm} \\ &= 3\Lambda_n^{\pm}\Lambda_t^{\pm}\end{aligned}$$

VI

$$\begin{aligned}\Delta(M_{VI}) &= 2[0, t^{+} \mid x^{\pm}, 0] + 2[x^{\pm}, 0 \mid x^{\pm}t^{+}] + 2[0, t^{-} \mid x^{\pm}t^{-}] \\ \#(\Delta(M_{VI})) &= 2(\Lambda_t^{+}\Lambda_n^{\pm}) + 2(\Lambda_n^{\pm}\Lambda_t^{+}) = 2(\Lambda_t^{-}\Lambda_n^{\pm}) \\ &= 4(\Lambda_n^{\pm}\Lambda_t^{+}) + 2(\Lambda_n^{\pm}\Lambda_t^{-})\end{aligned}$$

VII By symmetry with VI,

$$\#(\Delta(M_{VII})) = 4(\Lambda_n^{\pm}\Lambda_t^{-}) + 2(\Lambda_n^{\pm}\Lambda_t^{+})$$

VIII

$$\begin{aligned}
\Delta(M_{VIII}) &= [x^{\pm}, t^{+} \mid x^{\pm}, s^{+}] + 2[x^{\pm}, t^{-} \mid 0, s^{+}] + 2[x^{\pm}, t^{-} \mid 0, s^{-}(> t^{-})] \\
\#(\Delta(M_{VIII})) &= \Lambda_n^{\pm}\Lambda_t^{+}(\Lambda_t^{+} - 1) + 2\Lambda_n^{\pm}\Lambda_t^{-}\Lambda_t^{+} + 2\Lambda_n^{\pm} \sum_{\Lambda_t^{-} < t^{-} < -1} \#s : t^{-} < s < 0 \\
&= \Lambda_n^{\pm}\Lambda_t^{+}(\Lambda_t^{+} - 1) + 2\Lambda_n^{\pm}\Lambda_t^{-}\Lambda_t^{+} + 2\Lambda_n^{\pm} \sum_{-\Lambda_t^{-} < t^{-} < -1} |t^{-}| - 1 \\
&= \Lambda_n^{\pm}\Lambda_t^{+}(\Lambda_t^{+} - 1) + 2\Lambda_n^{\pm}\Lambda_t^{-}\Lambda_t^{+} + 2\Lambda_n^{\pm} \sum_{(-\Lambda_t^{-} - 1) > t > 1} t \\
&= \Lambda_n^{\pm}\Lambda_t^{+}(\Lambda_t^{+} - 1) + 2\Lambda_n^{\pm}\Lambda_t^{-}\Lambda_t^{+} + 2\Lambda_n^{\pm}(\Lambda_t^{-} - 1)\frac{(\Lambda_t^{-} - 1) + 1}{2} \\
&= \Lambda_n^{\pm}\Lambda_t^{+}(\Lambda_t^{+} - 1) + 2\Lambda_n^{\pm}\Lambda_t^{-}\Lambda_t^{+} + \Lambda_n^{\pm}\Lambda_t^{-}(\Lambda_t^{-} - 1)
\end{aligned}$$

IX

$$\begin{aligned}
\Delta(M_{IX}) &= 2[0, t^{+} \mid x^{\pm}, s^{+} < t^{+}] + 2[0, t^{-} \mid x^{\pm}, s^{-} > t^{-}] + 2[x^{\pm}, t^{-} \mid x^{\pm}, s^{+}] \\
\#(\Delta(M_{IX})) &= 2\Lambda_n^{\pm} \sum_{1 \leqslant t \leqslant \Lambda_t^{+}} (t - 1) + 2\Lambda_n^{\pm} \sum_{-\Lambda_t^{-} \leqslant t \leqslant -1} (|t| - 1) + 2\Lambda_n^{\pm}\Lambda_t^{-}\Lambda_t^{+} \\
&= 2\Lambda_n^{\pm}\left(\Lambda_t^{+}\frac{(\Lambda_t^{+} - 1) + 0}{2}\right) + 2\Lambda_n^{\pm}\left(\Lambda_t^{-}\frac{(\Lambda_t^{-} - 1) + 0}{2}\right) + 2\Lambda_n^{\pm}\Lambda_t^{-}\Lambda_t^{+} \\
&= \Lambda_n^{\pm}\Lambda_t^{+}(\Lambda_t^{+} - 1) + \Lambda_n^{\pm}\Lambda_t^{-}(\Lambda_t^{-} - 1) + 2\Lambda_n^{\pm}\Lambda_t^{-}\Lambda_t^{+}
\end{aligned}$$

X By symmetry with VIII

$$\Delta(M_X) = [x^{\pm}, t^{-} \mid x^{\pm}, s^{-}] + 2[x^{\pm}, t^{+} \mid 0, s^{-}] + 2[x^{\pm}, t^{+} \mid 0, s^{+}(< t^{+})]$$
$$\#(\Delta(M_X)) = \Lambda_n^{\pm}\Lambda_t^{-}(\Lambda_t^{-} - 1) + 2\Lambda_n^{\pm}\Lambda_t^{+}\Lambda_t^{-} + \Lambda_n^{\pm}\Lambda_t^{+}(\Lambda_t^{+} - 1)$$

XI

$$\Delta(M_{XI}) = [x^{\pm}, t^{+} \mid y^{\pm}, t^{+}] + 2[x^{\pm}, 0 \mid y^{\pm}, t^{-}]$$
$$\#(\Delta(M_{XI})) = \Lambda_n^{\pm}(\Lambda_n^{\pm} - 1)\Lambda_t^{+} + 2\Lambda_n^{\pm}(\Lambda_n^{\pm} - 1)\Lambda_t^{-}$$

XII

$$\Delta(M_{XII}) = [x^{\pm}, t^{-} \mid y^{\pm}, t^{-}] + 2[x^{\pm}, 0 \mid y^{\pm}, t^{+}]$$
$$\#(\Delta(M_{XII})) = \Lambda_n^{\pm}(\Lambda_n^{\pm} - 1)\Lambda_t^{-} + 2\Lambda_n^{\pm}(\Lambda_n^{\pm} - 1)\Lambda_t^{+}$$

XIII

$$\Delta(M_{XIII}) = [x^{\pm}, t^{\pm} \mid y^{\pm}, s^{\pm}]$$
$$\#(\Delta(M_{XIII})) = \Lambda_n^{\pm}\Lambda_t^{\pm}(\Lambda_n^{\pm} - 1)(\Lambda_t^{\pm} - 1)$$

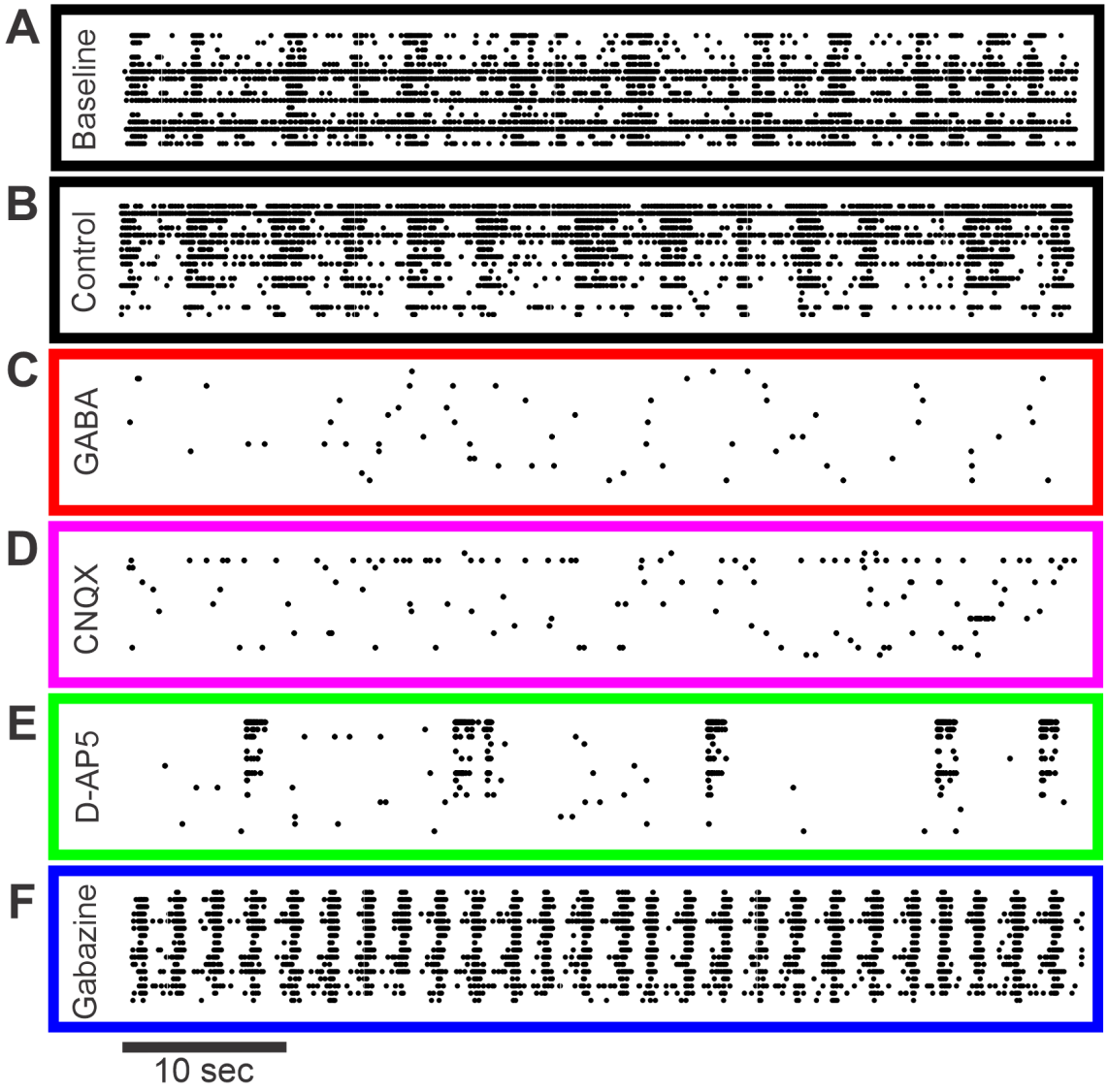


Figure 2.6: **Representative 1-minute epochs of rat cortical spike rasters per each experimental condition at 22 days *in vitro* (Kapucu et al, 2022).** Rat cortical cultures from microelectrode array well plates were exposed to the following pharmacological agents: baseline (A; black), control (B; black), GABA (C; red), CNQX (D; magenta), D-AP5 (E; green), and gabazine (F; blue). Note that we present the raster in two dimensions here while in our analysis for Fig. 2.5 we used the two spatial dimensions of the MEA and one temporal dimension.

Chapter 3

Detecting heterogeneous seizures in newborn infants using triple correlation

3.1 Introduction

Newborn infants have the highest risk of seizures among all age groups, seizures that both often have severe consequences and cannot always be visually detected (Glass et al, 2016; Padiyar et al, 2020): only EEG monitoring can detect a majority of these seizures (Glass et al, 2013). Unfortunately, the gold standard, continuous video EEG (cEEG) monitoring, is highly resource intensive, requiring specialized clinicians to evaluate high-frequency data around the clock (Kubota et al, 2018). To improve efficiency, clinicians often use transformed and time-compressed EEG signals to improve the visual salience of seizures, which allows them to read through recordings more quickly (Glass et al, 2013). However, even with more efficient review, ongoing real-time monitoring is still unrealistic for the vast majority of neonatal care. While catching the seizure in review at the end of the day is helpful, the

This chapter is pending submission as a manuscript of the same name, co-authored with Julia Henry and Wim van Drongelen.

quicker the detection the better the outcomes (Gotman, 1985; Gotman et al, 1997). Ultimately, we would like a detector that could alert clinicians in real-time to the possibility of a seizure.

Such a detector might naturally be constructed using the same transformation that enables clinicians to more efficiently review recordings: an improvement in visual salience should correspond to an improvement in numerical distinguishability. In neonatal intensive care units (NICUs), the standard transformed signal is amplitude-integrated EEG (aEEG) (Glass et al, 2013). An aEEG signal includes two traces subsampling a rectified and low-pass-filtered EEG to include only two points in a large time window (e.g. 15 seconds): one at an upper percentile, one at a lower percentile (e.g. 95th and 9th) (Zhang and Ding, 2013; Rakshasbhuvankar et al, 2015; Chen et al, 2019). Since this subsamples a typical 256Hz EEG over 2000x, with aEEG clinicians can quickly review several hours of recordings to identify time windows with trends suspicious for seizure activity. Additionally, neonatologists lacking the specialized training required to read cEEG can parse aEEG bedside Glass et al (2013). One particular trend indicative of seizure is an increase in the lower trace, which is a trend in theory easily detectable by simple thresholding (Glass et al, 2013).

In this paper, we implement a simple detector with a signal derived from triple correlation. We have shown elsewhere that triple correlation completely characterizes any neural recording (chapter 2), so it should reflect any difference between seizure and non-seizure epochs. We show that a simple, theory-inspired summary of triple correlation leads to detection on par with the aEEG lower margin. Moreover, unlike the aEEG's single axis of change, triple correlation reflects the heterogeneity of neonatal seizures.

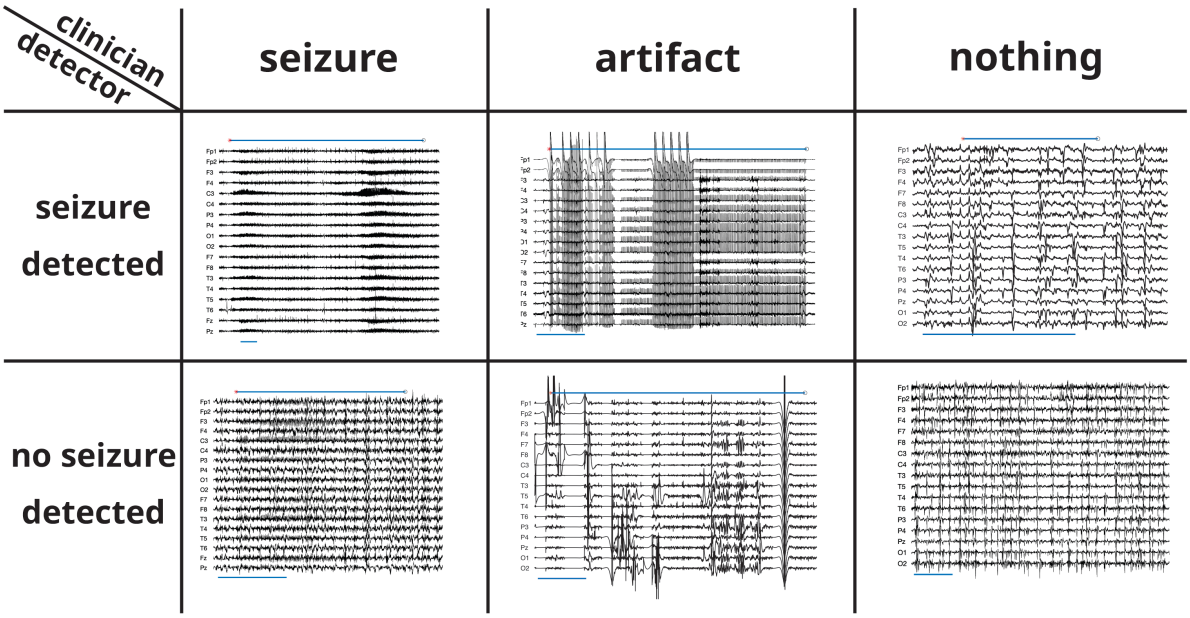


Figure 3.1: **Example pre-processed EEG snippets in confusion matrix with artifacts.** Snippets in the top row had salient TriCorr signals that were detected as seizures by most thresholds (see Fig. 3.3 for signal traces). Snippets in the bottom row did not have salient TriCorr signals. First-column snippets include a reviewer-consensus seizure (marked by upper bar). Middle-column snippets include a neurologist-annotated artifact (marked by upper bar). Third-column snippets include no expert-defined event, confirmed *post-hoc* by our neurologist (JH). The upper bar in the upper-right snippet indicates the erroneous detection. Green outlines indicate the detector matched the reviewer consensus, red indicates detector error. The blue bar below each recording indicates 60s.

3.2 Methods

3.2.1 Open-source EEG recordings

We used neonatal EEG recordings from an open-source dataset (Stevenson et al, 2019). These clinical recordings were sampled at 256 Hz and include evaluations by three expert reviewers for seizures. There were 79 patients in the dataset, with 343 seizures (defined by reviewer consensus) across 39 patients (seizing patient mean seizure count was 8.8 with std 11.9). The total duration of the recordings was 402,825s, of which 39,259s was marked as seizing (seizing patient mean seizure du-

ration was 1,006s with std 1,392s).

We ran our analyses both on the entire dataset, and on a subset with artifacts manually removed. Because of the time-consuming nature of expert artifact annotation, we only evaluated a subset of the recordings for artifacts. To choose that subset, we visually inspected all patients' reviewer consensus plots (see Fig. 3.2) and selected eight patients as having both 1) good reviewer consensus, as indicated by a lack of epochs with only one or two of the three reviewers indicating a seizure, and 2) a good balance of seizure and non-seizure epochs. We then additionally took the first fifteen recordings by the dataset's numbering, so that we evaluated a mix of "clean" and randomly-chosen recordings. A trained neurologist (JH) then annotated artifacts within these recordings. We discarded one patient (patient 21) from this subset because the vast majority of the recording was artifactual (92%).

There were 22 patients in the artifact-annotated dataset, with 134 seizures (defined by reviewer consensus) across 16 patients (artifact-annotated seizing patient mean seizure count was 8.4 with std 11.9). Total duration of the artifact-annotated recordings was 121,533s, of which 2,898s were marked as seizing (artifact-annotated seizing patient mean seizure duration was 929s with std 842s).

We excluded the respiratory and ECG channels from all analyses because they are not EEG signals. We additionally excluded Fz and Cz from all analyses due to channel-specific whole-recording artifacts in many patients.

3.2.2 Code

The below analyses were conducted with Julia v1.7.3 (Bezanson et al, 2017). All code is available on Github at grahamas/NeonateTriCorr. Filtering used the package JuliaDSP/DSP.jl. Statistical tests used JuliaStats/HypothesisTests.jl. ROC analysis used davidavdav/ROCAnalysis.jl.

3.2.3 EEG preprocessing

Prior to analysis, we preprocessed the EEG recordings. We first de-meaned per-channel; then bandstop filtered between 45 and 55 Hz with a sixth-order Butterworth (to account for European mains frequency of 50 Hz); then bandpass filtered between 0.1 Hz and 70 Hz with a second-order Butterworth; and finally divided by standard deviation, per-channel.

3.2.4 Amplitude-integrated EEG (aEEG)

To calculate the aEEG, we followed the example of Zhang and Ding (2013). Starting from the pre-processed EEG, we bandpass filtered between 2 and 20Hz using an FIR with a 50th order Hamming window, diverging from Zhang and Ding (2013) but in line with Werther et al (2017); Quigg and Leiner (2009). Next we computed the envelope by lowpassing below 0.32 Hz (corresponding to an RC circuit with a 500ms time constant, as in Zhang and Ding (2013)) with a 5th order Butterworth. Then, we divided the recording into sequential segments of length 15 seconds. Within each of these segments, we calculated the "margins": the value at the 9th and 93rd percentiles, the lower and upper margins respectively. These margins constitute the aEEG trace. To construct our detector, we used only the lower margin, since a common feature of seizures is an increase in the lower margin (Glass et al, 2013). In contrast to prior work, we opted to include all channels rather than pick a single channel.

3.2.5 Triple correlation over time

We calculated the triple correlation as in chapter 2, in one second snippets with lags in the ranges of -8:8 spatial and -25:25 temporal bins, using periodic spatial

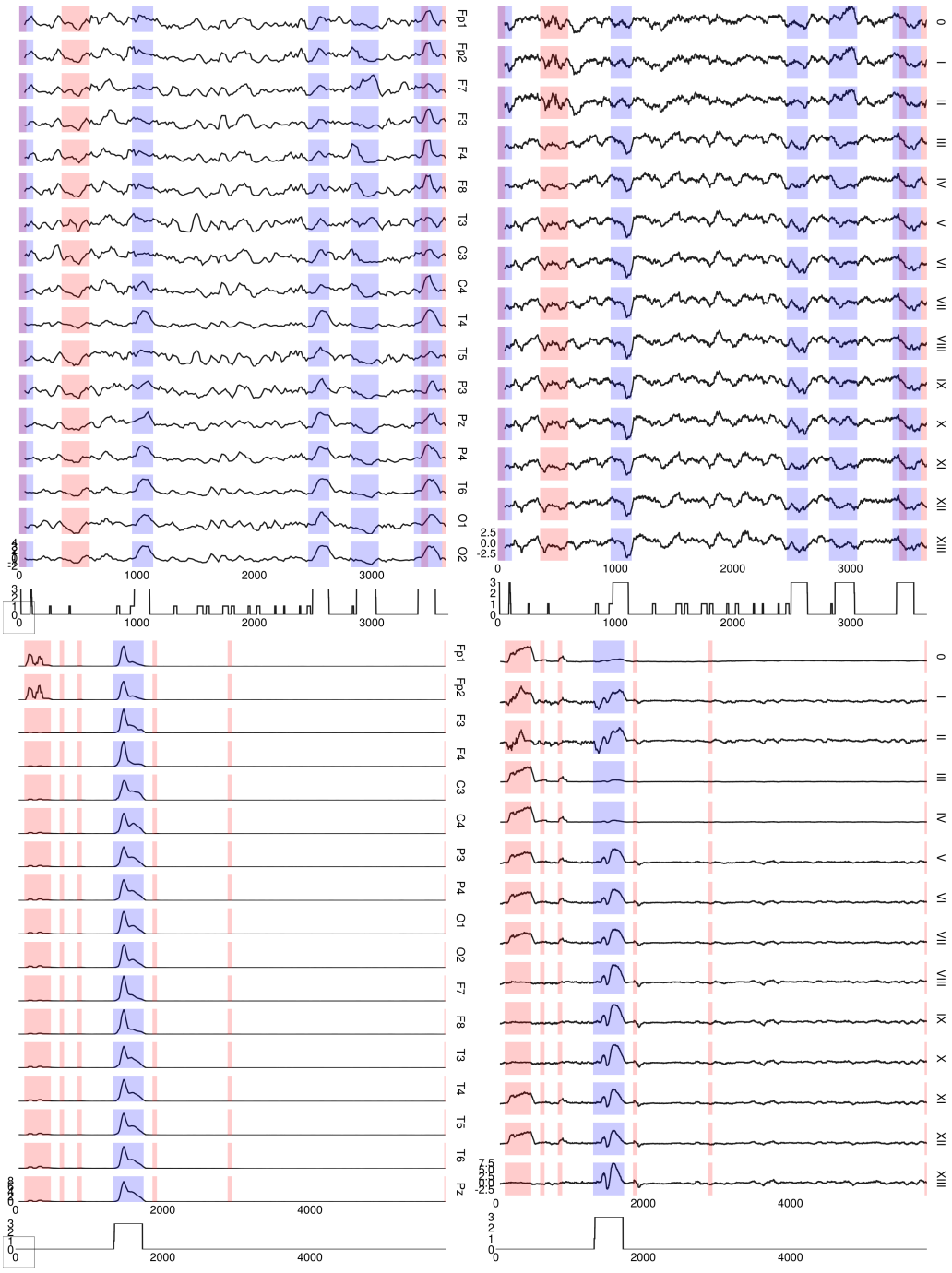


Figure 3.2: **TriCorr and aEEG signals used in simple detector. left** aEEG lower margins (9th percentile) for each EEG channel, for patients 7 (top) and 62 (bottom) **right** TriCorr contributions for each motif class within sequential 1s EEG snippets. The uncolored trace shows the number of reviewers indicating a seizure occured in each second, out of a possible three reviewers. Blue shading indicates reviewer consensus, plus one minute on either side. Red shading indicates minutes including seconds that were artifact-annotated by our expert reviewer

boundary conditions and data-padded temporal boundary conditions. With 17 EEG channels, a lag of 8 spatial bins encompasses the entire spatial domain. With a sampling frequency of 256 Hz, the temporal lags encompass approximately -98:98ms. We summarized the triple correlation by its fourteen motif classes, which partition the space of triplet motifs into qualitatively distinct classes embodying different computational patterns, e.g. synchrony, feedback, divergence, etc.

3.2.6 Distribution comparisons

We used an approximate two-sample Kolmogorov-Smirnov (KS) test to compare seizure and non-seizure snippets' signals (TriCorr or aEEG), excluding non-seizure snippets within 60 seconds of seizure onset or offset. The KS test tests the null hypothesis that two samples could have been drawn from the same unknown distribution. Thus, significant p-values indicate that the seizure and non-seizure signals likely had differing statistical distributions.

3.2.7 Detector construction

For both aEEG and triple correlation (TriCorr) we obtained multi-channel signals: in the case of aEEG, these are the seventeen EEG channels, and in the case of TriCorr, these are the fourteen motif classes. Within these channels, we z-scored either within or across patients: we subtracted either the across-patient or within-patient channel mean and divided by the all-patient or within-patient standard deviation. Then we computed the rolling temporal mean in 60s windows. These channels are depicted in Fig. 3.2.

In the case of TriCorr, we then took the absolute value of these signals, since seizures could be reflected by either positive or negative deviations from the mean.

In contrast, aEEG lower margin only reflects seizures through increases. We created a single timeseries by taking the maximum value across all channels.

To account for complications due to differing resolutions of TriCorr and aEEG (1s vs 15s snippets, respectively), we applied our detector per-minute: we divided the signals into 60s intervals and regarded any suprathreshold signal within an interval as a detection for the whole interval, and otherwise marked no detection for the interval. An interval was regarded as a seizure interval if all three reviewers marked a seizure within one minute of the start or end of the interval.

We depict examples of both aEEG and TriCorr detector signals in Figure 3.3.

3.2.8 Detector metrics

To evaluate the efficacy of our detectors, we constructed ROC curves using davi-davdav/ROCAnalysis.jl. These ROC curves are depicted in Fig. 3.4 with clinically-relevant versions of the true positive rate (TPR) and false positive rate (FPR). Rather than a per-interval TPR, we reported a per-seizure TPR, which indicates the proportion of seizures correctly detected. For FPR, we reported the false positives per hour (FP/Hr), which is simply the FPR multiplied by the number of non-seizure hours. Note that AUCs are reported for the identically-shaped ROC but with a per-interval FPR axis rather than FP/Hr, so that the AUC is between 0 and 1.

We calculated the clinician FP/Hour rate by marking detections when *any* reviewer indicated a seizure and comparing this to the reviewer-consensus "ground truth."

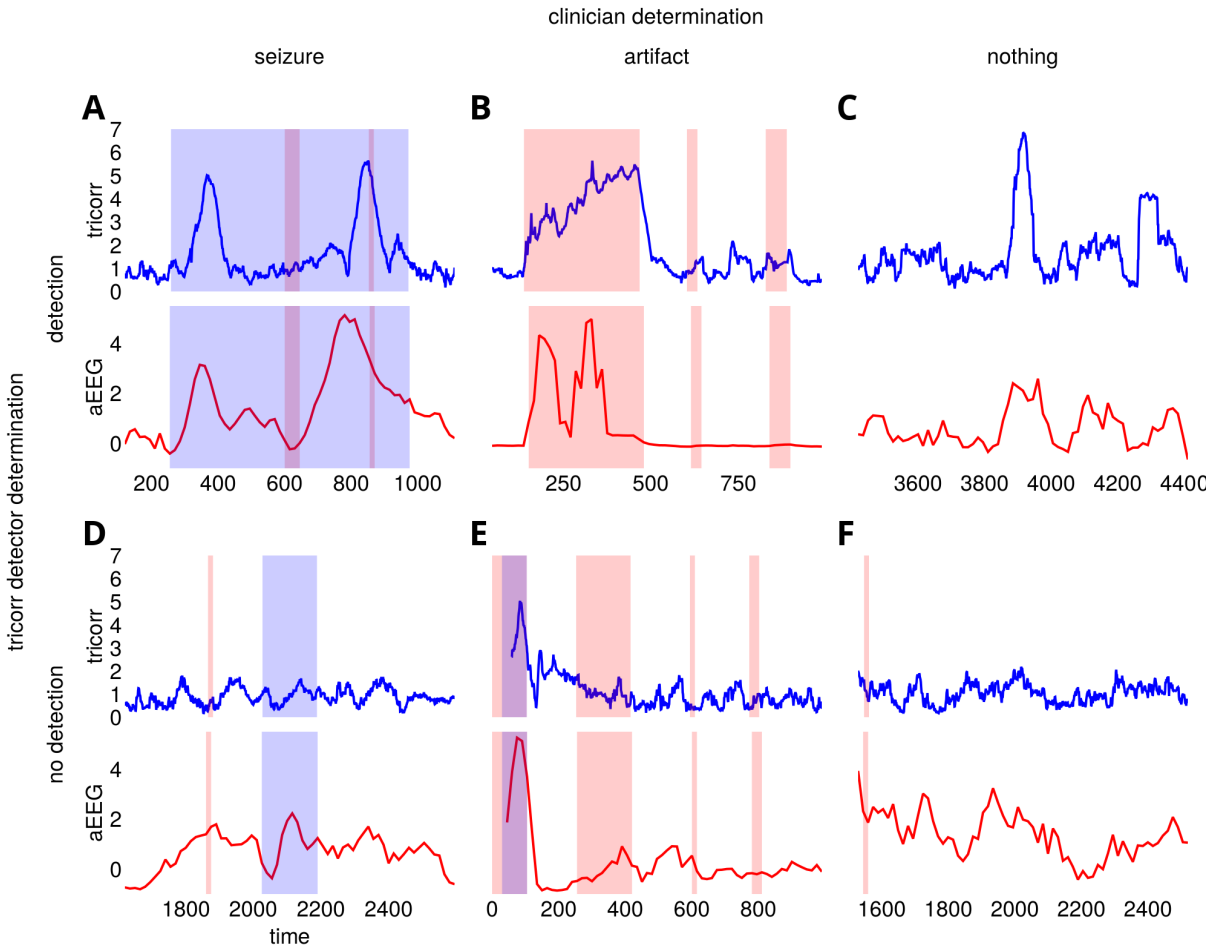


Figure 3.3: **Example detector traces in confusion matrix with artifacts.** Each trace is a 1000s long snippet including the example indicated by the row and column labels. These 1000s snippets encompass the same examples as in Fig. 3.1. We did not use a particular threshold to choose these examples. **A** true positives: a long seizure (blue highlight) with two high-amplitude detections by both TriCorr (blue line) and aEEG (red line). **B** artifact-induced false positive: three artifacts (red highlight), with the first producing salient signals in both TriCorr and aEEG. **C** a false positive in the TriCorr signal, though not so definitively in the aEEG. **D** false negative: a seizure not detected by either signal. **E** true negative (correct artifact rejection): the latter three artifacts are not marked as detections by either TriCorr or aEEG. However, the first artifact encompasses a seizure that is correctly detected by both TriCorr and aEEG, and so does not fall properly into any of these six categories. **F** a true negative with only a brief correctly rejected artifact.

3.3 Results

3.3.1 Triple correlation detects individual seizures comparably well to clinical standard

A signal based on triple correlation leads to similar per-seizure detection performance compared to one based on aEEG, the current clinical standard.

Neither is obviously better. The ROC curves are quite similar (Fig. 3.4), with the only substantial deviation being in the case where we z-score the signals across all patients for detecting individual seizures (Fig. 3.4B). In that case, the TriCorr detector has a higher AUC than the aEEG. This suggests TriCorr could be a better naive detector, i.e. that it would require less tailoring to the individual patient's signal statistics.

Comparing traces in individual patients, in some patients the aEEG signal is cleaner, in others the triple correlation signal. Nor is it clear that the "ground truth" is any better: trained clinicians routinely disagree with each other. In our data, the trained clinicians had a false positive rate of 6 FP/Hr when treating single reviewers as the detection signal and reviewer consensus as the ground truth.

3.3.2 Triple correlation can detect seizing patients comparably well to expert reviewers

For per-patient detection, at perfect sensitivity the TriCorr detector can have the same specificity as the expert reviewers (Fig. 3.4C; Table 3.1), in the most ideal case where we have an estimate for the distributions of motif-class contributions for each patient. Given that this level of uncertainty is inherent in our "ground truth," this performance is the best performance we could expect. In comparison, even

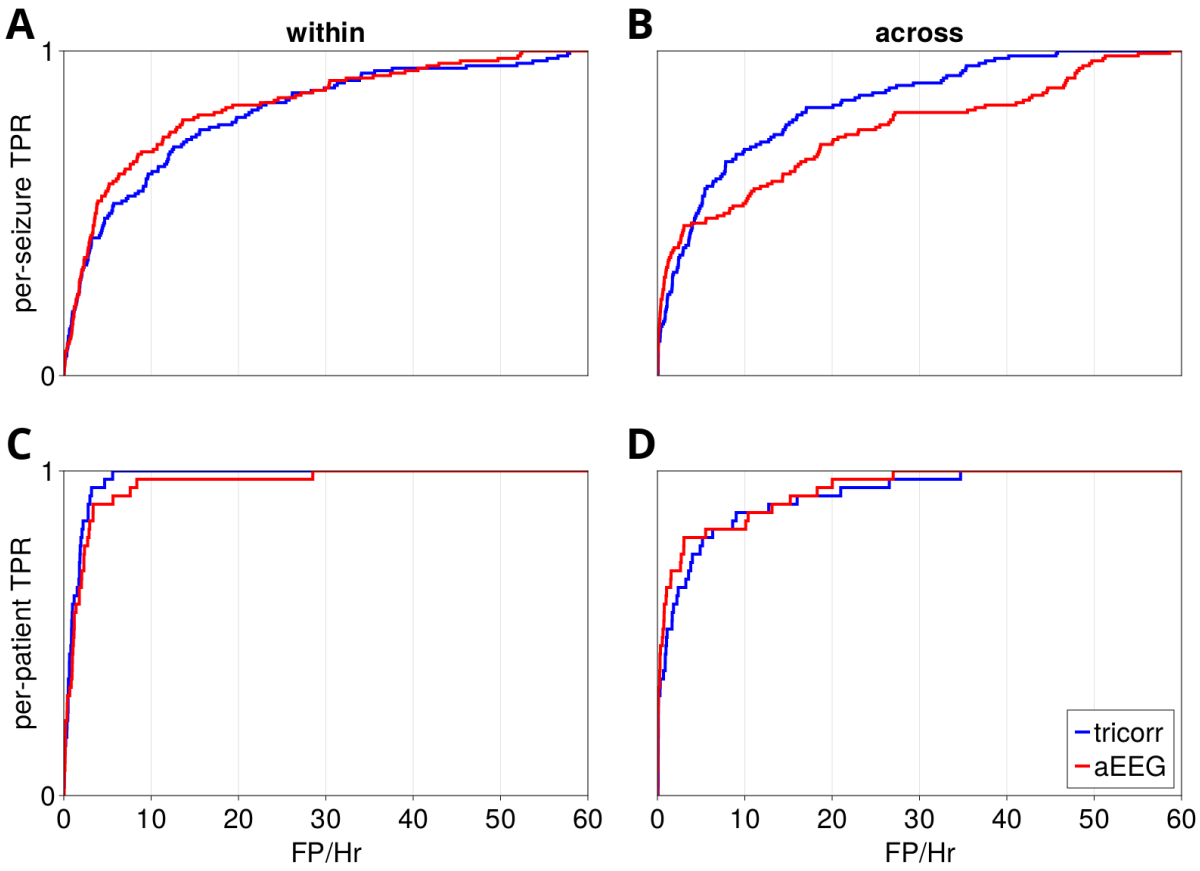


Figure 3.4: **Detector ROC curves.** These curves compare the success of detectors when defining success per-seizure, *top row*, or per-patient, *bottom row*, and when z-scoring the detector signals across all patients, *left*, or within each patient, *right*. In all cases the failure rate is defined as false positive minutes per hour.

Target	Standardization	Signal	AUC	TPR (5.7 FP/Hr)	FP/Hr (80% TPR)
patient	across	aEEG	0.94	0.82	5.5
patient	across	tricorr	0.92	0.79	6.3
patient	within	aEEG	0.96	0.92	2.9
patient	within	tricorr	0.98	1.0	2.1
seizure	across	aEEG	0.76	0.48	27.0
seizure	across	tricorr	0.84	0.58	16.0
seizure	within	aEEG	0.83	0.6	15.0
seizure	within	tricorr	0.81	0.52	21.0

Table 3.1: **Statistics for the ROC curves shown in Figure 3.4.** Target indicates whether the positives were per-patient or per-seizure; standardization indicates whether the signals were z-scored across or within patients. Signal indicates aEEG vs TriCorr. AUC is area under the ROC curve. TPR indicates the true positive rate (either per-patient or per-seizure) at FP/Hr = 5.7, which was chosen because that is the false positive rate of single-reviewer annotation. FP/Hr indicates the rate of false positives for a detector with 80% sensitivity.

in this ideal case aEEG does not achieve perfect sensitivity without more than five times the reviewers' FP/hour. The aEEG's sensitivity at the same 5.7 FP/Hr specificity is 92% (see Table 3.1).

3.3.3 Triple correlation distinguishes heterogeneity

Triple correlation reveals structural variations between patients' seizure responses, where aEEG only offers a monolithic signal. Our analysis reflects that neonatal seizures can involve either an increase or a decrease in triple correlation, and usually this change is consistent within a patient, i.e. each patient has a particular type of seizure. Our algorithm then reduces this complexity to ask if signals indicating any of these types of seizures are present. However, heterogeneity in seizure signal is visually apparent before this reduction Fig. 3.2.

Of the 39 patients with seizures, 13 had significant distribution differences with increases in triple correlation ictal vs interictally, 11 had decreases, 3 had heterogeneous effects among motif classes, and 11 had no significant difference in mo-

tif class contributions between ictal and interictal snippets, as determined by a Kolmogorov-Smirnov test.

3.4 Discussion

Triple correlation provides a rich field of potential signals. Here we used only the simplest and even so matched and perhaps exceeded aEEG, a signal transformation in wide clinical use. Even better, our approach distinguishes among the heterogeneous neonatal seizures, even here where we have reduced TriCorr to its grossest components (motif classes). In its fullness, triple correlation provides a plethora of building blocks with which to construct signals with the potential to differentiate between seizures. Perhaps we will one day be able to relate these differences to known differences in etiology and treatment.

The expressiveness of TriCorr signals could also help more immediately, at the bedside where aEEG is currently used. Commonly aEEG is used as a front-line bedside tool by neonatologists rather than neurologists because it offers a straightforward signal that doesn't require years of training the way that cEEG does. However, that simplicity comes at a tradeoff when some nuanced differences in signals have different diagnostic outcomes. For example, one critical predictor of neonatal outcomes is the EEG background, which can be read from the aEEG margins over an extended time. However, abnormal background voltage involves a decreased lower margin, which stands in contrast to seizure detection, where an increased margin indicates a seizure. This can lead long seizures to be mistakenly interpreted as normal background EEG, with concommitant incorrect treatment, and hence poor outcomes (Glass et al, 2013). With the expressiveness of TriCorr, we hypothesize seizure and normal background would be separable.

The further reaching advantage of TriCorr is that its expressiveness is of features more covert than those expressed by aEEG. aEEG makes overt the general trends of the maxima and minima of the EEG, less some noise. In contrast, TriCorr overtly represents the third-order structure. While an expert could in theory read such structure from the EEG, it would be with extensive training and experience, not unlike that of the trained epileptologist. TriCorr creates the potential to explicitly represent that structure to a more generally trained audience, such as a nurse or neonatologist, for whom noticing a seizure is but one of many tasks.

This brings us to the final caveat of this study: we do not claim this will replace the need for neurologists trained in the art of reading cEEG. We can only argue that this prepares the way for more reliable and timely alarm to bring cases to the attention of the neurologist for definitive diagnosis. We do, additionally, hope that TriCorr will augment the neurologist's reading, but this study does not begin to provide evidence for such a claim, merely the hope. Indeed, it is hard to see how such evidence could be acquired definitively, given that the ground truth must always be the expert reviewers themselves.

Chapter 4

Conclusion

Though my grander goal may be general BCIs, this dissertation addressed them only as motivation for pursuing population-scale neural activity. To conclude, I return to BCIs to contextualize the achievements of these chapters.

In the first chapter, I brought a theoretical model a step closer to reality with a minimal increase in complexity, with the particular goal of accounting for focal seizures. On a general level, any model sophisticated enough to underpin a BCI should be able to capture seizures, since they are relatively simple. I say "relatively" with full respect to the complexity of epilepsy: a seizure is simple in the way a whirlpool in an ocean is simple. But, then, a BCI is like understanding every molecule in that ocean, and then using lasers to generate ripples that write out Wikipedia in seven dimensions. To do that you need to have a really good model of flowing water, and a good enough model will include whirlpools. More specifically, though, you need to make sure your lasers don't accidentally make a whirlpool, since in this extended metaphor, whirlpools are bad. Any neural stimulation runs the risk of inducing seizures, so bidirectional BCIs will require modeling that can determine whether they will induce seizures.

In the second chapter, rather than a theoretical model, I described triple corre-

lation as a characterization of neural data that I showed is technically complete, meaning there is a one-to-one relationship between a dataset and its triple correlation. Neural activity is constantly moving in near infinitely many directions, but, for example, in EEG we only see two: up and down. Technically we can see multiple channels of ups and downs, but practically our direct perception of the signals on paper can only go so far away from "this line goes up and down." The essence of characterizing neural population activity is in figuring out which directions are important, so that up and down on paper reflects important movement in neural activity. The uniqueness of triple correlation tells us that those directions are probably important. The motif classes constitute various meaningful directions of motion, and our simple tests illustrate that our intuition about their importance is reasonably correct.

In the third chapter, with seizure detection, we have a concrete and substantial case: in a seizure, neural activity moves in very weird directions, quite strongly, so much so that those motions are often apparent even in the simple voltage up-down directions, as aEEG distills. This provides a good test for the utility of triple correlation: a good characterization should reveal those strange movements, and if we want to be able to understand, it should reveal those movements concisely. Our motif classes provide that concision, and our results show that even at such a simplified level, triple correlation reflects the changes due to seizures. Moreover, we see that the up-down movement of aEEG was masking a menagerie of directions that different patients' seizures can take. This is no surprise: neurologists have seen this with their deep insight into the vagaries of the full-resolution EEG, for all that they have to parse that insight from EEG's simple up-down. But, with training, neurologists can see spatiotemporal patterns in the up-down, patterns that we argue are made overt by triple correlation. This is the essence of what we need for BCIs:

overt representations of the essence of the brain's hidden movements, capturing complexity so that we have simple handles onto neural activity.

Thus the work in this dissertation falls into two broad categories that are necessary for the advancement of neuroscience research, and BCIs in particular. First, modeling: advancing theoretical understanding of the behavior of neural populations, particularly with scale-spanning mechanistic explanations, and even into pathological corner cases that BCIs must aim to avoid. Second, characterization: the representation of data in new directions that are both meaningful and explanatory. I hope that these contributions will provide fruitful ground for future advances in the understanding of neural population activity, and for the development of new tools that may at a minimum more effectively and efficiently detect patholological neural activity, and, in a perhaps distant future, capture neural activity with sufficient complexity to create a general BCI.

References

Abeles M (1982) *Local Cortical Circuits*, Studies of Brain Function, vol 6, 1st edn. Springer, Berlin, DOI: 10.1007/978-3-642-81708-3

Abeles M (1991) *Corticonics: Neural Circuits of the Cerebral Cortex*, 1st edn. Cambridge University Press, Cambridge, DOI: 10.1017/CBO9780511574566

Abeles M, Gerstein GL (1988) Detecting spatiotemporal firing patterns among simultaneously recorded single neurons. *J Neurophysiol* 60(3):909–924. DOI: 10.1152/jn.1988.60.3.909

Andrews JP, Ammanuel S, Kleen J, et al (2020) Early seizure spread and epilepsy surgery: a systematic review. *Epilepsia* 61(10):2163–2172. DOI: 10.1111/epi.16668

Anninos PA, Beek B, Csermely TJ, et al (1970) Dynamics of neural structures. *Journal of Theoretical Biology* 26(1):121–148. DOI: 10.1016/S0022-5193(70)80036-4

Baker SN, Lemon RN (2000) Precise spatiotemporal repeating patterns in monkey primary and supplementary motor areas occur at chance levels. *Journal of neurophysiology* 84(4):1770–1780. DOI: 10.1152/jn.2000.84.4.1770

Bartelt H, Lohmann AW, Wirnitzer B (1984) Phase and amplitude recovery from bispectra. *Applied Optics* 23(18):3121–3129. DOI: 10.1364/AO.23.003121

Beghi E, Giussani G, Nichols E, et al (2019) Global, regional, and national burden of epilepsy, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Neurology* 18(4):357–375. DOI: 10.1016/S1474-4422(18)30454-X

Benucci A, Frazor RA, Carandini M (2007) Standing waves and traveling waves distinguish two circuits in visual cortex. *Neuron* 55(1):103–117. DOI: 10.1016/j.neuron.2007.06.017

Beurle RL (1956) Properties of a mass of cells capable of regenerating pulses. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences* 240(669):55–94. DOI: 10.1098/rstb.1956.0012

Bezanson J, Karpinski S, Shah VB, et al (2012) Julia: A Fast Dynamic Language for Technical Computing. *arXiv* pp 1–27. DOI: 10.1109/APS.2015.7305314

Bezanson J, Edelman A, Karpinski S, et al (2017) Julia: A Fresh Approach to Numerical Computing. *SIAM Review* 59(1):65–98. DOI: 10.1137/141000671

Bianchi D, Marasco A, Limongiello A, et al (2012) On the mechanisms underlying the depolarization block in the spiking dynamics of CA1 pyramidal neurons. *Journal of Computational Neuroscience* 33(2):207–225. DOI: 10.1007/s10827-012-0383-y

Bojanek K, Zhu Y, MacLean J (2020) Cyclic transitions between higher order motifs underlie sustained asynchronous spiking in sparse recurrent networks. *PLOS Computational Biology* 16(9):e1007,409. DOI: 10.1371/journal.pcbi.1007409

Bou Assi E, Gagliano L, Rihana S, et al (2018) Bispectrum Features and Multilayer Perceptron Classifier to Enhance Seizure Prediction. *Sci Rep* 8(1):15,491. DOI: 10.1038/s41598-018-33969-9

Buzsáki G, Draguhn A (2004) Neuronal oscillations in cortical networks. *Science* 304(5679):1926–1929. DOI: 10.1126/science.1099745

Chen C, Sun C, Leonhardt S, et al (2019) Amplitude-Integrated Electroencephalography Applications and Algorithms in Neonates: A Systematic Review. *IEEE Access* 7:141,766–141,781. DOI: 10.1109/ACCESS.2019.2944531

Cowan JD, Neuman J, van Drongelen W (2016) Wilson–Cowan Equations for Neocortical Dynamics. *The Journal of Mathematical Neuroscience* 6(1):1. DOI: 10.1186/s13408-015-0034-5

de Curtis M, Avanzini G (2001) Interictal spikes in focal epileptogenesis. *Progress in Neurobiology* 63(5):541–567. DOI: 10.1016/S0301-0082(00)00026-5

Dechery JB, MacLean JN (2018) Functional triplet motifs underlie accurate predictions of single-trial responses in populations of tuned and untuned V1 neurons. *PLOS Computational Biology* 14(5):e1006,153. DOI: 10.1371/journal.pcbi.1006153

Diesmann M, Gewaltig MO, Aertsen A (1999) Stable propagation of synchronous spiking in cortical neural networks. *Nature* 402(6761):529–533. DOI: 10.1038/990101

van Drongelen W (2018) *Signal processing for neuroscientists: Introduction to the analysis of physiological signals.*, 2nd edn. Elsevier/Academic Press

Ermentrout GB, Cowan JD (1979a) A mathematical theory of visual hallucination patterns. *Biological Cybernetics* 34:137–150. DOI: 10.1007/BF00336965

Ermentrout GB, Cowan JD (1979b) Temporal oscillations in neuronal nets. *Journal of Mathematical Biology* 7:265–280. DOI: 10.1007/BF00275728

Ermentrout GB, Cowan JD (1980) Large scale spatially organized activity in neural nets. *SIAM Journal on Applied Mathematics* 38:1–21. DOI: 10.1137/0138001

Feller W (1950) *An introduction to probability theory and its applications*, 1st edn. Wiley, Oxford

Gagliano L, Bou Assi E, Nguyen DK, et al (2019) Bispectrum and Recurrent Neural Networks: Improved Classification of Interictal and Preictal States. *Sci Rep* 9(1):15,649. DOI: 10.1038/s41598-019-52152-2

Gibbs FA, Gibbs EL (1952) *Atlas of electroencephalography / Vol. 2, Epilepsy*. Addison-Wesley, Reading, MA [etc.]

Glass HC, Wusthoff CJ, Shellhaas RA (2013) Amplitude-Integrated Electroencephalography: The Child Neurologist's Perspective. *Journal of Child Neurology* 28(10):1342–1350. DOI: 10.1177/0883073813488663

Glass HC, Shellhaas RA, Wusthoff CJ, et al (2016) Contemporary Profile of Seizures in Neonates: A Prospective Cohort Study. *The Journal of Pediatrics* 174:98–103.e1. DOI: 10.1016/j.jpeds.2016.03.035

Gotman J (1982) Automatic recognition of epileptic seizures in the EEG. *Electroencephalography and Clinical Neurophysiology* 54(5):530–540. DOI: 10.1016/0013-4694(82)90038-4

Gotman J (1985) Seizure recognition and analysis. *Electroencephalography and clinical neurophysiology Supplement* 37:133–145

Gotman J, Gloor P (1976) Automatic recognition and quantification of interictal epileptic activity in the human scalp EEG. *Electroencephalography and Clinical Neurophysiology* 41(5):513–529. DOI: 10.1016/0013-4694(76)90063-8

Gotman J, Flanagan D, Zhang J, et al (1997) Automatic seizure detection in the newborn: methods and initial evaluation. *Electroencephalography and Clinical Neurophysiology* 103(3):356–362. DOI: 10.1016/S0013-4694(97)00003-9

Grün S (2009) Data-Driven Significance Estimation for Precise Spike Correlation. *Journal of Neurophysiology* 101(3):1126–1140. DOI: 10.1152/jn.00093.2008

Harris JD, Ermentrout B (2018) Traveling waves in a spatially-distributed Wilson–Cowan model of cortex: From fronts to pulses. *Physica D: Nonlinear Phenomena* 369:30–46. DOI: 10.1016/j.physd.2017.12.011

Harth EM, Csermely TJ, Beek B, et al (1970) Brain functions and neural dynamics. *Journal of Theoretical Biology* 26(1):93–120. DOI: 10.1016/S0022-5193(70)80035-2

Hebb DO (1949) *The organization of behavior; a neuropsychological theory*, 1st edn. Wiley, Oxford, England

Hertz J, Prügel-Bennett A (1996) Learning synfire chains: turning noise into signal. *Int J Neural Syst* 7(4):445–450. DOI: 10.1142/s0129065796000427

Hyvärinen T, Hyysalo A, Kapucu FE, et al (2019) Functional characterization of human pluripotent stem cell-derived cortical networks differentiated on laminin-521 substrate: comparison to rat cortical cultures. *Sci Rep* 9(1):17,125. DOI: 10.1038/s41598-019-53647-8

Ikegaya Y, Aaron G, Cossart R, et al (2004) Synfire chains and cortical songs: temporal modules of cortical activity. *Science* 304(5670):559–564. DOI: 10.1126/science.1093173

Iorio F, Saez-Rodriguez J, Bernardo Dd (2013) Network based elucidation of drug response: from modulators to targets. *BMC Syst Biol* 7:139. DOI: 10.1186/1752-0509-7-139

Jansen NA, Perez C, Schenke M, et al (2021) Impaired β-γ Coupling Indicates Inhibitory Dysfunction and Seizure Risk in a Dravet Syndrome Mouse Model. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 41(3):524–537. DOI: 10.1523/JNEUROSCI.2132-20.2020

Jehi L (2018) The Epileptogenic Zone: Concept and Definition. *Epilepsy Currents* 18(1):12–16. DOI: 10.5698/1535-7597.18.1.12

Jiang M, Gao X, An H, et al (2017) Reconstructing complex network for characterizing the time-varying causality evolution behavior of multivariate time series. *Scientific Reports* 7(1):10,486. DOI: 10.1038/s41598-017-10759-3

Jirsa VK, Stacey WC, Quilichini PP, et al (2014) On the nature of seizure dynamics. *Brain* 137(8):2210–2230. DOI: 10.1093/brain/awu133

Jovanović S, Rotter S (2016) Interplay between Graph Topology and Correlations of Third Order in Spiking Neuronal Networks. *PLOS Computational Biology* 12(6):e1004,963. DOI: 10.1371/journal.pcbi.1004963

Julesz B (1962) Visual Pattern Discrimination. *IRE Transactions on Information Theory* 8(2):84–92. DOI: 10.1109/TIT.1962.1057698

Julesz B, Gilbert EN, Shepp LA, et al (1973) Inability of humans to discriminate between visual textures that agree in second-order statistics-revisited. *Perception* 2(4):391–405. DOI: 10.1068/p020391

Kapucu FE, Vinogradov A, Hyvärinen T, et al (2022) Comparative microelectrode array data of the functional development of hPSC-derived and rat neuronal networks. *Scientific Data* 9(1):120. DOI: 10.1038/s41597-022-01242-4

Kellaway P, Crawley JW, Kagawa N (1959) A specific electroencephalographic correlate of convulsive equivalent disorders in children. *The Journal of Pediatrics* 55(5):582–592. DOI: 10.1016/S0022-3476(59)80242-0

Kim CM, Nykamp DQ (2014) Dynamics of a network of excitatory and inhibitory neurons induced by depolarization block. *BMC Neuroscience* 15(Suppl 1):P76. DOI: 10.1186/1471-2202-15-S1-P76

Kim CM, Nykamp DQ (2017) The influence of depolarization block on seizure-like activity in networks of excitatory and inhibitory neurons. *Journal of Computational Neuroscience* 43(1):65–79. DOI: 10.1007/s10827-017-0647-7

Kreuz T, Chicharro D, Greschner M, et al (2011) Time-resolved and time-scale adaptive measures of spike train synchrony. *J Neurosci Methods* 195(1):92–106. DOI: 10.1016/j.jneumeth.2010.11.020

Kreuz T, Chicharro D, Houghton C, et al (2013) Monitoring spike train synchrony. *Journal of Neurophysiology* 109(5):1457–1472. DOI: 10.1152/jn.00873.2012

Kubota Y, Nakamoto H, Egawa S, et al (2018) Continuous EEG monitoring in ICU. *Journal of Intensive Care* 6(1):39. DOI: 10.1186/s40560-018-0310-z

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436–444. DOI: 10.1038/nature14539

Liou Jy, Smith EH, Bateman LM, et al (2020) A model for focal seizure onset, propagation, evolution, and progression. *eLife* 9:e50,927. DOI: 10.7554/eLife.50927

Lu Y, Sato Y, Amari Si (2011) Traveling Bumps and Their Collisions in a Two-Dimensional Neural Field. *Neural Computation* 23(5):1248–1260. DOI: 10.1162/NECO_a_00111

Lüders HO, Najm I, Nair D, et al (2006) The epileptogenic zone: general principles. *Epileptic Disorders: International Epilepsy Journal with Videotape* 8 Suppl 2:S1–9

McCulloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5(4):115–133. DOI: 10.1007/BF02478259

Meijer HGE, Eissa TL, Kiewiet B, et al (2015) Modeling Focal Epileptic Activity in the Wilson–Cowan Model with Depolarization Block. *Journal of Mathematical Neuroscience* 5. DOI: 10.1186/s13408-015-0019-4

Milo R, Shen-Orr S, Itzkovitz S, et al (2002) Network Motifs: Simple Building Blocks of Complex Networks. *Science* 298(5594):824–827. DOI: 10.1126/science.298.5594.824

Mogensen PK, Riseth AN (2018) Optim: A mathematical optimization package for Julia. *Journal of Open Source Software* 3(24):615. DOI: 10.21105/joss.00615

Moszner Z (1980) Sur l'équation $f(x)f(y)f(-x-y) = g(x)g(y)g(-x-y)$. *aequationes mathematicae* 21(1):20–32. DOI: 10.1007/BF02189336

Muller L, Reynaud A, Chavane F, et al (2014) The stimulus-evoked population response in visual cortex of awake monkey is a propagating wave. *Nature Communications* 5(1):3675. DOI: 10.1038/ncomms4675

Muller L, Piantoni G, Koller D, et al (2016) Rotating waves during human sleep spindles organize global patterns of activity that repeat precisely through the night. *eLife* 5:e17,267. DOI: 10.7554/eLife.17267

Nauhaus I, Busse L, Carandini M, et al (2009) Stimulus contrast modulates functional connectivity in visual cortex. *Nature Neuroscience* 12(1):70–76. DOI: 10.1038/nn.2232

Nauhaus I, Busse L, Ringach DL, et al (2012) Robustness of Traveling Waves in Ongoing Activity of Visual Cortex. *Journal of Neuroscience* 32(9):3088–3094. DOI: 10.1523/JNEUROSCI.5827-11.2012

Padiyar S, Nusairat L, Kadri A, et al (2020) Neonatal seizures in the U.S. National Inpatient Population: Prevalence and outcomes. *Pediatrics & Neonatology* 61(3):300–305. DOI: 10.1016/j.pedneo.2019.12.006

Pao YH (1989) *Adaptive pattern recognition and neural networks*. Addison-Wesley Longman Publishing Co., Inc., USA

Quigg M, Leiner D (2009) Engineering Aspects of the Quantified Amplitude-Integrated Electroencephalogram in Neonatal Cerebral Monitoring. *Journal of Clinical Neurophysiology* 26(3):145–149. DOI: 10.1097/WNP.0b013e3181a18711

Rakshasbhuvankar A, Paul S, Nagarajan L, et al (2015) Amplitude-integrated EEG for detection of neonatal seizures: a systematic review. *Seizure - European Journal of Epilepsy* 33:90–98. DOI: 10.1016/j.seizure.2015.09.014

Rapela J (2018) Traveling waves appear and disappear in unison with produced speech. *arXiv:180609559 [q-bio]*

Riehle A, Grün S, Diesmann M, et al (1997) Spike synchronization and rate modulation differentially involved in motor cortical function. *Science* 278(5345):1950–1953. DOI: 10.1126/science.278.5345.1950

Roscoe DD, Hamm TM, Reinking RM, et al (1985) Detection of synchrony in the discharge of a population of neurons. II. Implementation and sensitivity of a synchronization index. *J Neurosci Methods* 13(1):51–64. DOI: 10.1016/0165-0270(85)90043-3

Rosenblatt F (1958) The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review* 65:386–408. DOI: 10.1037/h0042519

Russo E, Durstewitz D (2017) Cell assemblies at multiple time scales with arbitrary lag constellations. *eLife* 6:e19,428. DOI: 10.7554/eLife.19428

Saggio ML, Crisp D, Scott JM, et al (2020) A taxonomy of seizure dynamotypes. *eLife* 9:e55,632. DOI: 10.7554/eLife.55632

Sanzeni A, Akitake B, Goldbach HC, et al (2020) Inhibition stabilization is a widespread property of cortical networks. *eLife* 9:e54,875. DOI: 10.7554/eLife.54875

Saper CB, Fuller PM, Pedersen NP, et al (2010) Sleep State Switching. *Neuron* 68(6):1023–1042. DOI: 10.1016/j.neuron.2010.11.032

Sato TK, Nauhaus I, Carandini M (2012) Traveling Waves in Visual Cortex. *Neuron* 75(2):218–229. DOI: 10.1016/j.neuron.2012.06.029

Schevon CA, Weiss SA, McKhann G, et al (2012) Evidence of an inhibitory restraint of seizure activity in humans. *Nature Communications* 3:1060. DOI: 10.1038/ncomms2056

Schneidman E, Berry MJ, Segev R, et al (2006) Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* 440(7087):1007–1012. DOI: 10.1038/nature04701

Shimbel A, Rapoport A (1948) A statistical approach to the theory of the central nervous system. *The bulletin of mathematical biophysics* 10(1):41–55. DOI: 10.1007/BF02478329

Sholl DA (1953) Dendritic organization in the neurons of the visual and motor cortices of the cat. *Journal of Anatomy* 87(Pt 4):387–406.1

Singer W (1999) Neuronal Synchrony: A Versatile Code for the Definition of Relations? *Neuron* 24(1):49–65. DOI: 10.1016/S0896-6273(00)80821-1

Skaggs WE, McNaughton BL (1996) Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Science* 271(5257):1870–1873. DOI: 10.1126/science.271.5257.1870

Sporns O, Kötter R (2004) Motifs in Brain Networks. *PLoS Biology* 2(11):e369. DOI: 10.1371/journal.pbio.0020369

Stella A, Bouss P, Palm G, et al (2022) Comparing Surrogates to Evaluate Precisely Timed Higher-Order Spike Correlations. *eNeuro* 9(3). DOI: 10.1523/ENEURO.0505-21.2022

Stevenson NJ, Tapani K, Lauronen L, et al (2019) A dataset of neonatal EEG recordings with seizure annotations. *Scientific Data* 6:190,039. DOI: 10.1038/sdata.2019.39

Suresh J, Radojicic M, Pesce LL, et al (2016) Network burst activity in hippocampal neuronal cultures: the role of synaptic and intrinsic currents. *J Neurophysiol* 115(6):3073–3089. DOI: 10.1152/jn.00995.2015

Tamamaki N, Tomioka R (2010) Long-Range GABAergic Connections Distributed throughout the Neocortex and their Possible Function. *Frontiers in Neuroscience* 4:202. DOI: 10.3389/fnins.2010.00202

Tjepkema-Cloostermans MC, de Carvalho RCV, van Putten MJAM (2018) Deep learning for detection of focal epileptiform discharges from scalp EEG recordings. *Clinical Neurophysiology* 129(10):2191–2196. DOI: 10.1016/j.clinph.2018.06.024

Truccolo W, Donoghue JA, Hochberg LR, et al (2011) Single-neuron dynamics in human focal epilepsy. *Nature Neuroscience* 14(5):635–641. DOI: 10.1038/nn.2782

Tryba AK, Merricks EM, Lee S, et al (2019) Role of paroxysmal depolarization in focal seizure activity. *Journal of Neurophysiology* 122(5):1861–1873. DOI: 10.1152/jn.00392.2019

Victor JD (1994) Images, statistics, and textures: implications of triple correlation uniqueness for texture statistics and the Julesz conjecture: comment. *J Opt Soc Am A, JOSAA* 11(5):1680–1684. DOI: 10.1364/JOSAA.11.001680

Werther T, Olischar M, Naulaers G, et al (2017) Are All Amplitude-Integrated Electroencephalogram Systems Equal? *Neonatology* 112(4):394–401. DOI: 10.1159/000480008

Wilson HR, Cowan JD (1972) Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal* 12(1):1–24. DOI: 10.1016/S0006-3495(72)86068-5

Wilson HR, Cowan JD (1973) A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* 13(2):55–80. DOI: 10.1007/BF00288786

Yellott JI (1993) Implications of triple correlation uniqueness for texture statistics and the Julesz conjecture. *Journal of the Optical Society of America A* 10(5):777–793. DOI: 10.1364/JOSAA.10.000777

Yellott JI, Iverson GJ (1992) Uniqueness properties of higher-order autocorrelation functions. *Journal of the Optical Society of America A* 9(3):388–404. DOI: 10.1364/JOSAA.9.000388

Yu S, Yang H, Nakahara H, et al (2011) Higher-Order Interactions Characterized in Cortical Activity. *Journal of Neuroscience* 31(48):17,514–17,526. DOI: 10.1523/JNEUROSCI.3127-11.2011

Zhang D, Ding H (2013) Calculation of compact amplitude-integrated EEG tracing and upper and lower margins using raw EEG data. *Health* 5(5):885–891. DOI: 10.4236/health.2013.55116

Appendix

A Enumerated lag-sign motifs for triple correlation

Lag-Sign Motif	n_1	t_1	n_2	t_2	Constraints	Motif Class	Configuration
0	0	0	0	0		0	
1.0	0	0	0	+		I	
1.1	0	0	0	-		I	

Table A1: **Enumerated lag-sign motifs for triple correlation.** In the triple correlation equation, there are four lags that can be either zero or non-zero. Accordingly, the rows in this table are ordered in 16 (2^4) groups while the subgroups within these groups, represent the arrangement of the non-zero lags. The total number of rows is 169 lag-sign motifs for three-node motifs with distinct lag signs (Table S1). Each row represents a single three-node lag-sign motif, which are differentiated by the spike sequencing. The base node implicitly has lags $n_0 = 0$ and $t_0 = 0$, and the four lags of the other two nodes (n_1, n_2, t_1, t_2) are those in Equation (2.27). In the lag columns, '0' entries denote that the lag is zero while '+' entries denote a positive value, and '-' entries denote a negative value (respectively greater or less relative to the base node's spatiotemporal position). Note that in case where one modality (space or time) shares the same (non-zero) sign, we need additional constraints to specify the spike sequence, given by the column "Constraints." The 14 motif classes are labelled by Roman numerals in the third column group, "Motif Class". The "Configuration" column illustrates a representative schematic of the lag-motif, where the solid dot, small open circle, and large open circle indicate the base, second, and third nodes, respectively.

Table A1 continued.

2.0	0	0	+	0		III	
2.1	0	0	-	0		III	
3.0	0	0	+	+		V	
3.1	0	0	+	-		V	
3.2	0	0	-	+		V	
3.3	0	0	-	-		V	
4.0	0	+	0	0		I	
4.1	0	-	0	0		I	
5.0.0	0	+	0	+	$\lvert t_1\rvert = \lvert t_2\rvert$	I	
5.0.1	0	+	0	+	$\lvert t_1\rvert > \lvert t_2\rvert$	II	
5.0.2	0	+	0	+	$\lvert t_1\rvert < \lvert t_2\rvert$	II	
5.1	0	+	0	-		II	
5.2	0	-	0	+		II	
5.3.0	0	-	0	-	$\lvert t_1\rvert = \lvert t_2\rvert$	I	
5.3.1	0	-	0	-	$\lvert t_1\rvert > \lvert t_2\rvert$	II	
5.3.2	0	-	0	-	$\lvert t_1\rvert < \lvert t_2\rvert$	II	
6.0	0	+	+	0		VI	

Table A1 continued.

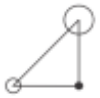
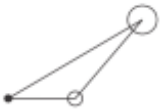
6.1	0	+	-	0		VI	
6.2	0	-	+	0		VII	
6.3	0	-	-	0		VII	
7.0.0	0	+	+	+	$\|t_1\| = \|t_2\|$	VII	
7.0.1	0	+	+	+	$\|t_1\| > \|t_2\|$	IX	
7.0.2	0	+	+	+	$\|t_1\| < \|t_2\|$	X	
7.1	0	+	+	-		VIII	
7.2.0	0	+	-	+	$\|t_1\| = \|t_2\|$	VII	
7.2.1	0	+	-	+	$\|t_1\| > \|t_2\|$	IX	
7.2.2	0	+	-	+	$\|t_1\| < \|t_2\|$	X	
7.3	0	+	-	-		VIII	
7.4	0	-	+	+		X	
7.5.0	0	-	+	-	$\|t_1\| = \|t_2\|$	VI	

Table A1 continued.

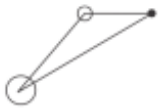
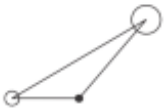
7.5.1	0	-	+	-	$\|t_1\| > \|t_2\|$	IX	
7.5.2	0	-	+	-	$\|t_1\| < \|t_2\|$	VIII	
7.6	0	-	-	+		X	
7.7.0	0	-	-	-	$\|t_1\| = \|t_2\|$	VI	
7.7.1	0	-	-	-	$\|t_1\| > \|t_2\|$	IX	
7.7.2	0	-	-	-	$\|t_1\| < \|t_2\|$	VIII	
8.0	+	0	0	0		III	
8.1	-	0	0	0		III	
9.0	+	0	0	+		VI	
9.1	+	0	0	-		VII	
9.2	-	0	0	+		VI	
9.3	-	0	0	-		VII	
10.0.0	+	0	+	0	$\|n_1\| = \|n_2\|$	III	

Table A1 continued.

10.0.1	+	0	+	0	$\lvert n_1\rvert > \lvert n_2\rvert$	IV	
10.0.2	+	0	+	0	$\lvert n_1\rvert < \lvert n_2\rvert$	IV	
10.1	+	0	-	0		IV	
10.2	-	0	+	0		IV	
10.3.0	-	0	-	0	$\lvert n_1\rvert = \lvert n_2\rvert$	III	
10.3.1	-	0	-	0	$\lvert n_1\rvert > \lvert n_2\rvert$	IV	
10.3.2	-	0	-	0	$\lvert n_1\rvert < \lvert n_2\rvert$	IV	
11.0.0	+	0	+	+	$\lvert n_1\rvert = \lvert n_2\rvert$	VI	
11.0.1	+	0	+	+	$\lvert n_1\rvert > \lvert n_2\rvert$	XII	

Table A1 continued.

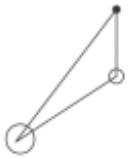
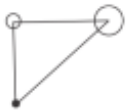
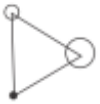
11.0.2	+	0	+	+	$\lvert n_1\rvert < \lvert n_2\rvert$	XII	
11.1.0	+	0	+	-	$\lvert n_1\rvert = \lvert n_2\rvert$	VII	
11.1.1	+	0	+	-	$\lvert n_1\rvert > \lvert n_2\rvert$	XI	
11.1.2	+	0	+	-	$\lvert n_1\rvert < \lvert n_2\rvert$	XI	
11.2	+	0	-	+		XII	
11.3	+	0	-	-		XI	
11.4	-	0	+	+		XII	
11.5	-	0	+	-		XI	
11.6.0	-	0	-	+	$\lvert n_1\rvert = \lvert n_2\rvert$	VI	
11.6.1	-	0	-	+	$\lvert n_1\rvert > \lvert n_2\rvert$	XII	

Table A1 continued.

11.6.2	-	0	-	+	$\lvert n_1\rvert < \lvert n_2\rvert$	XII	
11.7.0	-	0	-	-	$\lvert n_1\rvert = \lvert n_2\rvert$	VII	
11.7.1	-	0	-	-	$\lvert n_1\rvert > \lvert n_2\rvert$	XI	
11.7.2	-	0	-	-	$\lvert n_1\rvert < \lvert n_2\rvert$	XI	
12.0	+	+	0	0		V	
12.1	+	-	0	0		V	
12.2	-	+	0	0		V	
12.3	-	-	0	0		V	
13.0.0	+	+	0	+	$\lvert t_1\rvert = \lvert t_2\rvert$	VII	
13.0.1	+	+	0	+	$\lvert t_1\rvert > \lvert t_2\rvert$	X	
13.0.2	+	+	0	+	$\lvert t_1\rvert < \lvert t_2\rvert$	IX	
13.1	+	+	0	-		X	

Table A1 continued.

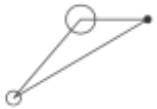
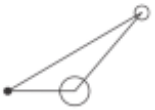
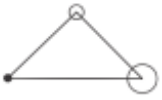
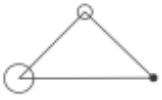
13.2	+	-	0	+		VIII	
13.3.0	+	-	0	-	$\|t_1\| = \|t_2\|$	VI	
13.3.1	+	-	0	-	$\|t_1\| > \|t_2\|$	VIII	
13.3.2	+	-	0	-	$\|t_1\| < \|t_2\|$	IX	
13.4.0	-	+	0	+	$\|t_1\| = \|t_2\|$	VII	
13.4.1	-	+	0	+	$\|t_1\| > \|t_2\|$	X	
13.4.2	-	+	0	+	$\|t_1\| < \|t_2\|$	IX	
13.5	-	+	0	-		X	
13.6	-	-	0	+		VIII	
13.7.0	-	-	0	-	$\|t_1\| = \|t_2\|$	VI	
13.7.1	-	-	0	-	$\|t_1\| > \|t_2\|$	VIII	
13.7.2	-	-	0	-	$\|t_1\| < \|t_2\|$	IX	
14.0.0	+	+	+	0	$\|n_1\| = \|n_2\|$	VI	

Table A1 continued.

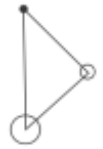
14.0.1	+	+	+	0	$\lvert n_1\rvert > \lvert n_2\rvert$	XII	
14.0.2	+	+	+	0	$\lvert n_1\rvert < \lvert n_2\rvert$	XII	
14.1	+	+	-	0		XII	
14.2.0	+	-	+	0	$\lvert n_1\rvert = \lvert n_2\rvert$	VII	
14.2.1	+	-	+	0	$\lvert n_1\rvert > \lvert n_2\rvert$	XI	
14.2.2	+	-	+	0	$\lvert n_1\rvert < \lvert n_2\rvert$	XI	
14.3	+	-	-	0		XI	
14.4	-	+	+	0		XII	
14.5.0	-	+	-	0	$\lvert n_1\rvert = \lvert n_2\rvert$	VI	

Table A1 continued.

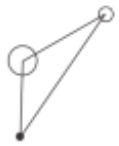
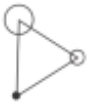
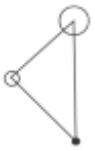
14.5.1	-	+	-	0	$\|n_1\| > \|n_2\|$	XII	
14.5.2	-	+	-	0	$\|n_1\| < \|n_2\|$	XII	
14.6	-	-	+	0		XI	
14.7.0	-	-	-	0	$\|n_1\| = \|n_2\|$	VII	
14.7.1	-	-	-	0	$\|n_1\| > \|n_2\|$	XI	
14.7.2	-	-	-	0	$\|n_1\| < \|n_2\|$	XI	
15.0.0	+	+	+	+	$\|n_1\| = \|n_2\|$ $\|t_1\| = \|t_2\|$	V	
15.0.1	+	+	+	+	$\|n_1\| = \|n_2\|$ $\|t_1\| > \|t_2\|$	VIII	
15.0.2	+	+	+	+	$\|n_1\| = \|n_2\|$ $\|t_1\| < \|t_2\|$	VIII	
15.0.3	+	+	+	+	$\|n_1\| > \|n_2\|$ $\|t_1\| = \|t_2\|$	XI	
15.0.4	+	+	+	+	$\|n_1\| > \|n_2\|$ $\|t_1\| > \|t_2\|$	XIII	

Table A1 continued.

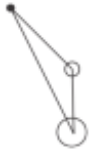
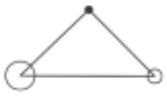
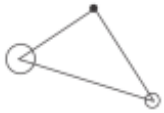
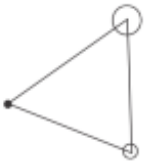
15.0.5	+	+	+	+	$\lvert n_1\rvert > \lvert n_2\rvert$ $\lvert t_1\rvert < \lvert t_2\rvert$	XIII	
15.0.6	+	+	+	+	$\lvert n_1\rvert < \lvert n_2\rvert$ $\lvert t_1\rvert = \lvert t_2\rvert$	XI	
15.0.7	+	+	+	+	$\lvert n_1\rvert < \lvert n_2\rvert$ $\lvert t_1\rvert > \lvert t_2\rvert$	XIII	
15.0.8	+	+	+	+	$\lvert n_1\rvert < \lvert n_2\rvert$ $\lvert t_1\rvert < \lvert t_2\rvert$	XIII	
15.1.0	+	+	+	-	$\lvert n_1\rvert = \lvert n_2\rvert$	IX	
15.1.1	+	+	+	-	$\lvert n_1\rvert > \lvert n_2\rvert$	XIII	
15.1.2	+	+	+	-	$\lvert n_1\rvert < \lvert n_2\rvert$	XIII	
15.2.0	+	+	-	+	$\lvert t_1\rvert = \lvert t_2\rvert$	XI	
15.2.1	+	+	-	+	$\lvert t_1\rvert > \lvert t_2\rvert$	XIII	
15.2.2	+	+	-	+	$\lvert t_1\rvert < \lvert t_2\rvert$	XIII	

Table A1 continued.

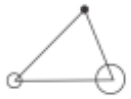
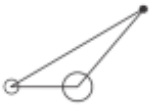
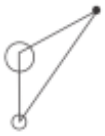
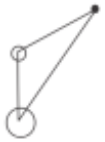
15.3	+	+	-	-		XIII	
15.4.0	+	-	+	+	$\|n_1\| = \|n_2\|$	IX	
15.4.1	+	-	+	+	$\|n_1\| > \|n_2\|$	XIII	
15.4.2	+	-	+	+	$\|n_1\| < \|n_2\|$	XIII	
15.5.0	+	-	+	-	$\|n_1\| = \|n_2\|$ $\|t_1\| = \|t_2\|$	V	
15.5.1	+	-	+	-	$\|n_1\| = \|n_2\|$ $\|t_1\| > \|t_2\|$	X	
15.5.2	+	-	+	-	$\|n_1\| = \|n_2\|$ $\|t_1\| < \|t_2\|$	X	
15.5.3	+	-	+	-	$\|n_1\| > \|n_2\|$ $\|t_1\| = \|t_2\|$	XII	
15.5.4	+	-	+	-	$\|n_1\| > \|n_2\|$ $\|t_1\| > \|t_2\|$	XIII	
15.5.5	+	-	+	-	$\|n_1\| > \|n_2\|$ $\|t_1\| < \|t_2\|$	XIII	
15.5.6	+	-	+	-	$\|n_1\| < \|n_2\|$ $\|t_1\| = \|t_2\|$	XII	

Table A1 continued.

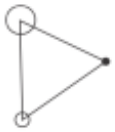
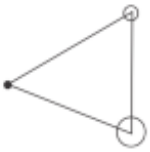
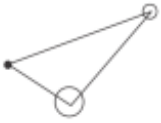
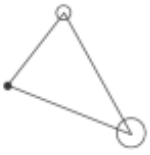
15.5.7	+	-	+	-	$\lvert n_1\rvert < \lvert n_2\rvert$ $\lvert t_1\rvert > \lvert t_2\rvert$	XIII	
15.5.8	+	-	+	-	$\lvert n_1\rvert < \lvert n_2\rvert$ $\lvert t_1\rvert < \lvert t_2\rvert$	XIII	
15.6	+	-	-	+		XIII	
15.7.0	+	-	-	-	$\lvert t_1\rvert = \lvert t_2\rvert$	XII	
15.7.1	+	-	-	-	$\lvert t_1\rvert > \lvert t_2\rvert$	XIII	
15.7.2	+	-	-	-	$\lvert t_1\rvert < \lvert t_2\rvert$	XIII	
15.8.0	-	+	+	+	$\lvert t_1\rvert = \lvert t_2\rvert$	XI	
15.8.1	-	+	+	+	$\lvert t_1\rvert > \lvert t_2\rvert$	XIII	
15.8.2	-	+	+	+	$\lvert t_1\rvert < \lvert t_2\rvert$	XIII	

Table A1 continued.

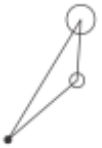
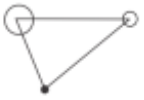
15.9	-	+	+	-		XIII	
15.10.0	-	+	-	+	$\lvert n_1\rvert = \lvert n_2\rvert$ $\lvert t_1\rvert = \lvert t_2\rvert$	V	
15.10.1	-	+	-	+	$\lvert n_1\rvert = \lvert n_2\rvert$ $\lvert t_1\rvert > \lvert t_2\rvert$	VIII	
15.10.2	-	+	-	+	$\lvert n_1\rvert = \lvert n_2\rvert$ $\lvert t_1\rvert < \lvert t_2\rvert$	VIII	
15.10.3	-	+	-	+	$\lvert n_1\rvert > \lvert n_2\rvert$ $\lvert t_1\rvert = \lvert t_2\rvert$	XI	
15.10.4	-	+	-	+	$\lvert n_1\rvert > \lvert n_2\rvert$ $\lvert t_1\rvert > \lvert t_2\rvert$	XIII	
15.10.5	-	+	-	+	$\lvert n_1\rvert > \lvert n_2\rvert$ $\lvert t_1\rvert < \lvert t_2\rvert$	XIII	
15.10.6	-	+	-	+	$\lvert n_1\rvert < \lvert n_2\rvert$ $\lvert t_1\rvert = \lvert t_2\rvert$	XI	
15.10.7	-	+	-	+	$\lvert n_1\rvert < \lvert n_2\rvert$ $\lvert t_1\rvert > \lvert t_2\rvert$	XIII	
15.10.8	-	+	-	+	$\lvert n_1\rvert < \lvert n_2\rvert$ $\lvert t_1\rvert < \lvert t_2\rvert$	XIII	
15.11.0	-	+	-	-	$\lvert n_1\rvert = \lvert n_2\rvert$	IX	

Table A1 continued.

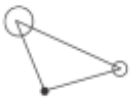
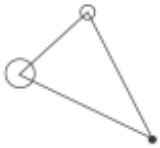
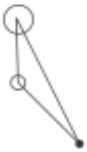
15.11.1	-	+	-	-	$\lvert n_1\rvert > \lvert n_2\rvert$	XIII	
15.11.2	-	+	-	-	$\lvert n_1\rvert < \lvert n_2\rvert$	XIII	
15.12	-	-	+	+		XIII	
15.13.0	-	-	+	-	$\lvert t_1\rvert = \lvert t_2\rvert$	XII	
15.13.1	-	-	+	-	$\lvert t_1\rvert > \lvert t_2\rvert$	XIII	
15.13.2	-	-	+	-	$\lvert t_1\rvert < \lvert t_2\rvert$	XIII	
15.14.0	-	-	-	+	$\lvert n_1\rvert = \lvert n_2\rvert$	IX	
15.14.1	-	-	-	+	$\lvert n_1\rvert > \lvert n_2\rvert$	XIII	
15.14.2	-	-	-	+	$\lvert n_1\rvert < \lvert n_2\rvert$	XIII	
15.15.0	-	-	-	-	$\lvert n_1\rvert = \lvert n_2\rvert$ $\lvert t_1\rvert = \lvert t_2\rvert$	V	
15.15.1	-	-	-	-	$\lvert n_1\rvert = \lvert n_2\rvert$ $\lvert t_1\rvert > \lvert t_2\rvert$	X	

Table A1 continued.

15.15.2	-	-	-	-	$\lvert n_1\rvert = \lvert n_2\rvert$ $\lvert t_1\rvert < \lvert t_2\rvert$	X	
15.15.3	-	-	-	-	$\lvert n_1\rvert > \lvert n_2\rvert$ $\lvert t_1\rvert = \lvert t_2\rvert$	XII	
15.15.4	-	-	-	-	$\lvert n_1\rvert > \lvert n_2\rvert$ $\lvert t_1\rvert > \lvert t_2\rvert$	XIII	
15.15.5	-	-	-	-	$\lvert n_1\rvert > \lvert n_2\rvert$ $\lvert t_1\rvert < \lvert t_2\rvert$	XIII	
15.15.6	-	-	-	-	$\lvert n_1\rvert < \lvert n_2\rvert$ $\lvert t_1\rvert = \lvert t_2\rvert$	XII	
15.15.7	-	-	-	-	$\lvert n_1\rvert < \lvert n_2\rvert$ $\lvert t_1\rvert > \lvert t_2\rvert$	XIII	
15.15.8	-	-	-	-	$\lvert n_1\rvert < \lvert n_2\rvert$ $\lvert t_1\rvert < \lvert t_2\rvert$	XIII	

Description of Supplementary Files

In addition to the Github links in the text above, alongside this PDF I publish zip files containing the code for each chapter at time of publication. However, I recommend the use the Github repositories linked above whenever possible. These online repositories offer a more streamlined installation, and the possibility of updated code. In all cases, installation and usage instructions are within the repository READMEs.