

SYNAPTIC SHAPING OF CORTICAL DYNAMICS AT SPECIES SPECIFIC SCALES

JOHN BENINGER

THESIS SUBMITTED TO THE UNIVERSITY OF OTTAWA IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY IN NEUROSCIENCE

DEPARTMENT OF CELLULAR AND MOLECULAR MEDICINE
FACULTY OF MEDICINE
UNIVERSITY OF OTTAWA

© JOHN BENINGER, OTTAWA, CANADA, 2026

ABSTRACT

Events in the world are structured both spatially and temporally: while brains vary greatly in size across species, they must coordinate neuronal populations in response to shared environmental timescales. How are spatial and temporal structure scalably encoded in cortical dynamics? Individually, synapses show preferential responses to particular temporal patterns of activity enabled by transient changes in connection strength known as short-term plasticity. At a group level, one idea is that sensory stimulation evokes cortical travelling waves which share information between spatially distant neurons and act as a record of the recent past. However, individual synaptic dynamics display a large degree of heterogeneity that is informally separated into classes despite the precise number and properties of such classes being unclear. Moreover, species with larger brains have more neurons distributed over larger spatial volumes: for cortical waves to integrate information from topologically organized receptive fields in different species, they must propagate via mechanisms that scale proportionately. To make sense of this complexity, this thesis leverages modelling spanning from the scale of synapses to networks, as well as machine learning tools, a large open synaptic physiology dataset, and novel experiments performed by collaborating scientists. We present a picture of the cortical microcircuit where synaptic dynamics belong to a nuanced dictionary of functional subtypes and where spatially structured synaptic connectivity encodes local perturbations into cortical travelling waves. In rodent data, we find five functional clusters of synaptic dynamics that partially converge with transgenically defined subtypes. Strikingly, the application of the same clustering method in human data infers a highly similar number of subtypes, supportive of stable clustering. At the network level, our modelling elucidates how cortical synaptic connectivity may selectively encode localized perturbations into transient travelling waves, a prediction consistent with new experimental results from electrical microstimulation of rodent cortical tissue. Crucially, our modelling finds these rodent data consistent with two rival connectivity schemes: one in which perturbations of both human and mouse tissue produce waves of similar scale and a second in which evoked waves propagate more than twice as far in models of human cortex, consistent with a role in species specific scaling of spatiotemporal integration. Taken together, this thesis provides a cross-species account of how the structure of cortical connectivity shapes dynamics across networks including a nuanced dictionary of specific functional subtypes of cortical connections and a falsifiable

mechanistic prediction about how neuronal populations of species specific scale may be coordinated in response to shared environmental scales.

ACKNOWLEDGMENTS

I have been distinctly lucky to receive the support of an entire community of collaborators, mentors, friends and family without whom this thesis would not exist.

I would like to thank my supervisor Richard Naud for providing both patient support and rigorous intellectual challenge, teaching me a great deal, and encouraging me to take risks and work across boundaries.

I would also like to thank my co-supervisor Katalin Tòth without whom I would not have developed the understanding of basic physiology necessary to make science out of models.

Thank you to Mark Harnett for allowing me to visit in his lab, expanding my understanding of physiology, and broadening my foundation in neuroscience.

Thank you to my thesis advisory committee, Leonard Maler, Jérémie Lefebvre, and Jean-Philippe Thivierge for years of feedback and guidance.

I would like to thank the many other formal collaborators who made the work in this thesis and in my broader PhD possible: Cynthia Rais, Jade Poirier, Julian Rossbroich, Sydney Cash, and Daniel Trotter.

I would also like to thank the many people who contributed informally to these works or to my journey as a scientist more broadly especially Zachary Friedenberger, Maxim Piatine, Nate Jones, Jan Ranier Sidiangco, Nima Raman, Minh Tran, Edward Sun, the other past and present members of the Naud, Bèïque, Tòth, Chen, and Harnett labs, and Andrew Hamilton-Wright without whom I may not have pursued a path in science to begin with.

Reading a thesis is itself a substantial commitment and for that I thank David DiGregorio, Jean-Claude Bèïque, Jean-Philippe Thivierge, and Mathieu Lavallée-Adam.

Finally, I would like to emphatically thank my siblings, my grandparents, and above all my mother for their years of support.

This work was made possible by financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC), the Government of Ontario, the Bonnie

and Don Poole Parkinson's Research Fellowship, and the University of Ottawa, as well as project level grants from the The Canadian Institutes of Health Research (CIHR) and the National Institutes of Health (NIH).

AUTHOR CONTRIBUTIONS

I am the first author of the paper presented in chapter 2 which has been published in the journal Cell Reports [1]. In that work, we established a novel pipeline to test for the existence of functional subtypes of synaptic dynamics and identified a collection of such subtypes of excitatory synapses in human and rodent data. For that paper I wrote all of the novel code except for a database access and initial data cleaning script which was written by my co-author Julian Rossbroich. I also designed the methods and performed all of the analysis, produced all of the figures, and wrote the initial draft of the manuscript. These tasks were performed with the contribution of regular critical feedback and suggestions from my supervisor Richard Naud and my co-supervisor Katalin Tóth as well as periodic technical discussions with Julian Rossbroich. All authors participated in editing and revision of the manuscript.

Chapter 3 is adapted from a manuscript on which I share first authorship. The manuscript is presently in preparation for submission as a publication. The work emerged from a close international collaboration with our experimental partners in the labs of Mark Harnett (Massachusetts Institute of Technology), and Sydney Cash (Massachusetts General Hospital) during which I spent several months as a visiting PhD student in the lab of Mark Harnett at the Massachusetts Institute of Technology. I designed the spatially structured spiking neural network model and the field model, wrote the related simulation packages, performed the simulations, performed the related analysis, and produced the related figures. I also worked closely with the other first author, Cynthia Rais, to refine experimental protocols appropriate for testing the predictions of our theoretical models. Cynthia Rais, performed all of the experiments as well as the related analysis and figure production. Richard Naud performed the axonal cable recruitment modeling, as well as the associated analysis, and figure production. All members of the project (Cynthia Rais, John Beninger (myself), Sydney Cash, Mark Harnett, and Richard Naud) were regularly involved in shaping the direction of both the theoretical and experimental work. Initial writing of the manuscript was performed jointly by myself, Cynthia Rais, and Richard Naud. All authors contributed to editing and revision of the manuscript.

Chapter 4 contains the results of simulations that I have performed in order to better understand the implications of the results in chapter 3 for systems neuroscience. I

performed all of work and writing in the chapter while benefiting from ongoing scientific discussions with, and feedback from, Richard Naud.

During my doctoral studies, I also had a middle author role on the paper which introduced the Spike Response Plasticity (SRP) model of short term plasticity in the journal PLOS Computational Biology [2]. My role in that work was centered on packaging the model software and improving software design quality and I was also involved in editing and review of the manuscript. Following publication of the work, I have acted as the maintainer for the SRP Mmodel software package and have steadily refined it. The content of the SRP model paper is not included in a standalone manner in this thesis, but the model was a key tool for the paper in chapter 2 and is introduced in that context.

During my doctoral studies, I performed research for which I am a co-first author, alongside Jade Poirier, on a paper presently that is presently under review and is available as a preprint [3] in which we have applied the pipeline I developed for the work in chapter 2 of this thesis to the analysis of the functional organization of short term plasticity in GABAergic synapses. I developed the initial code base for that work as well a large share of the methods, performed parts of the analysis directly, played a minority role in early figure making and writing, contributed to editing and review of the manuscript, and provided training and guidance throughout the project. While I discuss that paper in the conclusions of this thesis, the content of paper is not included in this thesis document.

During my doctoral studies, I performed work for which I am a co-first author, again alongside Jade Poirier, on a paper published at the journal Star Protocols [4] that includes refinements to the SRP model of short term plasticity and related software as well as detailed documentation and instructions designed to help people effectively use the model with their own data. Jade and I performed relatively even shares of the writing, software engineering and figure production for that paper with feedback and editing from Richard Naud. While software related to the SRP model played a role in all of my analyses of short term plasticity and was a part of my overall multi-year stewardship of the related software package, the content of that paper is not directly included in this thesis document.

CONTENTS

ABSTRACT	ii
ACKNOWLEDGMENTS	iv
AUTHOR CONTRIBUTIONS	vi
LIST OF FIGURES	xi
LIST OF TABLES	xiii
1 INTRODUCTION	1
1.0.1 Short-Term Plasticity	4
1.0.2 Electrical Microstimulation	7
1.0.3 Cortical Waves	9
I Synaptic Dynamics	16
2 FUNCTIONAL SUBTYPES OF SYNAPTIC DYNAMICS IN MOUSE AND HUMAN	18
2.1 Introduction	19
2.2 Results	22
2.2.1 Model captures the diversity of synaptic dynamics	22
2.2.2 Unsupervised clustering reveals functional subtypes	25
2.2.3 Functional subtypes cannot be explained by continuum hypotheses	28
2.2.4 Conserved functional subtypes are found in human synapses	32
2.2.5 Supervised prediction of transgenic identity relies on similar functional subtypes	34

2.3	Discussion	38
2.3.1	Limitations of the study	41
2.4	Methods	43
2.4.1	Resource Availability	43
2.4.2	METHOD DETAILS	44
2.4.3	Spike Response Plasticity (SRP) Model	44
2.4.4	Data Processing	46
2.4.5	Model Fitting	47
2.4.6	Unsupervised Learning	48
2.4.7	Supervised Learning	48
2.4.8	QUANTIFICATION AND STATISTICAL ANALYSIS	50
2.4.9	Fit Error Statistics	50
2.4.10	Testing Continuum Hypotheses	50
2.4.11	Cluster and Group Similarity Measure	51
2.4.12	Cluster Number Occurrence Probability	52
2.4.13	Author Contributions	54
2.4.14	Declaration of interests	55
2.4.15	Supplementary Figures	56

II Cortical Waves 63

3	BIOPHYSICAL MECHANISMS UNDERLYING THE CORTICAL RESPONSE TO ELECTRICAL MICROSTIMULATION	65
3.1	Introduction	67
3.2	Results	70
3.2.1	Two modes of microstimulation-evoked activation produce distinct somatic signatures	70
3.2.2	Axial current flow accounts for microstimulation-evoked potentials .	73
3.2.3	Axotomy reveals relative contribution of dendrites and axons	73
3.2.4	Differences in axonal cable properties predict a recruitment bias across cell types	74
3.2.5	PV INs are more readily recruited by microstimulation	80
3.2.6	Transient Expansion is Consistent with Transient Travelling Waves in a Two Population Field Model with Refractoriness	81

3.2.7	Transient Expansion in Spiking Models Despite Preferential Recruitment of Inhibitory Cells	84
3.2.8	Transient expansion in acute cortical slices	87
3.3	Discussion	89
3.4	Methods	92
3.4.1	Ex vivo experiments	92
3.4.2	Analytical Methods	96
3.4.3	Field model	104
3.4.4	Spiking Neural Network Model	109
3.4.5	Table of Symbols, Sections 3.4.1 and 3.4.2	116
3.4.6	Table of Symbols, Section 3.4.4	117
4	STRUCTURED CONNECTIVITY MAY ENCODE LOCAL PERTURBATIONS IN CORTICAL TRAVELLING WAVES	122
4.1	Introduction	122
4.2	Extended Structured Cortical Spiking Neural Network Model	124
4.3	Results	126
4.4	Discussion	134
III	Conclusion	136
5	CONCLUSION	137
	REFERENCES	145

LIST OF FIGURES

1.1	Topographic Propagation of Sensory Evoked Waves, Reproduced under a creative commons license	13
2.1	The Spike Response Plasticity (SRP) model captures responses across a diversity of synapse types.	24
2.2	Unsupervised clustering reveals 5 clusters in rodent data.	27
2.3	Testing uniform and continuum hypotheses in rodent.	30
2.4	Unsupervised clustering shows functional alignment between rodent and human data.	33
2.5	Clusters correspond to model features that are helpful for predicting pre- or post-synaptic cell identity.	36
2.6	Kernels of SRP model fits by pre-synaptic	56
2.7	Clusters by pair identity.	57
2.8	Example fits by cluster.	58
2.9	Testing uniform and continuum hypotheses in human.	59
2.10	Supervised accuracy by extracellular calcium concentration matched for dataset size.	60
2.11	Supplemental results by dataset and representation	60
2.12	Supplemental out-of-sample results for alternate classification algorithms .	61
2.13	Supplemental within-sample results for alternate classification algorithms .	62
3.1	Near-electrode patch clamp distinguishes somatic from cable effects of microstimulation	71
3.2	Axotomy distinguishes dendritic and axonal contributions of microstimulation	75
3.3	Cable properties predict differential recruitment of cortical cell types confirmed in slice experiments	76

3.4	Strong stimulation induces a transient travelling wave pulse in a field model with refractory terms	79
3.5	Spiking neural network model of the cortical response to electrical microstimulation	81
3.6	Transient spatial amplification by imbalance of E-I	88
3.7	Application of the theory to biphasic pulses	97
3.8	Relationship between activation radius of axon and soma	99
3.9	Comparison of field models with and without refractory variable	107
3.10	Time steps from example pulse response of 2D spiking model corresponding to supplemental animation.	119
3.11	Characterization of evoked responses for pyramidal and PV cells	120
3.12	Comparison of the effect of synaptic blockers on peak potential for different distances	121
4.1	Spatially structured spiking neural network model	125
4.2	Species specific spatial scaling of predicted microstimulation induced transient travelling waves	126
4.3	Human and Rodent Microstimulation Responses	128
4.4	Synaptically evoked waves in mouse SNN	130
4.5	Synaptically evoked waves in human SNN	132
4.6	Spatial response distribution of responses to different temporal patterns . .	133

LIST OF TABLES

2.1	Median SRP parameters of rodent synapses used in clustering	56
2.2	Median SRP parameters of Human synapses used	57
3.1	Table of Symbols, Sections 3.4.1 and 3.4.2	116
3.2	Table of Symbols, Section 3.4.4	118

INTRODUCTION

For millennia many of the world's greatest thinkers have contemplated the nature of the mind and many considered the capacity for thought to be so extraordinary that they did not believe it could not arise as a property of physical matter [5, 6]. We have since identified the brain as a physical substrate that hosts this incredible power, neurons as fundamental building blocks of brains, and events of microscopic scale and incredible speed known as action potentials as the basic indivisible unit of information by which neurons communicate [7]. But how are these syllables combined to form the basic language of the brain and how does this language differ between the brains of humans and other animals? The human brain is the most powerful computational system ever encountered and is differentiated in part by the relative size and complexity of a set of layers known as the neocortex. These layers incorporate incredible organizational complexity, including considerable recurrent connectivity that gives rise to rich nonlinear dynamics and supports sophisticated

cognitive functions [8]. Neurons within this circuitry do not communicate each action potential (message) equally; they use dynamical rules to selectively communicate some at the expense of others [9–11]. The prevalence and complexity of such short-term plasticity highlights the importance of understanding its role in shaping information transmission. These rules are thought to be associated with specific pairs of neuron types, and those neuron types are in turn thought to have specific properties and connectivity patterns. This collection of properties is thought to largely define the basic, region-specific, circuitry of the brain. Moreover, this circuitry underpins complex patterns of activity that in turn underlie cognitive functions.

How is external information encoded in these dynamics? A large body of work in non-human animals and computational models has emerged to support the idea that sensory stimulation evokes travelling waves which integrate information across both space and time [12–14]. Synaptic cascades underlying “waves” of cortical activity, may encode sensory information [15–19] and reward history [20], and may direct attention by modifying perceptual sensitivity [21]. In order to understand these patterns mechanistically we pose two critical questions: First, what is the basic dictionary of neural communication? Second, how does the connectivity of the mammalian cortex shape the propagation and sharing of information?

Chapter 2 deal’s with the basic language of synaptic communication. Using the world’s largest paired patch clamp synaptic physiology dataset, synaptic modeling, and machine learning techniques we show that short-term plasticity in excitatory connections recorded in mice may organize into five functional categories (i.e “types”) while that in humans may

organize into six. We validated that these categories cannot be accounted for by continuum hypotheses, are similar between species, and partly, but incompletely, overlap transgenic associated cellular identities.

Having established a basic dictionary of synaptic dynamics, we then asked what information these synapses decode. Essentially, if these functional subtypes interpret the “words”, then what is the language of the brain? One idea is that sensory stimulation evokes cortical travelling waves which share information between spatially distant neurons allowing for the integration of information from distinct, topologically organized receptive fields and acting as a record of the recent past [12]. Because the empirical study of such dynamics will ultimately require a method by which to causally evoke cortical waves, in chapter 3 we developed cable, mean-field, and spiking neural network modelling to mechanistically establish the conditions under which electrical microstimulation can evoke transient travelling waves in cortical tissue. We also present the empirical results of our work to validate the predictions of these models in rodent cortical slices in partnership with experimental collaborators.

Finally, in Chapter 4, we use open data on synaptic connectivity in rodents and humans, as well as large spiking neural network simulations to predict the response of human cortical tissue to electrical microstimulation, and to explore the temporal features of synaptic input capable of evoking similar waves in the presence of background activity. While regular synaptic input on the order of 10 Hz did not trigger these waves, brief periods of high frequency (60 Hz) stimulation evoked travelling waves in both species that strikingly propagated approximately twice as far in models of human cortex as in models of mouse

cortex. This corresponds to travelling waves in the models of human cortex that amplify the same spatially localized signals to a greater degree and may coordinate larger neuronal populations in response to shared environmental timescales. These findings link a complete mechanistic account of the transient amplification and propagation of cortical information to the basic organizational principles of cortical synapses.

The remaining sections of this chapter introduce three areas of literature that are largely siloed at present: short-term plasticity, electrical microstimulation of cortex, and cortical waves. If short-term plasticity can decode the temporally coincident events the passage of travelling waves triggers, then the two may form a powerful temporally precise information processing system. Further, if microstimulation can evoke travelling waves, then it could be an excellent causal tool to study communication in this system. Whether such synergies exist remains an open question, and the foundation for studying them requires a functional taxonomy of cortical synaptic dynamics, a detailed mechanistic account of the cortical response to microstimulation, and understanding of travelling waves in cortex at the level of the individual spikes that compose them.

1.0.1 SHORT-TERM PLASTICITY

Examining the mechanisms through which neurons convey information offers insight into the fundamental language of the brain. This language is, in part, shaped by synapses that exhibit preferential responses to specific temporal patterns of activity. These responses are facilitated by short-term plasticity (STP), which refers to transient (10 ms - 10 s)

changes in connection strength [9–11]. The most commonly considered categories of STP are facilitation and depression, which involve transient increases or decreases in synaptic strength, respectively [22–25]. However, a far greater degree of complexity has been observed empirically [8].

In order to understand the fundamental motifs guiding information transmission in neural circuits, it is therefore important to pose several questions regarding the functional characteristics and underlying taxonomy of these dynamics. How should synapses be classified based on these properties, and how do functional properties vary across the resulting classes? Can these properties be traced back to an underlying genetic basis, and if so, what is the relevant level of genetic description? Is there a continuum of these properties, and if so, when is it appropriate to separate this continuum into distinct functional regions? Finally, what kinds of functional groupings are necessitated by different theories of neuronal information processing, and do these groupings align with empirical observations?

Short-term plasticity has long been thought to allow for the separation of messages to different post-synaptic cells from the same axon [26–29]. Based on timescale, STP can support detection of fine differences in low versus high firing rates [30–32] or between bursts and single spikes [31]. Short-term depression may also preserve time delay information at multiple levels [33–35] by acting as adaption mechanism [36]. STP may also support multiple network level roles such as enhancing correlated signals through facilitation [37] and decorrelating inputs through depression [38]. STP has been proposed to support working memory by shaping the emergence of persistent attractor dynamics at a network level [39] where it may also facilitate storage and retrieval of memories [40] or by acting as directly

as a basis of working memory [41]. In the context of circuits that apply multiplexing, STP can decode multiplexed signals by separating distinct signals from a given spike train [42–44]. This is of particular relevance to biological learning because multiplexed code has been used to solve the “credit assignment problem” which entails coordinating learning across multiple areas of the brain [45]. This application also highlights the importance of clearly understanding STP because proposed solutions often assume a particular relationship between neuron types, connectivity patterns, and STP properties [46]. STP may also contribute to solving the credit assignment problem by acting as a reservoir of timing information, allowing for sub-second timing to be used for the learning of temporal associations [47, 48]. These applications illustrate the importance of a precise understanding of the functional organisation of STP properties in the brain.

Despite the considerable heterogeneity of synaptic responses that exists within the brain, the most commonly applied classification of STP is the dichotomy of facilitation and depression. Nonetheless, more complex forms of STP exist including sub- and supra-linear forms of facilitation or depression [8, 49], biphasic STP which overlays both facilitation and depression [50, 51], and delayed burst-dependent STP [52]. This heterogeneity in STP properties also extends to synapses that share genetic and regional identities [8]. Importantly, the most appropriate scheme for classifying these dynamics has remained ambiguous.

1.0.2 ELECTRICAL MICROSTIMULATION

Implantable brain-computer interface (iBCI) technologies show incredible promise both as a means to infer useful information from the brain, and as a way to influence neural activity. The inference function of these technologies has been used clinically via brain-spine interfaces to restore walking ability to a paraplegic man [53] and near normal walking ability to a patient whose Parkinson's had progressed beyond other kinds of treatment [54], as well as restoring a host of other life changing functions [55]. Such functions are relevant to a global patient population increasing by well above a quarter of a million people per year [55, 56]. Similarly, iBCIs which provide electrical stimulation to correct dysregulated neural activity are applicable to many conditions including Epilepsy, Parkinson's, Alzheimer's OCD, depression, and more [57, 58].

While powerful frameworks have been built to allow for the modelling and design of electrical stimulation in the peripheral nervous system [59] and spinal cord [60, 61], we lack an equivalent tool for direct electrical stimulation (DES) of cortical regions. This is despite the incredible demonstrated clinical potential of cortical microstimulation [62, 63]. Models for peripheral nervous system and spinal cord often rely on selective explicit modelling of spatially detailed neuronal biophysics which is difficult in a cortical context given the incredible spatial complexity and recurrent connectivity of cortical tissue [64, 65]. Further, even cutting edge work has sometimes relied on computationally intensive grid search based optimization to plan electrode configurations and stimulus patterns [53]. Moreover, activity in the brain propagates through neural circuits that have complicated,

nonlinear dynamics [8] which can result in similarly complicated response patterns. As such, our ability to efficiently interpret these dynamics is presently limited [66, 67].

Experimental work has probed the mechanisms of cortical microstimulation, including its spatial propagation under ex-vivo conditions [68, 69] and more recently in vivo conditions [70, 71]. Meanwhile, reinforcement learning has been studied as a tool to optimize peripheral nervous stimulation [72] and machine learning is being extensively applied to the decoding of neural activity in the context of iBCIs [73–75]. However, despite advances allowing for the application of gradient based optimization to spiking neural networks [76] the effect of electrical microstimulation in humans is still studied through parameter search over complex forward simulations [65]. Meanwhile, differentiable simulation has been applied to optogenetic perturbations [77], which can only be achieved in animal models, but we lack a computationally efficient spatially structured model of direct electrical stimulation in rodent and human cortex.

When building models of electrical stimulation of the peripheral nervous system “Hybrid Models” are often used to bridge the gap between spatially detailed physics and neuronal activity [59]. First, detailed spatial descriptions of the target axons and resistive properties of tissue at the intended implant site are assembled [59]. Next, finite element models allow prediction of the relative electric field at each point in space while the activation threshold at each node of each axon is predicted using activation functions inferred from cable properties [78–81]. These methods can be extended to unmyelinated axon segments which we expect in cortex, by treating the axon as a series of discrete disks [80].

Cortex differs from the peripheral nervous system and is characterized by considerable recurrence and intricate axonal morphologies [64]. Owing to this complexity, in chapter 3, we take a different approach from the hybrid modelling strategy described above. Instead, we infer class specific recruitment probabilities from basic biophysical properties and then describe the direct effect of stimulation in cortex using a computationally tractable phenomenological equation. However, we also construct our model in a modular fashion so that predictions of the direct effect can be substituted if precise implant specific morphological characterization is available.

1.0.3 CORTICAL WAVES

Standing on the coast of any ocean, it is hard not to be struck by the sheer power and scale of waves approaching the shore. Many such waves travel thousands of kilometers across open ocean, carrying both substantial energy and an impression of the distant past events which formed them [82]. Yet the water shaped by these waves has not travelled with them: water is the medium through which waves propagate, but the waves themselves are transmitted effects carrying energy and information across the medium. In this sense, the waves are dynamic records of the oceans' events. These records link the shape of the ocean's surface in space to its history in time: if the waves close to shore tell the story of a distant storm, then waves beyond the horizon may tell the story of how that same storm ended. Given time, those distant waves will arrive at the shore. As such, a wave with a specified speed is, in effect, a dynamic mechanism by which a change in amplitude (e.g.

the height of the water wave) over time at a fixed location is translated into a change in the spatial location of the wave's peak at a fixed time. In a single spatial dimension, this relationship between space x and time t is often presented in terms of the second derivatives of the wave amplitude z as follows:

$$\frac{\partial^2 z}{\partial x^2} = \frac{1}{v^2} \frac{\partial^2 z}{\partial t^2} \quad (1.1)$$

Where v is the speed of the wave in the medium (this can differ across the medium which has implications for the dynamics of the wave). Wave equations can be generalized into any number of dimensions. In two dimensions the equivalent equation becomes

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = \frac{1}{v^2} \frac{\partial^2 z}{\partial t^2} \quad (1.2)$$

where y is the added spatial dimension. In two-dimensional systems with circular symmetry, such as the case of circular ripples evoked by a pebble dropped into a pond, two dimensional waves can be represented in terms of their distance r from their origin (e.g. the location of the pebble's impact) as follows:

$$\frac{\partial^2 z}{\partial r^2} + \frac{1}{r} \frac{\partial z}{\partial r} = \frac{1}{v^2} \frac{\partial^2 z}{\partial t^2} \quad (1.3)$$

In this equation it becomes clear that the amplitude of the travelling waves declines as a function of distance r from the location of the initial event because the energy of the wave is spread across a circular wave with a perimeter that increases as a function of the

radius r . As a result, the travelling waves created by the impact of a pebble persist as a record of the impact even once the pebble itself has disappeared beneath the surface. The waves carry that record away from the location of the event over time, but their peaks decline in amplitude as they travel.

WAVES IN EXCITABLE MEDIA

Ripples evoked by a pebble dropped into a pond lose amplitude as they travel because of the conservation of energy in physical media, but waves travel differently in environments where they can recruit energy as they propagate. For example, imagine a case where lightning strikes a tree in the middle of a dry forest and ignites a fire. In this scenario, the initial fire ignites surrounding trees which creates a larger fire, which in turn ignites even more trees. The result is a circular wave of flame with an amplitude (heat) that is not forced to decline as a function of distance because trees at the front provide additional energy to the fire as it propagates. Such environments are known as excitable media [83, 84] and their study originates in the study of the brain. The classic model of a travelling pulse in an excitable medium is the FitzHugh-Nagumo model [85] when used to describe action potential propagation along axons. This is a simplified model of action potential propagation along axons which describes membrane potential V in terms of a stimulus current I and a recovery variable W as follows:

$$\frac{dV}{dt} = V - \frac{V^3}{3} - W + I \tag{1.4}$$

$$\frac{dW}{dt} = 0.08(V + 0.7 - 0.8W) \quad (1.5)$$

When projected across the length of an axon, the cubic nonlinearity in equation 1.4 corresponds to strong positive feedback from rapidly activated channels that regenerates the action potential as it propagates, analogous to trees converted into heat as they are consumed by an expanding fire. In contrast, W has slower dynamics that rise as a function of V and suppress activity, preventing backward pulses by imposing a refractory period known as a “trailing shock” [83] during which events are suppressed. In the fire analogy, the effect of W is similar to the constraint that the fire wave is unable to reverse course because it consumes all available material in the areas it passes; trees must regrow before another wave of flame can pass. For a similar reason, if two pulses meet in excitable media, they often both cease to exist because neither can cross the trailing shock left by the passage of the other wave. These properties contrast with those of classical waves which can generally pass through each other and reverse course at boundaries [82]. In chapter 3, we show that due to the high degree of recurrent connectivity in cortex and refractoriness in neuronal firing, it may be appropriate to describe sheets of cortical neurons as excitable media capable of supporting multi-dimensional wave pulses.

INFORMATION PROCESSING WITH CORTICAL WAVES

Because waves translate activity occurring across space into activity distributed in time at fixed locations, they are ideal carriers of information. For example, humans commu-

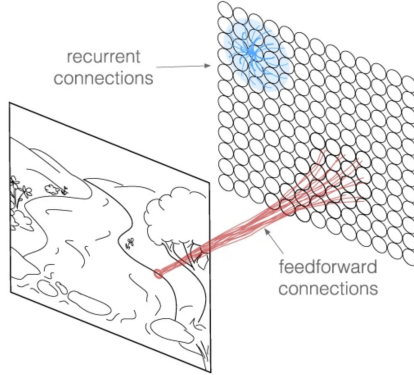
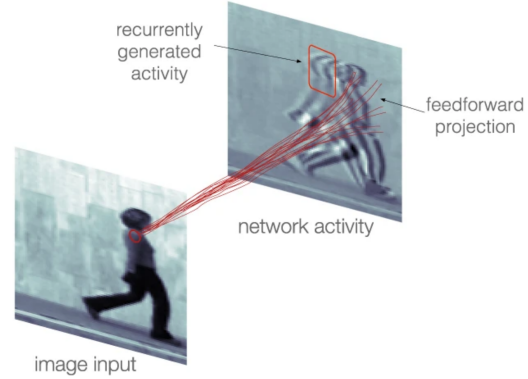
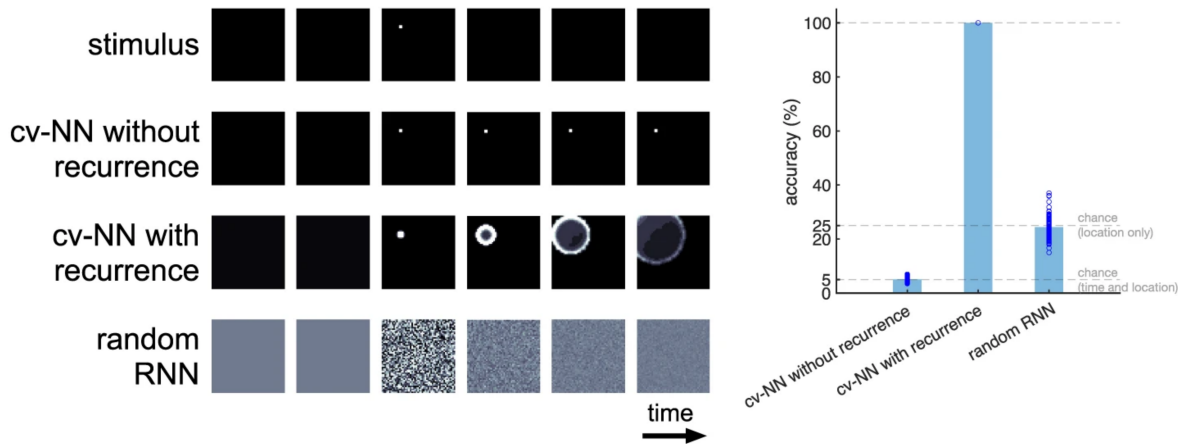
a**b****c**

Figure 1.1: Reproduced Topographic Propagation of Sensory Evoked Waves This figure is reproduced from [12] under a creative commons license and illustrates wave based horizontal propagation of activity across a topographically organized map of visual space. **a** Illustrates the organization of a recurrently connected artificial neural network (RNN) as a topographically structured set of visual receptive fields. Different fields receive input corresponding to different parts of the visual scene in order to mimic retinotopy, and are connected through lateral recurrent connectivity. **b** Shows an example of how the RNN generates intrinsic activity corresponding to an imprint of the visual scene despite restricted feedforward input. **c** Left, the stimulus (top) evokes responses that are shown for a topographic network without recurrence (second row), a topographic network with recurrence (third row), and an RNN without topographic structure. Notice the horizontally propagating circular wave in the topographically structured network with recurrence. Accuracy is shown for a linear classifier that was tasked with predicting the time or location of the stimulus event, the recurrently connected topographically organized network was the only one that performed well in the task.

nicate complex structured language through sound waves carried in air and, given the ubiquity of digital communications, it is highly likely that you, dear reader, received this thesis after it was transported in the form of electromagnetic wave pulses. Within cortex, travelling waves are thought to be evoked by sensory stimulation [15–19] and propagate laterally across topographically organized sensory areas. Fig. 1.1a, reproduced from [12], provides an illustration of one such topographically organized artificial neural network where feedforward sensory input is directed to specific spatially localized receptive fields. In that model, the receptive fields themselves are structured topographically so that their proximity in terms of connectivity matches the spatial proximity of the visual information they receive from the input scene. As shown in fig. 1.1c, the recurrent connectivity in the model was engineered in such a way that sensory input evoked circular travelling waves that propagated laterally and allowed for the integration of spatial information across distinct receptive fields.

In computational work these types of waves have been shown to permit recurrent artificial neural networks (RNNs) that otherwise lack a mechanism for sharing relevant information to solve tasks involving integration of complex spatially distributed information [14], temporally structured patterns requiring encoding of the recent past [13], and predictions of upcoming sensory sequences [12]. It has also been claimed that travelling waves can eliminate the requirement for error gradients to be propagated backward in time during artificial recurrent neural network training [86] by acting as a type of working memory that stores representations of recent events. Nonetheless, it remains unclear which synaptic conditions evoke travelling waves, how neurons in spiking networks could decode the

information encoded in travelling waves, and whether the precise temporal features of travelling waves can be leveraged for the coordination of biologically plausible learning rules [87–92]. We also lack a well understood experimental perturbation to precisely evoke and control cortical travelling waves which is a critical barrier limiting causal study of these potentially rich spatiotemporal dynamics. For these reasons, in chapter 3 we seek to develop a rigorous, experimentally validated, mechanistic model of the cortical response to electrical microstimulation, and in chapter 4 we present simulations comparing travelling waves evoked by this microstimulation to responses evoked by endogenous synaptic input.

PART I

SYNAPTIC DYNAMICS

Insofar as we think of mental processes as computational (hence as formal operations defined on representations) it will be natural to take the mind to be, inter alia, a kind of computer. That is, we will think of the mind as carrying out whatever symbol manipulations are constitutive of the hypothesized computational processes.

J. A. FODOR, METHODOLOGICAL SOLIPSISM
CONSIDERED AS A RESEARCH STRATEGY IN
COGNITIVE PSYCHOLOGY (1980)

FUNCTIONAL SUBTYPES OF SYNAPTIC DYNAMICS IN MOUSE AND HUMAN

ABSTRACT

Synapses preferentially respond to particular temporal patterns of activity with a large degree of heterogeneity that is informally or tacitly separated into classes. Yet, the precise number and properties of such classes is unclear. Do they exist on a continuum, and if so, when is it appropriate to divide that continuum into functional regions? In a large dataset of glutamatergic cortical connections, we perform model-based characterization to infer the number and characteristics of functionally distinct subtypes of synaptic dynamics. In rodent data, we find five clusters that partially converge with transgenic-associated subtypes. Strikingly, the application of the same clustering method in human data infers a

highly similar number of clusters, supportive of a stable clustering. This nuanced dictionary of functional subtypes shapes the heterogeneity of cortical synaptic dynamics and provides a lens into the basic motifs of information transmission in the brain.

2.1 INTRODUCTION

Studying how neurons communicate information provides a window into the fundamental language of the brain. This language is partly articulated by synapses that show preferential responses to particular temporal patterns of activity. This selectivity is enabled by transient (10 ms - 10 s) changes in connection strength termed short term plasticity (STP) [9–11]. STP is found in most brain regions and helps sculpt neural information processing in vivo [93, 94]. The most basic types of STP are short term facilitation (STF) and short term depression (STD), which entail a transient increase or decrease in synaptic strength, respectively [22–25, 95]. More complex forms of STP have been reported, including biphasic STP [50, 51], delayed burst-dependent STP [52], and sub- and supra-linear facilitation [49]. The prevalence and complexity of STP highlights the importance of understanding its role in shaping information transmission [35].

Short term plasticity has long been viewed as a mechanism to allow for communication of different messages to different postsynaptic targets from the same axon [26–29]. Depending on the timescale of STP, it supports detection of subtle changes in low versus high firing rates [30–32] or subtle changes in firing patterns such as bursts versus single spikes [31]. A number of other network-level functions have also been advanced: Depressing synapses

may serve to decorrelate inputs [38] while facilitating synapses can enhance detection of correlated signals [37]. STD can be seen as an adaptation mechanism [36], allowing for preservation of time-delay information across multiple levels of intensity [33–35]. STP has also been proposed to support working memory either by serving directly as its basis [41] or by acting at a network level to sculpt the emergence of persistent attractor dynamics [39]. In addition, STP may facilitate network-level storage and retrieval of memories [40]. In spike trains encoding a multiplexed code, STP properties can be used to decode independent streams of information [42–44]. Decoding multiplexed code is broadly relevant to biological learning because this type of multiplexed coding has been used to address the “credit assignment problem”, providing an account of how learning can be coordinated across multiple areas [44–46, 96]. Finally, the diversity of STP across individual synapses can form a reservoir of temporal patterns, which is beneficial to solving temporal credit assignment problems that involve learning temporal associations [47, 48]. These myriad of roles of STP underscore the importance of clearly understanding the functional subtypes of STP that exist in biology.

A major impediment to relating STP to its function is the large degree of response heterogeneity that exists across synapses [8, 97]. This heterogeneity is routinely and tacitly separated into broad classes such as facilitation and depression. A number of studies have established that, even when controlling for post-synaptic neuron subclass, response heterogeneity appears to be structured [8, 48, 97, 98]. While this heterogeneity arises from a combination of intrinsic variability, plasticity and genetically regulated synapse types, how these factors come together remains a matter of debate. Three hypotheses can be

proposed: 1) STP properties walk a continuum, with cell types associated with parts of the continuum, 2) STP properties are clustered by genetically defined functional subtypes, with intrinsic variability and plasticity acting in a small neighborhood around the cluster center, 3) STP properties cluster around genetically defined subtypes, but a continuum of STP properties between subtypes are allowed. A prediction of hypotheses 2 and 3 but not of hypothesis 1 is that functional subtypes should emerge from an unsupervised clustering of the dynamics of a large number of synaptic responses. An important question relating to such clusters is to what extent they might be stable across species, even if a perfect alignment is not expected given reported differences in human and rodent STP recovery timescales [8, 99] . Theoretically, clusters found by unsupervised methods should also contain features that are predictive of the genetic expression of cells on both sides of the synapse. These questions can be tested in a sufficiently large dataset and are the main focus of this article.

Here we leverage the Allen Institute Synaptic Physiology Dataset [8], which includes an unprecedented number of paired-patch recordings performed in physiological external calcium concentration. Crucially, this dataset provides enough synapses to meet the training requirements of data efficient machine learning techniques and open an avenue of fine grained analysis. We have applied a highly adaptable short-term plasticity model (referred to here as the Spike Response Plasticity (“SRP”) model) [2] to generate empirical fits descriptive of STP. Using unsupervised learning across the 224 excitatory connections (see Methods for selection criteria), we found 5-6 stable clusters of STP with a striking match between human and mouse. We then used the cell subclass labels provided by the

Allen institute for supervised analysis of STP properties and contrasted the features used for supervised (label aware) versus unsupervised (label free) learning. We find distinct functional STP subtypes more nuanced than the binary partition between facilitation and depression. These form biologically plausible building blocks of the information processing properties of synapses and were found to only partially align with the cellular subclass level of transcriptomic description as defined in Ref. [100].

2.2 RESULTS

Our analysis focused on glutamatergic cell connections in mouse and human cortex from the Allen Institute Synaptic Physiology Dataset [8]. In this publicly available dataset, paired-patch recordings were performed by stimulating the pre-synaptic neurons with diverse temporal patterns and recording the excitatory post-synaptic amplitude. The experiments were performed with genetically-identified cells for both pre- and post-synaptic partners via transgenic mouse lines. The genetic markers used in this dataset correspond to the subclass level of hierarchical clustering of transcriptomic data [100–102].

2.2.1 MODEL CAPTURES THE DIVERSITY OF SYNAPTIC DYNAMICS

In order to generate a biologically plausible dictionary of functional STP models, we fit instances of the Spike Response Plasticity (SRP) model from [2] to synaptic data from the Allen Institute Synaptic Physiology Dataset [8]. As shown in Figure 2.1A, the SRP

model predicts post-synaptic efficacy from an input spike train while incorporating short-term changes in synaptic strength. The SRP model works by convolving an input spike-train with a spike-response kernel, taking a nonlinear readout of the resulting trace at the time of the spikes, and then stochastically sampling synaptic efficacies according to this readout. The spike-response kernel encapsulates the trace left by every spike on the degree of facilitation/depression. It can combine facilitation and depression on different timescales (Fig. 2.1A, right). The parameters of this model regulate the shape and baseline of the spike-response kernel as well as the standard deviation of the stochastic sampling step.

In order to capture the differing dynamics present within the Allen Institute Synaptic Physiology Dataset, we used standard optimization methods (Simplicial Homology Global Optimization [103]; see methods) to fit distinct kernel parameters for every cell-pair. Figure 2.1B illustrates the range of stimulus protocols for which data were available. In all cases, eight stimuli were applied pre-synaptically at a fixed frequency, followed by a delay and four more fixed frequency stimuli. We chose to normalize all responses against the average of first responses in order to focus on the dynamics of the synapse and abstract away the amplitude which is in great part a cellular property. A range of stimulation frequencies and delay durations were experimentally examined for each synapse. Model parameters were optimized via maximum likelihood on the sequence of recorded post-synaptic amplitudes. The model captured both facilitating and depressing synapses, as shown in Fig. 2.1C for four exemplar pairs. We also found strong average fitting performance when we pooled protocols and pairs that corresponded to the same post-synaptic cell subclass (Fig. 2.1D). For each post-synaptic subclass, the mean squared error remained low (mean MSE and

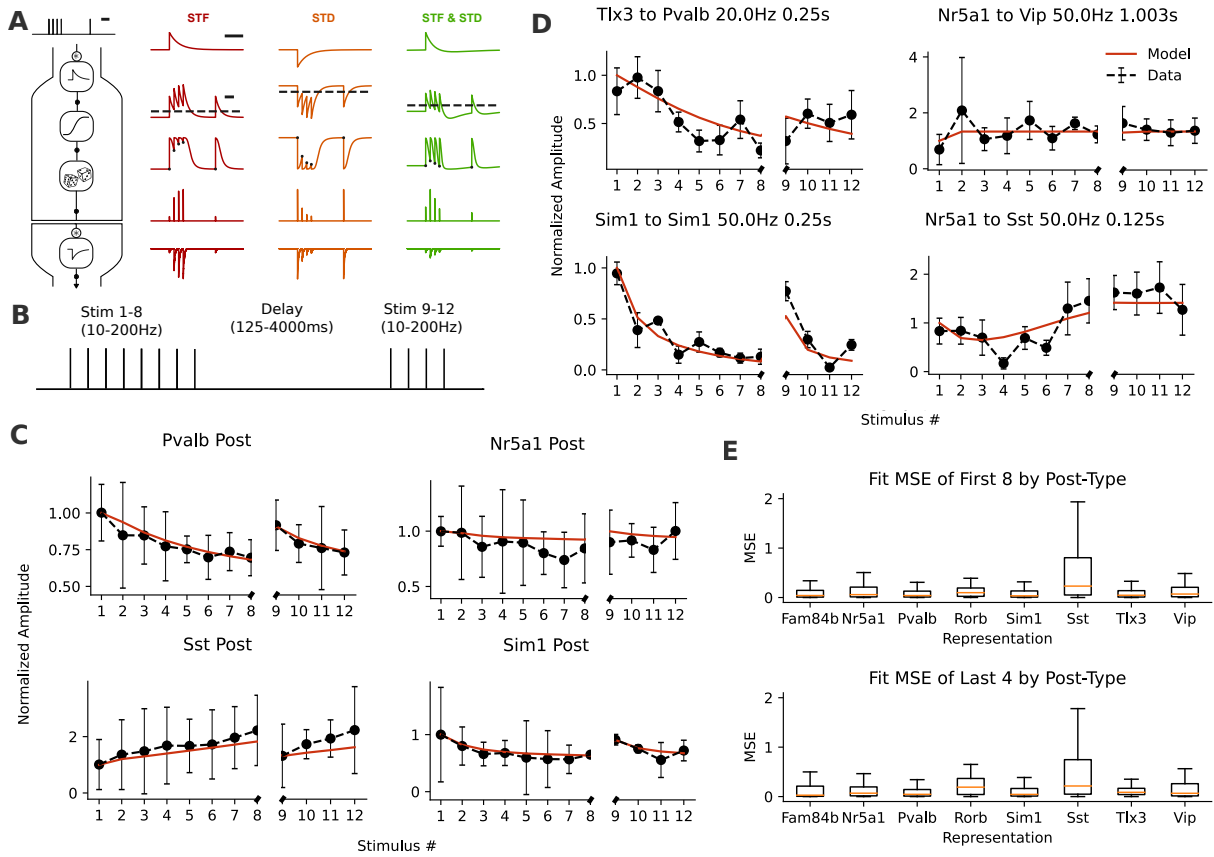


Figure 2.1: The Spike Response Plasticity (SRP) model captures responses across a diversity of synapse types. **A:** shows the SRP model of short-term synaptic dynamics from [2] which predicts excitatory post-synaptic potentials (EPSPs) based on spike timing input. **B:** shows the timing of input given to pre-synaptic cells in a paired patch clamp experiment, the model in A is fitted to predict EPSPs from this input. **C:** shows mean SRP model fits (red), and data (black) for all synapses onto specific post-synaptic cell types (29 onto Pvalb (2124 responses), 11 onto Nr5a1 (838 responses), 11 onto Sst (875 responses), 25 onto Sim1 (902 responses)) and demonstrates high overall fit quality. Error bars show standard error of the mean (SEM) across all pooled fits for the indicated subclass. **D:** Shows fits of example synapses between different transgenically defined subclasses for a particular protocol. The Tlx3 to Pvalb example corresponds to 20 Hz stimulation frequency and a 250 ms delay across 5 runs, the Nr5a1 to Vip example is for 50 Hz stimulation frequency and a 1003 ms delay over 5 runs, the Nr5a1 to Sst example is for 50 Hz stimulation frequency and a 125 ms delay over 8 runs, and the Sim1 to Sim1 example is for 50 Hz stimulation frequency and a 250 ms delay over 4 runs. Error bars show standard error of the mean (SEM) for the specific protocol and pair. **E:** shows mean squared error of fit in units of normalized amplitude, to individual runs for the first eight stimuli before a delay (top) and the four stimuli after a delay (bottom), grouped by postsynaptic cell subclass. The orange line indicates the median mean squared error, the box boundaries show interquartile range, outlier values beyond the whiskers are not plotted. Because not all pairs were tested for the same number of runs, the fits are across the following numbers of responses for first 8 and last 4 runs respectively: Fam84b: 108, 55, Nr5a1: 838, 421, Pvalb: 2124, 1060, Rorb: 72, 36 Sim1: 902, 453 Sst: 875, 438 Tlx3: 144, 72 Vip: 267, 134

standard error of the mean of 0.38 ± 0.02 and 0.47 ± 0.06 for the first eight stimuli and the last four stimuli across all post-subclasses, respectively; Fig. 2.1E). Synapses onto somatostatin (Sst) expressing cells were notable for having a higher proportion of poorly fitting outliers. Specifically, the mean MSE and standard error of the means of all synapses onto somatostatin expressing cells was 1.4 ± 0.1 ($p < 0.001$) for the eight pre-delay simulations and 1.9 ± 0.4 ($p < 0.001$) for the four post-delay simulations. Because of the strong average fitting across all other subclasses and good fitting performance for non-outlier Sst synapses, the model parameters can be used as a compressed representation of the dynamics [104].

2.2.2 UNSUPERVISED CLUSTERING REVEALS FUNCTIONAL SUBTYPES

In order to develop insights about the functional subtypes of STP in rodents, we used unsupervised machine learning techniques to perform clustering (Fig. 2.2A). Clustering allows us to group synaptic dynamics based solely on the relative distance of synapses in a projected space. We selected OPTICS (Ordering Points To Identify the Clustering Structure) [105] as a clustering algorithm. OPTICS defines clusters given a parameter for the minimum cluster size (in number of elements per cluster). This parameter indirectly affects the number of clusters. In order to determine the optimal minimum cluster size parameter for the OPTICS algorithm, we tested a range of minimum cluster sizes and computed an indicator of clustering quality (the mean silhouette score, see Methods). The highest silhouette scores corresponded to minimum cluster sizes of 6, 7, and 8 in rodents as shown in Figure 2.2B) from which we selected a minimum cluster size of 8. With this choice

of parameter, OPTICS identified five clusters (abbreviated as 'R0-R5'), corresponding to five subtypes of synaptic dynamics.

The main characteristics of these functional subtypes can be seen by inspecting the shape of the SRP kernel, its baseline, and the parameter regulating the standard deviation (SD) of responses (all of which are used for clustering). Mean SRP kernels of each rodent cluster (Fig. 2.2C) and SD (Fig. 2.2D) revealed qualitatively distinct dynamics. Clusters 'R3' and 'R4' were characterized by kernels with large positive amplitude, which implements facilitation. These two clusters can be distinguished by considering the timescale of facilitation, which is faster for cluster 'R4' and the baseline parameter which is high for 'R3' limiting the magnitude of dynamic changes. The baseline parameter controls the response strength relative to its maximal and minimal values, such that a zero baseline enforces equidistance and a higher baseline brings the upper bound closer to the initial state. Clusters 'R1' and 'R2' are characterized by low but distinct baselines combined with kernels that have negative amplitudes and are associated with depressing synapses. These two clusters can be further distinguished by the SD parameter, as cluster 'R2' was associated with the smallest degree of variability. Cluster 'R0' displayed a median kernel that is bimodal, which produces a slow and gradual form of facilitation, and can even produce depression for very fast stimulation.

To visualize how these clusters relate with transgenically defined subclasses, we first plotted the model parameters for each pair projected into the space of their first principal components (Fig. 2.2E). As expected, each cluster occupied a restricted subregion of the space defined by the first two principal components. The median kernels and SD

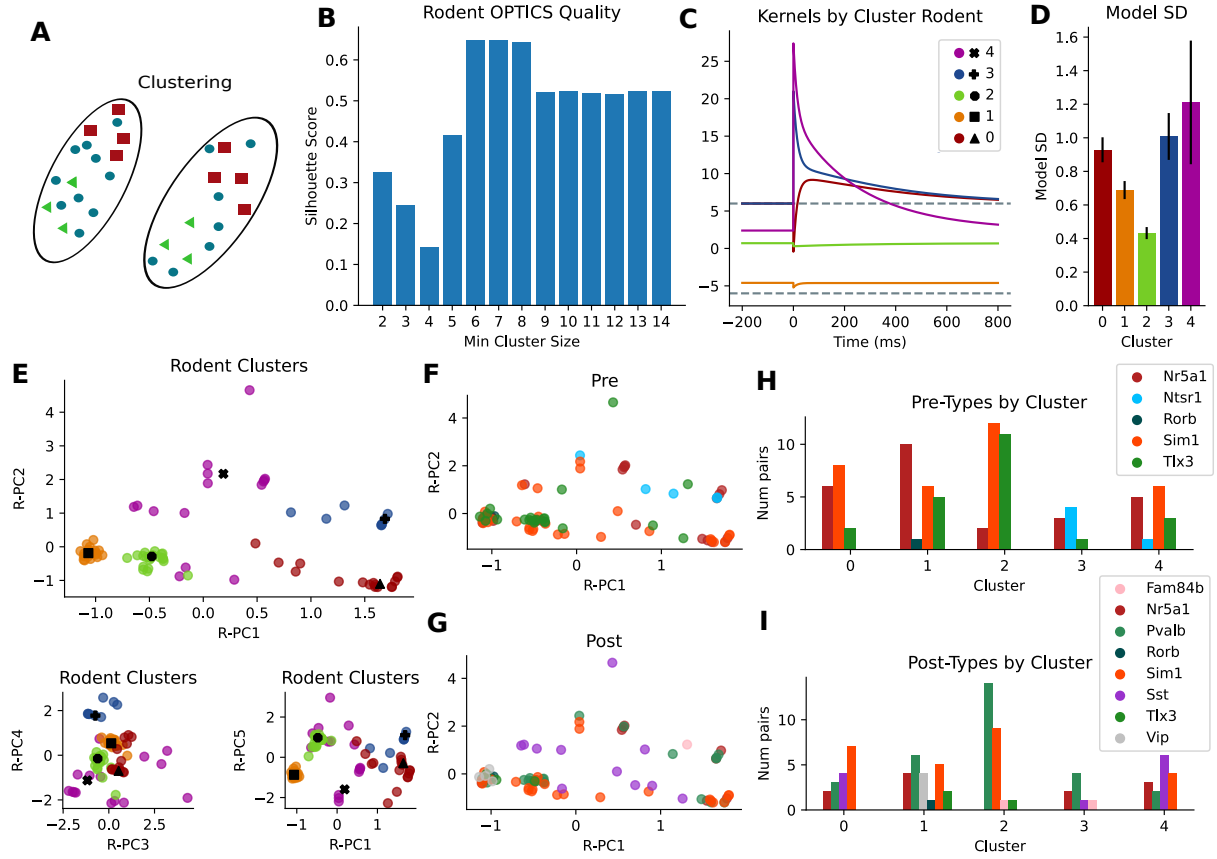


Figure 2.2: Unsupervised clustering reveals 5 clusters in rodent data. **A:** visual schematics of the process of unsupervised clustering. Synapses are grouped together based on model parameter similarity without taking labels such as cell subclass into account. **B:** silhouette coefficient, a measure of clustering quality, is shown against minimum cluster size. We selected a minimum cluster size of 8 for all rodent clustering. **C:** median kernels of each cluster are shown for each cluster based on the clustering of 86 glutamatergic synapses, median values of all parameters of each cluster are shown in table S1. Grey lines indicate the saturation values of the nonlinear readout in the SRP model. **D:** using the same color code as in C, model SD is shown for each cluster. This is the average of SDs from each fit (estimation error) which we also use as a metric to capture variability in synaptic efficacy on a synapse by synapse basis. Error bars show standard error of the mean (SEM) for clusters with 15, 22, 25, 8, and 15 connections respectively (ordered from R0 to R4). **E:** clusters generated by the OPTICS algorithm are plotted in different principal components of rodent data (R-PCs, where R-PC1 is the first principal component) in colours corresponding to C-D. Black markers show the location in PC space of the median kernels shown in C. **F-G:** in the same low-dimensional space as E (top), each dot represents a synapse coloured according to pre-synaptic subclass (F) or post-synaptic subclass (G). **H-I:** the number of synapses of each pre-synaptic subclass (H) and each post-synaptic subclass (I) is shown for each cluster with colours matched to C-E (see Fig. 2.7 for Pre-Post matrices and Fig. 2.8 for example fits by cluster).

parameters of each colour coded cluster are shown in Fig. 2.2C-D. In the same space, we then labeled the pairs according to the pre-synaptic (Fig. 2.2F) or the post-synaptic (Fig. 2.2G) cell subclass. All clusters contained a mixture of cell subclasses both pre- and post-synaptically. This is consistent with the large heterogeneity of synaptic dynamics associated with transgenically defined subclasses [8, 98] (Fig. 2.6). Counting the number of pre-synaptic (Fig. 2.2H) or post-synaptic (Fig. 2.2I) subclasses found within each cluster shows that cell types tend to belong to specific functional subtypes. For instance, post-synaptic Pvalb tends to belong to cluster ‘R2’ (14 of the 29 connections) and cluster ‘R1’ (6 of the 29 connections), the two depressing subtypes. As another example, post-synaptic Sst tends to belong to cluster ‘R0’ (4 of the 11 connections) and cluster ‘R4’ (6 of the 11 connections), two facilitating subtypes. These observations are therefore consistent with the previous experimental reports of facilitating connections onto Sst and depressing connections on Pvalb-positive cells [27–29]. Together, our results support a view where STP subtypes distribute in all cellular subclasses, but in subclass-specific proportions [100].

2.2.3 FUNCTIONAL SUBTYPES CANNOT BE EXPLAINED BY CONTINUUM HYPOTHESES

Next, we sought to determine whether the appearance of clusters in parameter space is an artifact of our methodology. One such artifact emerges from the OPTICS clustering algorithm as it is designed to find structure even in uniform data, particularly when the number of data points is low. To this end, we generated 3000 batches of 86 model synapses

by sampling parameters from a uniform distribution constrained by the maximum and minimum parameters observed using the data (Fig 2.3A-C). We then applied our OPTICS clustering procedure on each batch of surrogate parameters. Depending on the batch, the methodology found both low and high numbers of clusters (Fig. 2.3B). Calculating the silhouette coefficients for each batch, we found that a uniform distribution of model parameters leads to coefficients distributed between 0 and 0.2 (Fig. 2.3C). This distribution contrasted with the peak Silhouette Coefficients we found in rodent (green line). Thus, a uniform distribution of model parameters could not explain the observed clustering.

We reasoned that model parameters are not uniformly distributed, but that they distribute more evenly along the first principal component of the data. This gives a natural axis for the clustering algorithm to parcel the data, which may in turn artificially inflate the silhouette coefficient. To test whether our observation of separate clusters along this axis can be explained by a continuum model, we generated 3000 batches of surrogate model parameters by sampling from a multivariate Gaussian distribution matched to the statistics of the biological data. By design, these artificially generated parameters span the entire range of STP dynamics treated in this work but do not correspond to distinct underlying groups. We applied PCA and unsupervised clustering using the exact same procedure we had applied to the biological synapses and analyzed the clusters produced. Fig. 2.3F-G again shows scatter plots for two examples. As quantified by the comparison between silhouette coefficients arising from the continuum as those observed in the data, the continuum model could not explain the emergence of distinct clusters across 3000 batches.

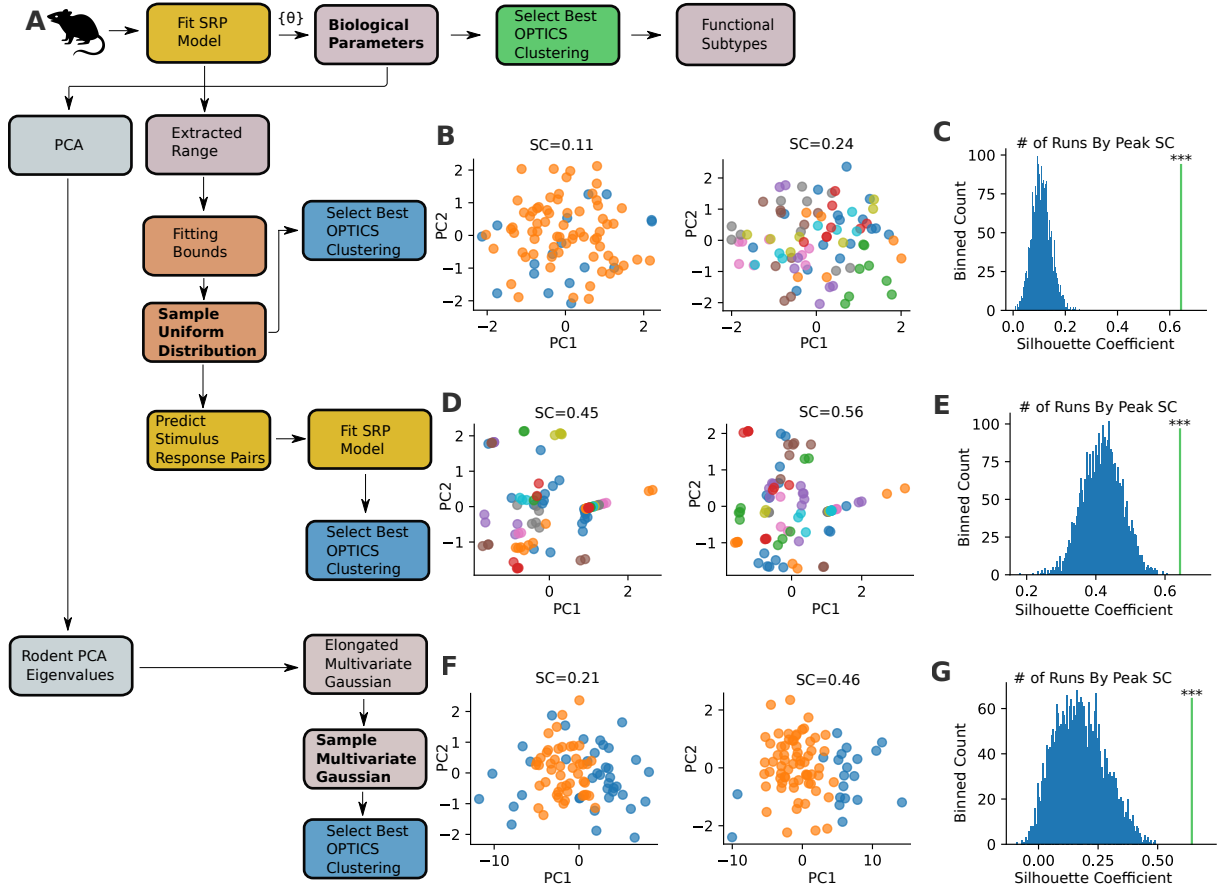


Figure 2.3: Testing uniform and continuum hypotheses in rodent. **A** visual schematic showing the procedure we used to characterize functional subtypes in biological data (first row). The range of fitted model parameters and the output of the PCA are used to generate our null hypotheses. For the uniform hypothesis, we sampled from a uniform distribution using the extracted range (second row). For the continuum hypothesis, we sampled from a multivariate Gaussian distribution whose variance structure matched the variance structure of the observations (third row). Each hypothesis was sampled 3000 times. Further testing the effect of applying our SRP fitting procedure to output simulated from the models sampled from the uniform distribution (fourth row). **B-G** In each row, two exemplar scatter plots (**B**, **D**, **F**) show the result of clustering (different colors) in the space spanned by the first two principal component dimensions. The overall Silhouette Coefficient distributions associated with each hypothesis are shown in the right (**C**, **E**, **G**) with the observed Silhouette Coefficient shown with a green vertical line. Triple stars indicate that the value observed in biology occurred less than once in 1000 random samplings of the corresponding hypothesis ($P < 0.001$, Bootstrap resampling).

Next, we considered that the fitting procedure for the SRP model could generate artificial clusters. For instance, parameters may be pushed against hard bounds during the inference procedure, creating a cluster where none existed. To test whether the clusters we observed in biology could be explained by combining uniformly distributed parameters and the fitting procedure, we used the sets of parameters sampled from a uniform distribution and simulated output of the stochastic SRP model to produce surrogate traces of membrane potentials. This surrogate data matched the collection of experimental protocols associated with the Allen Institute Synaptic Physiology Dataset. Next, we used our SRP fitting procedure to re-extract model parameters from this surrogate data. The clustering procedure was then applied to the re-fitted parameter set. Fig. 2.3D shows two example batches where some artificial clusters are visible along parameter bounds. We also noted an increase in elongation along the first eigenvalue (1.33 in uniform model parameters, 1.85 in re-fitted parameters). Further, we found that this procedure resulted in a much higher number of clusters (with a large number of batches between 23 and 31 clusters) than what we found in real rodent and human data and resulted in a shift towards more facilitating values. Nonetheless, the Silhouette Coefficients were significantly below those observed in the biological data. In summary, we found that the clusters we observed in biology could not be explained by a uniform distribution, nor by a continuum distribution, nor by artifacts of fitting the SRP model when these tests were constrained by the biological parameter range and variance structure.

2.2.4 CONSERVED FUNCTIONAL SUBTYPES ARE FOUND IN HUMAN SYNAPSES

A question of salience to this work is whether clusters emerging from our unsupervised process are consistent between species (Fig. 2.4A). We repeated the fitting and clustering procedure on excitatory (E-E and E-I, see methods for selection criteria) human data from the Allen Institute Synaptic Physiology Dataset. The minimum cluster size was selected to match the highest Silhouette Coefficient (minimum cluster size of 10, Fig. 2.4B). With this clustering parameter, OPTICS found 6 clusters associated with a Silhouette Coefficient of 0.71. We again verified that these 6 clusters could not be explained by model fitting, uniform models, or continuum models (Fig. 2.9).

We asked whether a similar optimal cluster number should be expected in independent datasets collected from distinct species. Figure 2.4C shows the distribution of the optimal cluster numbers when assuming uniformly distributed parameters that underwent re-fitting (Fig. 2.3D). We overlay the optimal cluster numbers for rodent (green) and human (purple), illustrating that the number of clusters we observed are unlikely to occur under this null hypothesis. Moreover, we calculated the probability of finding a clustering with an optimal cluster number between 5 and 7 immediately after finding a clustering with an optimal value of 6 and found it to be highly unlikely under our refit continuum hypothesis ($p < 0.001$). Thus, even if a perfect overlap is not expected [99], this convergence of the number of clusters across the two species hints at similar functional organisation of STP between species and strongly supports the existence of functional subtypes of STP.

Given that clustering in humans and rodents yields similar numbers of clusters, the

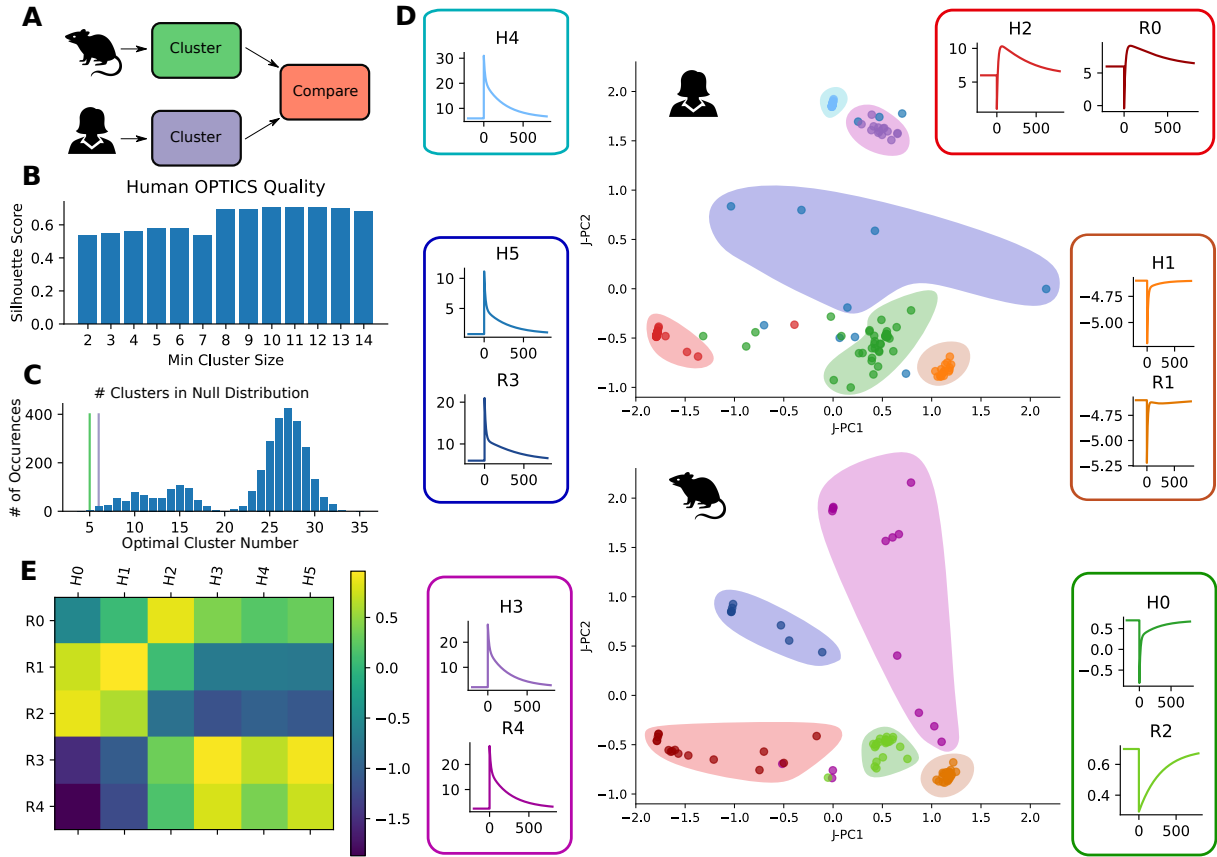


Figure 2.4: Unsupervised clustering shows functional alignment between rodent and human data. **A:** visual schematic demonstrating that the same clustering protocol is applied to both human and rodent synapses and groups of functionally similar synapses are compared without using references to pre-defined types. **B:** as in Fig. 2.2B, different minimum cluster sizes are compared for human data based on Silhouette coefficient. The peak Silhouette coefficient value corresponds to a minimum cluster size of 10 which was used for all human clustering. **C:** plot showing the number of clusters associated with the best Silhouette Coefficient after fitting the SRP model to data drawn from a uniform distribution (see also: fig 2.3D). Given the tendency in the distribution towards higher optimal cluster numbers, the chance occurrence of a clustering with an optimal value of 6 followed by another clustering with an optimal cluster number different by no more than 1 (i.e between 5 and 7) is very highly unlikely ($P < 0.001$, null hypothesis matches fig 2.3D). **D:** scatter plot showing human (139 pairs) and rodent (86 pairs) clusters in the first jointly defined two principal component dimensions (J-PCs). The closest matches between human (ex: 'H2') and rodent (ex: 'R0') clusters are indicated with similarly coloured clouds and show strong qualitative alignment across species for all but one cluster. Corresponding median kernels are also presented and median values for all parameters of each human cluster are shown in table S2. **E:** the similarity measure between the parameters of rodent and human clusters used to assign the closest matches is shown as a heat map (see Methods).

next question is whether these clusters display functional overlap. We illustrate the similarity between clusters in Figure 2.4D where the clusters of both species are plotted onto a set of principal components defined on the joint cross-species dataset (J-PC). The primary areas occupied by individual clusters are highlighted with clouds that are colour matched to closely corresponding clusters from the other species. Median kernels of the most similar clusters are shown next to each other, suggesting similar subtypes in the different species. While one more cluster was found in the human synapses, this manifests itself through two clusters with very similar facilitating dynamics ('H4' and 'H5' in Fig. 2.4D). While relatively small, other notable differences exist, including stronger, but shorter lived depression in human cluster 'H0' than in its corresponding rodent cluster ('R2'). To quantify the level of consistency between clusters, we computed a similarity measure incorporating a weighted combination of the temporal correlation of median kernels of each cluster and the difference in the modelled variance of their constituent synapses (Fig. 2.4E). In sum, although the exact number of subtypes may not be five (see Discussion), human synapses display similar functional modes to those found in the rodent data.

2.2.5 SUPERVISED PREDICTION OF TRANSGENIC IDENTITY RELIES ON SIMILAR FUNCTIONAL SUBTYPES

Having demonstrated that neither the pre- nor the post-synaptic cell subclass aligns perfectly with one particular functional subtype, we wondered whether features of synaptic dynamics could predict molecular subclass. Despite imperfect alignment, it is still possi-

ble that the preponderance of specific functional subtypes within specific cell subclasses delineates such genetically-defined categories. To test this, we used supervised learning to predict cell subclass labels using only synaptic dynamics data (Fig 2.5A). We compared various machine learning algorithms as shown in Figure 2.12 and selected a multi-layer perceptron (MLP) with three hidden layers because it displayed the highest overall performance. The MLP was trained to predict cell subclass either from a set of physiological features (e.g. paired-pulse ratio, release probability, etc. See Methods) or from the SRP model parameters. In both cases, when predicting post-synaptic subclass, cross-validated accuracy was better than a ‘baseline’ classifier (31.76%) which always predicted the most frequent label in the training set (Fig. 2.5B) but this difference only reached significance in the case of the model representation (model accuracy 45.14%, $P < 0.001$, phys accuracy 33.71%, $P = 0.297$, Bootstrap test for difference to baseline prediction). When predicting pre-synaptic subclass, we found no significant difference with baseline prediction (36.16%) either using model parameters (41.36%, $P = 0.136$, Bootstrap test for difference with baseline) or using physiological features (33.71%, accuracy below baseline, Fig. 2.5B). Directly comparing classification accuracy using model parameters against using physiological features did not show a significant difference either in pre-synaptic ($P = 0.133$, Bootstrap test) or post-synaptic ($P = 0.074$) subclass prediction. In summary, synaptic dynamics support non-trivial prediction with respect to post-synaptic but not pre-synaptic subclass.

We next trained an MLP on physiological features and SRP model parameters of data obtained under a higher external calcium concentration (2 mM). In this case it was not possible for the MLP to outperform the baseline (Fig. 2.5C). However, there were more

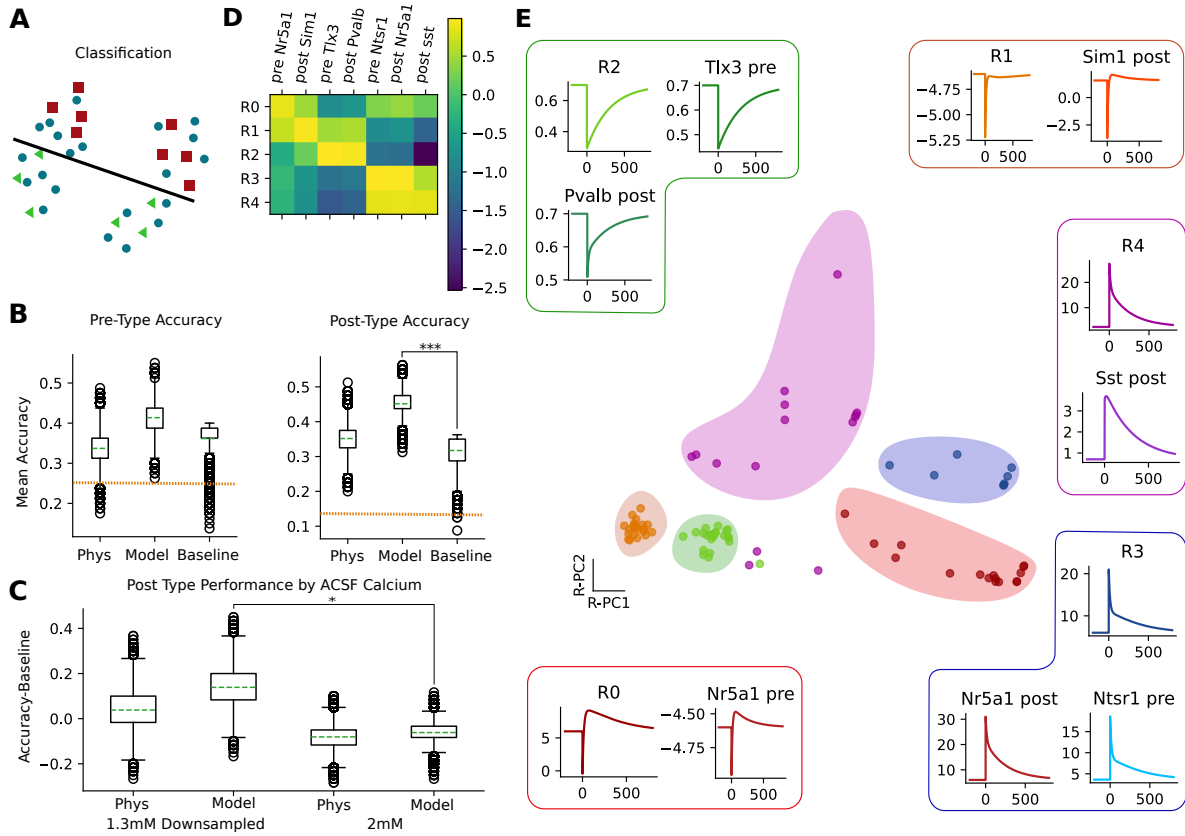


Figure 2.5: Clusters correspond to model features that are helpful for predicting pre- or post-synaptic cell identity. **A:** a visual schematic for illustrating that supervised learning separates data by optimally separating known subclasses, potentially at the expense of how the data is clustered. **B:** mean cross-validation accuracy for predicting pre-type (left) and post-type (right) is shown for three models. The green line indicates mean predictive accuracy across Bootstrapped shuffles, the box boundaries show interquartile range, and outlier values beyond the whiskers are plotted as circles. The dotted orange line indicates a “naive” estimate of chance level accuracy obtained by assuming all predicted classes occur with the same probability ($\frac{1}{4}$ for pre-type, $\frac{1}{7}$ for post-type). The ‘baseline’ model uses the most frequently occurring label as the prediction and therefore captures the extent to which predictive accuracy can be explained by class imbalances. The ‘Phys’ uses a multi-layer perceptron (MLP) with three hidden layers trained on physiological features. This classifier showed the best predictive performance in an out-of-sample classifier comparison shown in Fig. 2.12, Fig.S8 shows the same comparison for within-simple (i.e. previously seen) data. The ‘Model’ used an MLP with the same configuration trained on model parameters. For pre-type prediction (left), the mean accuracy of the ‘Baseline’ is not outperformed by the ‘Model’ ($P=0.136$, Bootstrap resampling) nor ‘Phys’ ($P=0.764$). For post-type prediction (left), the mean accuracy of the ‘Baseline’ is outperformed by the ‘Model’ ($P<0.001$, Bootstrap resampling) but not ‘Phys’ ($P=0.297$, Bootstrap resampling). All comparisons were based on 1,000,000 Bootstrap samples. **C:** shows the difference in accuracy above baseline in post-synaptic subclass prediction based on data collected in 1.3 mM vs 2 mM. In this case the 1.3 mM dataset is downsampled to match the number of 2 mM exemplars remaining (67 after our data processing and quality controls). The ‘Model’ representation trained on 1.3 mM data outperformed an equivalent representation in 2 mM data ($P=0.012$, Bootstrap resampling) and the difference in performance between physiology representation in 1.3 mM and 2 mM trended towards significance ($P=0.107$, bootstrap resampling). This comparisons were made based on 100,000 Bootstrap samples. Histograms of differences resulting from the Bootstrapping procedure used to generate these p-values are shown in 2.11. **D:** shows a similarity measure (see Methods) between unsupervised clusters and groups predicted by supervised learning incorporating temporal correlations between median kernels and the difference of mean variance parameters. These measures demonstrate alignment between functionally defined STP subtypes and the features relevant to the prediction of cell subclasses. Only the predicted subclasses with the highest measure of similarity to a rodent cluster are shown. **E:** scatter plot of model parameters with colouring based on cluster identity (as in Fig. 2.2E). Insets highlight the qualitative similarity of median kernels of the different cell subclasses and functionally defined classes shown in **D**.

data available in physiological conditions (1.3 mM calcium) than in the higher calcium condition (2 mM), so we controlled for the effect of dataset size by downsampling the amount of physiological calcium data to match that in higher calcium. Even after downsampling it was still possible to predict post-synaptic subclass with higher accuracy than the baseline ($P=0.040$, difference from baseline, Bootstrap test). We further controlled for the effect of different baseline prediction in high and physiological conditions by directly computing the differences between model-based and baseline prediction accuracy in both conditions. We found that classifiers based on physiological calcium data outperformed those trained on higher external calcium concentration data when predicting post-synaptic subclass ($P=0.012$, 100,000 Bootstrap shuffle for each, Fig. 2.10). In all cases the strongest predictive power was attained by combining our model-based representation with data collected in physiological calcium conditions. This suggests that some predictive features present under physiological conditions were not present in the presence of higher external calcium concentration.

We next compared the model features exploited by the MLP to the clusters obtained using the OPTICS algorithm. Specifically, the median kernel parameters and mean variance parameters of those synapses predicted by the MLP to belong to a particular subclass reveal features that may have been used by the machine learning algorithm to perform the classification. Similarity measures computed between such a subclass stereotype and the representative parameters associated to a given cluster ('R1', 'R2', ...) show selective alignment between cell subclass and clusters (Fig. 2.5D). Each rodent cluster has strong correspondence with at least one specific pre- or post- synaptic subclass. In Figure 2.5E, we

illustrate the correspondence between label-aware and label-agnostic groupings by pairing the median kernels of the most similar groups based on the similarity measures shown in Figure 2.5D. Paired median kernels are shown in distinct panels that are colour matched to corresponding clusters in the space defined by the first two principal components of rodent SRP model fits. Together, our findings suggest that despite their heterogeneity functional subtypes are characterized by similar sets of dynamical features as those used to predict cellular subclasses. These findings are consistent with subtypes of synaptic dynamics that are only partially aligned with the cell subclasses.

2.3 DISCUSSION

This work uses machine learning techniques and mathematical modelling to infer distinct functional subtypes from the heterogeneity of cortical short-term plasticity in both humans and rodents. We consistently observed in both humans and rodents that synaptic dynamics formed five to six subtypes. These clusters were genetically heterogenous yet they retain an alignment with transcriptomics that was sufficient to provide predictive power with regards to cellular subclasses. We have validated that these functional subtypes are not simply the product of noise by deliberately testing our classification in out of sample data, controlling for the effect of model fitting on clustering, and testing against continuum and uniform models. Our results also indicate that the dynamics of synapses belonging to different cell subclasses become less distinctive in data collected in an external calcium concentration above physiological levels (i.e. 2 mM Ca).

Ultimately, are there really precisely 5 functional subclasses of STP? Like all clustering approaches, our work is limited by the amount of data available, with more data typically enabling reliable identification of smaller clusters [101, 102, 106]. Furthermore, we have shown that both the clustering procedure and the model fitting process will report clusters out of uniformly distributed data. Therefore there still remains some uncertainty regarding the exact number of STP subtypes. What appears clear is that, consistent with previous work identifying multiple types of STD [48, 107, 108], the dictionary of functional STP subtypes must extend further than the simple division of STD and STF.

What does the existence of our 5 to 6 subtypes change in our understanding of STP? A diversity of facilitating types suggests the possibility of multiple types of high frequency coding schemes. Moreover, when these facilitating types are paired with the bimodal STP cluster we observed in both species, it hints at the possibility of decoding multi-level multiplexing schemes [43] wherein three or more different channels of information are encoded into a combination of long bursts, short bursts, and isolated events. Our STD subtypes were largely differentiated on the basis of variability which may indicate a difference in release probability or asynchronous release [109]. This may indicate a difference in the number of post-synaptic contacts which have been hypothesized to relate inversely to release probability as part of a mechanism to conserve energy at the synapse [110]. We provide the full set of our fitted SRP model parameters labelled by group and by transgenically defined pre-synaptic and post-synaptic subclass (see data and code availability) as a resource to aid the community in understanding the implications of these subtypes.

What determines if a connection belongs to a given functional subtype? We, like others [8], have observed that there is likely a transcriptomic basis for these subtypes, as clusters are aligned to features predictive for the transgenically defined subclass of either pre- or postsynaptic neuron. Given, however, the weak predictive power and the observation that each subclass has connections that belong to almost every cluster, we consider it likely that stronger correspondence between cell identities and functional dynamics will emerge from finer grain cell identity, such as at the type or sub-type level [100, 111]. Factors such as model choice, the number of categories the classifier must distinguish between, and the number of examples from each category can strongly influence these results. Relatedly, other groups using distinct models and datasets have attempted to predict cellular subclass based on STP properties [112, 113], but direct comparison is complicated by different reporting standards, different numbers of categories, and in particular the absence of clear comparison with a class balance aware baseline classifier. If no accurate, finer-grained transcriptomic basis is found, there remains the explanation that plasticity of STP properties is an important factor to explain their diversity [114].

These findings highlight several key takeaways for future experimental work. Since the most salient feature separating the STF subtypes is the timescale of facilitation, our results highlight the need for continued collection of cell-pair specific data containing stimulation protocols conducive to separating longer timescales. Relatedly, our supervised learning methods show better separation in near physiological external calcium concentrations (Fig. 2.5C), supporting its importance as an experimental parameter [49, 87, 115]. In this case we theorize that higher external calcium concentrations partially homogenize STP,

obfuscating meaningful cell identity specific dynamics. Lastly, our results support using a model-based representation instead of selected features of the recordings (Fig. 2.5B). Future experimental data can be integrated with our publicly available model parameters and model inference algorithm to help provide deeper insight about the dynamical nature of brain connectivity.

2.3.1 LIMITATIONS OF THE STUDY

Like any study using unsupervised learning our methodological choices entail certain caveats. Firstly, as discussed in our testing of continuum hypotheses clustering methods will attempt to generate clusters by design, regardless of whether the data are inherently structured as groups. We mitigate this concern in several ways: firstly we rely on PCA as our primary dimensionality reduction method, this allows us to preserve distance relationships through a linear transformation rather than applying a more complex transformation (ex: T-SNE, UMAP [116, 117]) that has the potential to non-linearly modify relative distance relationships. We also directly test against two specific models that do not have clustered STP dynamics. Nonetheless, it remains possible that the heterogeneity of STP can be explained by some more complex continuum that we have not tested. Our findings are limited to showing that if such a continuum exists it is not a uniform or multivariate Gaussian defined in a range consistent with the data we have examined.

Another potential drawback of our clustering pipeline is the reliance on high Silhouette Score as a method for the selection of the “minimum cluster size” OPTICS meta-parameter.

This is notably a limitation in the case of rodent data where three minimum size values can be considered strong (6, 7 and, 8) but minimum values of 6 or 7 result in different labeling of a small subset of the dataset from the value we selected (8). The difference we observe between the two parameterizations does not change the results of our testing against continuum hypotheses (i.e. fig 2.3), our cross species finding of 5-6 clusters (though values of 6 or 7 yield 6 clusters instead of 5), or our observation of similar cross species clustering (the difference is not in the well-aligned R0-H2, R1-H1, and R2-H0 clusters but rather is the more dispersed, more facilitating portion of the J-P2C axis, see fig 2.4). We also expect greater ambiguity in the alignment of facilitating types because a greater share of facilitating kernels are thresholded by our fitting bounds (see ‘data and code availability’ for all fitted values). This is because we use a single tuned version of our fitting procedure to capture data for all synapses, rather than tuning meta-parameters on a synapse by synapse basis. We deliberately chose meta-parameters that would properly capture dynamics in facilitating synapses without compromising fit quality in depressing synapses (ex: fig 2.1, fig 2.1C). Part of how we implemented this facilitating bias is an asymmetry in the fitting bounds for the baseline parameters of our models which ensures that the mean predicted response of our modelled synapses to the first stimulus in a train is strictly positive (though it can be small). In addition to permitting better fitting across types this bias ensures that our models are compatible with simulations that rely on methods which require a defined error gradient (e.g. surrogate back propagation [76]). There do exist a variety of other STP models [8, 118], and other formulations of the SRP model [2] which are likely to capture the dynamics in of these connections with different biases. For example assigning more

complexity (parameters) to the modelling of variability could generate a clearer picture of the difference in types separated on the basis of variability, and perhaps permit resolution of other distinct subtypes at the price of greater fitting complexity and associated data requirements.

2.4 METHODS

2.4.1 RESOURCE AVAILABILITY

LEAD CONTACT

Further information and requests should be directed to and will be fulfilled by the lead contact, Richard Naud (rnaud@uottawa.ca)

MATERIALS AVAILABILITY

This study did not generate new unique reagents.

DATA AND CODE AVAILABILITY

- This paper analyzes existing, publicly available data. The accession numbers for the datasets are listed in the key resources table. We have also packaged individual SRP fits by connection for each cluster with our code (see below).

- All original code is archived on Zenodo and available on Github as of the date of publication. Repository information is listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

2.4.2 METHOD DETAILS

2.4.3 SPIKE RESPONSE PLASTICITY (SRP) MODEL

We modelled synaptic dynamics using the following flexible model from [2] which is defined as follows:

Assume each pre-synaptic spike train gives rise to a post-synaptic current trace $I(t)$ which corresponds to the sum of post-synaptic currents each triggered at a pre-synaptic time t_j :

$$I(t) = \sum_j \mu_j k_{PSC}(t - t_j), \quad (2.1)$$

Where μ_j is the response to the j th spike in a normalized spike train and k_{PSC} is a kernel that describes a stereotyped PSC time course. We then defined an efficacy train $E(t)$ as the product of a pre-synaptic spike train $S(t)$ with a time varying signal $\mu(t)$ that represents the synaptic efficacy at a given time:

$$E(t) = \mu(t)S(t). \quad (2.2)$$

This allows us to rewrite the current trace as a convolution $I = \mathbf{k}_{PSC} * E$ treating $\mathbf{k}_{PSC}$ as a kernel corresponding to the stereotyped PSC time course. We then model the efficacies μ_j as the output of a nonlinear readout of a linear filtering:

$$\mu = \frac{1}{f(b)} f(\mathbf{k}_\mu * S + b) \quad (2.3)$$

For this work, we select the kernel k_μ to be a sum of three exponential functions with time constants fixed to 15 ms, 200 ms, and 300 ms. These time constants were hand selected to permit good fitting (as defined by mean-squared error) of synapses displaying both synaptic facilitation and synaptic depression. The amplitudes of these functions and of the baseline parameter b are then free to be fitted for individual synapses providing a compact deterministic description of the dynamics of each synapse. We further describe the stochastic dynamics of the synapse by considering a sample of synaptic efficacies as a random variable Y_j with a mean value corresponding to the output of the deterministic description provided above:

$$\langle Y_j \rangle \equiv \mu_j = \frac{1}{f(b)} f(\mathbf{k}_\mu * S(t_j) + b). \quad (2.4)$$

We then described the stochastic properties of the synapses by extracting the standard deviation σ_j from our mean squared error based fitting of mean synaptic behaviour. This

allowed us to model the distribution of synaptic efficacies using a Gaussian distribution. While it is unclear which single parameter distribution is optimal for modelling the combination of synaptic response variability and sampling noise, we found this choice allowed for effective response fitting and a Gaussian noise model has been used effectively in similar work [112]. Using this approximation, synaptic efficacies can be predicted for different stimulation protocols for a synapse by sampling from the following distribution:

$$p(y_j|S, \theta) = g(y_j|\mu_j, \sigma), \quad (2.5)$$

Where $g(y_j|\mu_j, \sigma)$ is a Gaussian distribution with mean μ_j defined by equation 4, and fixed standard deviation σ . For a more extensive explanation and evaluation of the SRP model see Ref. [2].

2.4.4 DATA PROCESSING

We fit our data to a specific subset of the Allen Institute Synaptic physiology dataset [8] selected via the following criteria: Unless otherwise specified we selected data with a bath calcium concentration of 1.3 mM and bath magnesium concentration of 1 mM magnesium. The one exception is that for comparisons with 2 mM calcium data we selected data with 2 mM bath calcium. All queries required that the Allen Institute Synaptic Physiology dataset label a cell pair as a ‘synapse’, as excitatory, and not as a gap junction. Human or rodent data were specified for the respective analyses. We further excluded voltage clamp recordings, any pairs that were not associated with a ‘PulseResponseFit’, and any pairs

with pre-synaptic labels corresponding to gabaergic interneuron types (Sst, Pvab, Vip). We then pooled run data with 1-3 ms variability in delay duration (125 ms, 126 ms, 127 ms, and 128 ms to 125 ms and 250 ms, 251 ms, 252 ms, and 253 ms to 250 ms) and padded runs missing responses to some stimuli with ‘nan’ values. In rodent data we excluded all synapses for which either presynaptic or postsynaptic cells were not transgenically identified. We also excluded synapses that lacked an experimental trial with 50 Hz induction frequency and 250 ms delay, which we used in calculation of our ‘physiological’ representation. We also excluded all synapses with paired pulse ratios over 4 to exclude a small number with unrealistically high values likely corresponding to an initial failure. On a protocol by protocol basis we excluded responses with initial response values above 0.01 volts and replaced any values below 10^{-9} volts with the value 10^{-9} volts to remove extreme values. We then normalized all values in a protocol by dividing them by the average of the first responses after the aforementioned exclusions and floor value replacements.

2.4.5 MODEL FITTING

Once data had been pre-processed, we fit SRP model amplitudes, baselines, and standard deviations using simplicial homology global optimization [103] from the Scipy scientific computing library [122] version ‘1.6.2’. We selected amplitude and baseline parameters for each synapse by minimizing the mean squared error of our model prediction compared to recorded data across all runs and protocols. Finally, we computed the standard deviation for each synapse model based on the associated fitting errors. The baseline parameter was

constrained to a value range between -4.5 and 6 while the value ranges of the amplitudes of three summed exponential functions of our kernel were constrained to -150 to 150, -1000 to 1001, and -3000 to 3000 for corresponding time constants of 15 ms, 200 ms, and 300 ms.

2.4.6 UNSUPERVISED LEARNING

Prior to clustering and supervised learning we applied PCA and scaling to the parameters (including SD) captured by the SRP model using the Scikit-learn package [123]. Once the data were processed and scaled we ran the OPTICS (Ordering Points To Identify the Clustering Structure) algorithm from the Scikit-learn library [123] with a ‘sqeuclidean’ metric for minimum cluster sizes between 2 and 14. In both species we identified the minimum cluster size corresponding to the highest Silhouette Coefficient (a measure of cluster quality). We then performed all further analyses using OPTICS with the minimum cluster size set to this default, the metric set to ‘sqeuclidean’ and all other parameters set to the defaults from Scikit-learn version ‘0.24.1’. In rodents the minimum cluster size we used was 8 while in humans we used a minimum cluster size of 10.

2.4.7 SUPERVISED LEARNING

In order to perform supervised learning on our data we generated two representations corresponding to our SRP model fits and a basket of traditional physiological measures. The SRP representation included the amplitudes of fitted kernels, the baseline parameter, and

the standard deviation parameter for each synapse. The physiological representation included the mean summed value of the first eight stimuli, the mean ratio of the fifth response over the first response, the mean ratio of the second response over the first response (pair pulse ratio) for all runs with a 50 Hz induction frequency. Additionally, the physiological representation also included the release probability and the mean value of the last four responses in a train (those which follow the ‘delay’ divided by the mean of the first four for all protocols). We computed predictive accuracy in the prediction of all synapses onto each transgenically defined presynaptic or postsynaptic type that met the inclusion criteria described in in our data processing section. We also removed one synapse because it was the single remaining exemplar of its subclass and therefore could not be both trained and tested on.

MLP PARAMETERS

We used the Scikit-learn package [123], version ‘0.24.1’ to implement a multilayer perceptron (MLP) classifier with three hidden layers of 50 units each and all other parameters set to the default configuration parameters of the package.

COMPARISON CLASSIFIER PARAMETERS

We used the Scikit-learn package [123], version ‘0.24.1’ to implement gradient boosting, logistic regression, Adaboost, random forest, and support vector machine classifiers for comparison with the selected MLP architecture. We set the maximum number of iterations for the logistic regression to 10000, and the number of estimators for gradient boosting

to 500, otherwise all parameters were set to the default configuration parameters of the package.

2.4.8 QUANTIFICATION AND STATISTICAL ANALYSIS

2.4.9 FIT ERROR STATISTICS

The mean MSE and standard error of the mean for the first 8 and last 4 stimuli were obtained by computing the mean and standard error of the mean (SEM) of these all responses pooled across all post-synaptic subclasses. We also computed p-values for the probability that the mean MSE values associated with synapses onto somatostatin (Sst) cells for the first 8 and last 4 stimuli was different from the pooled mean MSE of all other synapses. These p-values were obtained using a standard t-test implemented using the Scipy scientific computing package in Python [122].

2.4.10 TESTING CONTINUUM HYPOTHESES

We generated SRP fits by sampling from a uniform distribution using the same bounds as those for the biological data, and assumed SD varied between values of 0 and 3. Within these bounds we sampled 3000 batches of 86 synapses from a uniform distribution by sampling each parameter between our defined bounds. For our multivariate Gaussian hypothesis we defined a multivariate Gaussian such that it was elongated along its primary axis of variance (PC1) by the same amount as the biological data (rodent), and displayed

the same degree of remaining variance evenly along the same number of remaining dimensions as the biological data. Finally, we tested the combination of uniform distribution sampling with our model fitting procedure. We first performed the same procedure described for the uniform distribution case to generate synapses and then generated artificial stimulus-response pairs by using the models to predict responses to different stimulus protocols from the Allen Institute Synaptic Physiology Dataset. We then used our SRP model fitting procedure to fit synapse models to the artificial stimulus-response pairs resulting in 3000 batches of 86 fitted SRP synapse models generated from a distribution with no inherent cluster structure. P-values were obtained by determining the share of samples in which each respective case matched or exceeded the Silhouette Coefficient associated with the clustering of real biological data.

2.4.11 CLUSTER AND GROUP SIMILARITY MEASURE

In order to calculate cluster similarity as in 2.4E and 2.5D we computed a combination of the correlation of median kernels and the difference in the Gaussian standard deviation parameter of the SRP model. Specifically, for each cluster pair we computed a Pearson correlation of the cluster median kernels and subtracted the absolute difference of SRP SD means between the two clusters divided by the standard deviation of the SRP SD parameters across the entire distribution of fitted synapses. This amounts to combining a correlation based measure of efficacy response with a measure representing the difference in SD between clusters relative to the overall spread of SRP SD values across all synapses.

This resulted in a similarity metric where a value of 1 corresponds to perfect temporal correlation with zero difference in mean SD parameters and lower values correspond to increasingly dissimilar groups.

2.4.12 CLUSTER NUMBER OCCURRENCE PROBABILITY

While performing our continuum null hypothesis testing (above) we also kept a record of number of clusters associated with the best Silhouette Coefficient in each run. Having recorded these values divided the number of occurrences of each count in the best performing continuum hypothesis (uniform distribution+SRP fitting) by the total number of counts (3000) to establish the probability of each optimal number of clusters. We then computed the statistic we refer to in the caption of 2.4C by summing the products of the probability associated with a value of 6 with probabilities associated with values of 5, 6, and 7.

CROSS-VALIDATION PROTOCOL

In order to ensure that all synapses would be used in both model fitting (training) and for model assessment (testing) a standard cross-validation protocol was applied. The data corresponding to all excitatory rodent cells recorded in 1.3 mM calcium conditions were presented in one of two representations (SRP model parameters or traditional physiological measures), randomly shuffled, and then partitioned into subgroups of 10 samples each which dictated the number of folds. After partitioning, all but one fold was used to fit

an untrained classifier and then the performance of that model was measured based on predictive accuracy on the remaining held out samples. This procedure was repeated for each fold to ensure testing on all available data. Further, the entire procedure was repeated for 1,000,000 random shuffles or 100,000 shuffles for downsampled calcium concentration comparisons. The random shuffling was applied to minimize order and grouping effects and repeated a large number of times to minimize variance in estimates of mean accuracy.

At each stage of the shuffling the difference between supervised predictive accuracy and accuracy baseline was computed and the distribution of these differences was used for significance calculations. We computed an accuracy baseline by using the ‘DummyClassifier’ class from Scikit-learn [123] to predict the most frequently occurring label from each training set for all entries in the corresponding test set and calculate the corresponding predictive accuracy. These accuracy baseline values provide a measure of how accurate a model could become by simply learning which subclasses occurred most regularly without learning anything about underlying synaptic dynamics. We computed the differences of mean cross validation accuracies between representations and against baseline for each shuffle. We then computed p-values by considering the proportion of shuffles in which representations outperformed baseline or each other.

ACKNOWLEDGEMENTS

We thank Luke Campagnola for feedback on the project as well as Zachary Friedenberger, Emerson Harkin, Xingyun Wang, Michael Lynn, Somayah Alsayadi, Leonard Maler, and Jérémie Lefebvre for helpful discussions. This work was supported by The Canadian Insti-

tutes of Health Research (CIHR):Project Grant RN38364 (RN and KT, providing salary to JB), by the Natural Sciences and Engineering Research Council of Canada: Postgraduate Scholarships-Doctoral (JB, providing salary to JB), and by a Canada research Chair to KT. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We also acknowledge the creative commons licensed synapse image (<https://doi.org/10.7875/togopic.2020.200> used with modification) and function image ('image: Flaticon.com') in our graphical abstract.

2.4.13 AUTHOR CONTRIBUTIONS

- Conceptualization: Richard Naud, John Beninger.
- Data curation: John Beninger, Julian Rossbroich.
- Formal analysis: John Beninger.
- Funding acquisition: Katalin Tóth, Richard Naud, John Beninger.
- Investigation: John Beninger.
- Methodology: John Beninger, Richard Naud, Katalin Tóth.
- Project Administration: John Beninger.
- Resources: Richard Naud.
- Software: John Beninger, Julian Rossbroich.

- Supervision: Richard Naud, Katalin Tóth.
- Validation: John Beninger.
- Writing – original draft: John Beninger.
- Writing – review & editing: John Beninger, Richard Naud, Katalin Tóth, Julian Rossbroich.

2.4.14 DECLARATION OF INTERESTS

The authors declare no competing interests.

2.4.15 SUPPLEMENTARY FIGURES

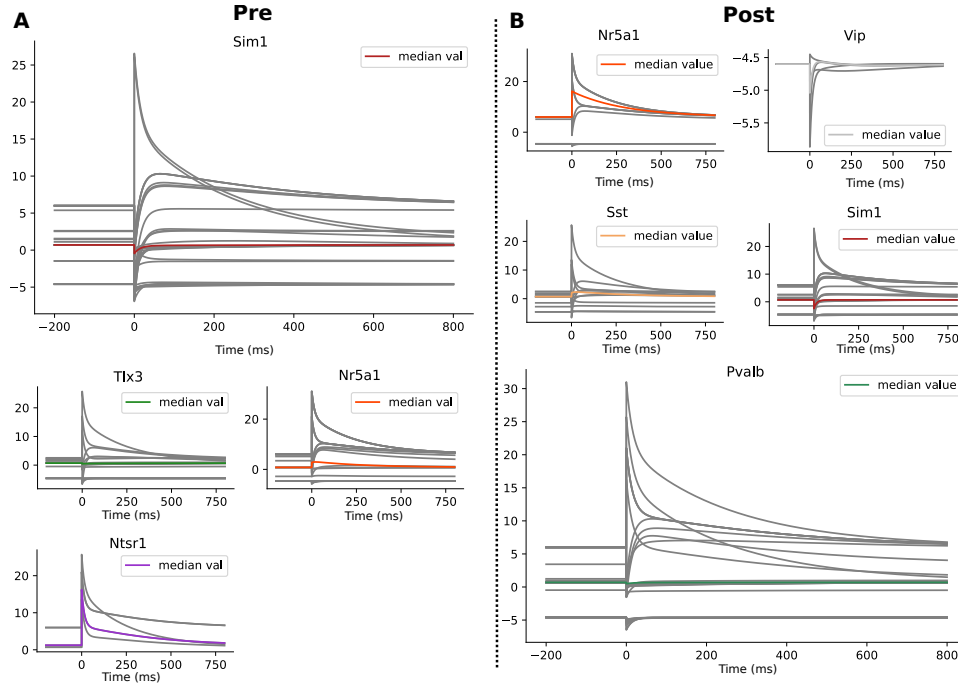


Figure 2.6: Kernels of SRP model fits by pre-synaptic (A) and post-synaptic (B) cell subclass. *Related to figures 1, 2, and 5.*

Cluster #	Baseline	A_1	A_2	A_3	SD
R0	5.996	-150.0	-1000.0	2570.635	0.929
R1	-4.6	-9.419	46.448	-67.945	0.689
R2	0.7	0.0	0.5	-122.902	0.432
R3	6.0	150.0	-1000.0	3000.0	1.008
R4	2.39	150.0	1001.0	3000.0	1.211

Table 2.1: Median SRP parameters of rodent synapses used in clustering (Fig 2.2, Fig 2.4, and Fig 2.5) of 1.3mM Ca data labelled by pre-synaptic subclass, post-synaptic subclass, and cluster identity. The authors will provide fits to outstanding synapses from the Allen Institute Synaptic Physiology Dataset upon request (See also: data and code availability). *Related to Figures 2, 4, and 5.*

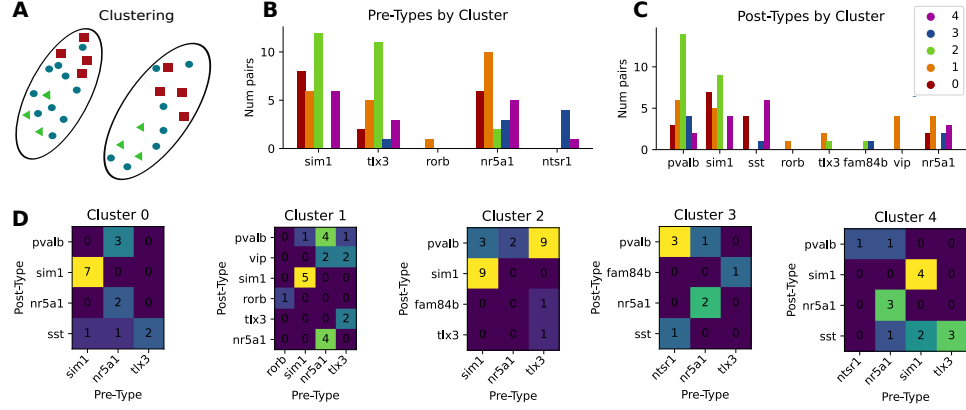


Figure 2.7: Clusters by pair identity. **A:** visual schematic illustrates unsupervised clustering. **B-C:** the functional cluster composition of synapses of each pre-synaptic (B) and post-synaptic (C) subclass. **D:** matrices showing the number of synapses of each pre-post subclass pairing in each cluster. *Related to figures 2 and 5.*

Cluster #	Baseline	A_1	A_2	A_3	SD
H0	0.7	-16.515	0.5	-123.938	0.367
H1	-4.6	-7.866	-10.322	-6.996	0.651
H2	6.0	-150.0	-1000.0	3000.0	0.624
H3	1.957	150.0	1001.0	3000.0	0.806
H4	6.0	150.0	1001.0	3000.0	0.521
H5	3.063	120.753	691.768	1387.867	0.758

Table 2.2: Median SRP parameters of Human synapses used in Fig 2.4 labelled by cluster identity. *Related to Figure 4.*

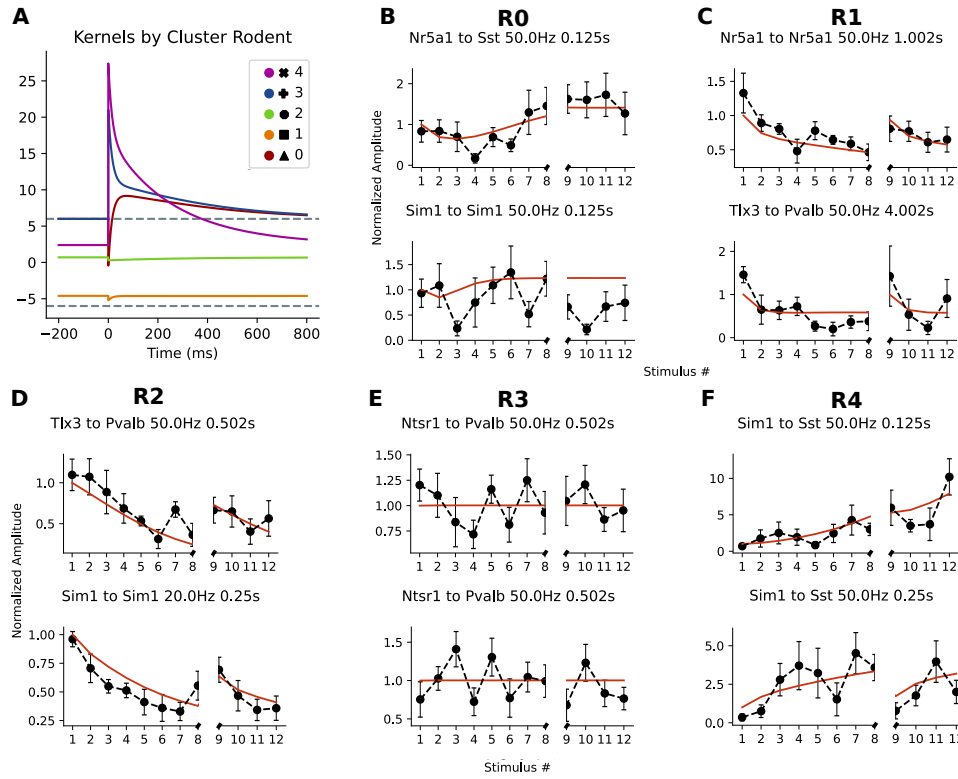


Figure 2.8: Example fits by cluster. **A** shows the median kernels associated with each cluster (replicated from 2.2C). **B-F** show fits predicting responses to a single protocol for two different connections from clusters R0-R4 respectively. The number of runs by fit are as follows: R0 Nr5a1 to SST: 8 runs, R0 Sim1 to Sim1: 5 runs, R1 Nr5a1 to Nr5a1: 5 runs, R1 Tlx3 to Pvalb: 4 runs, R2 Tlx3 to Pvalb: 5 runs, R2 Sim1 to Sim1: 5 runs, R3 Ntsr1 to Pvalb: 10 runs (top) and 11 runs (bottom), R4 Sim1 to Sst: 5 runs (top), 6 runs (bottom). **Related to Figure 2.**

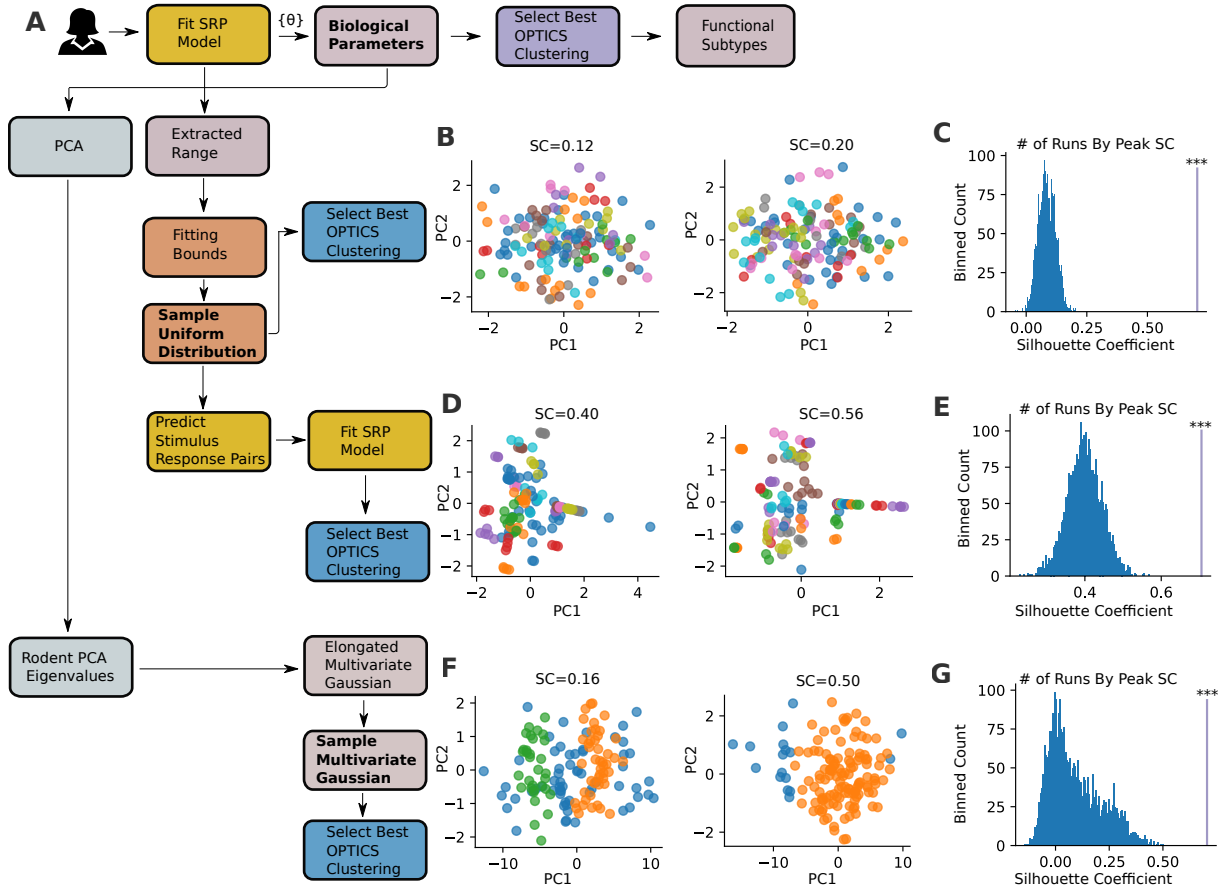


Figure 2.9: Testing uniform and continuum hypotheses in human. **A:** visual schematic showing the procedure we used to characterize functional subtypes in biological data (first row). The range of fitted model parameters and the output of the PCA are used to generate our null hypotheses. For the uniform hypothesis, we sampled from a uniform distribution using the extracted range (second row). We also tested the effect of applying our SRP fitting procedure to output simulated from the models sampled from the uniform distribution (third row). For the continuum hypothesis, we sampled from a multivariate Gaussian distribution whose variance structure matched the variance structure of the observations (fourth row). Each hypothesis was sampled 3000 times. **B-G** In each row, two exemplar scatter plots (**B**, **D**, **F**) show the result of clustering (different colors) in the space spanned by the first two principal component dimensions. The overall Silhouette Coefficient distributions associated with each hypothesis are shown in the right (**C**, **E**, **G**) with the observed Silhouette Coefficient shown with a purple vertical line. Triple stars indicate that the value observed in biology occurred less than once in 1000 random samplings of the corresponding hypothesis ($P < 0.001$, Bootstrap resampling). **Related to Figures 3 and 4.**

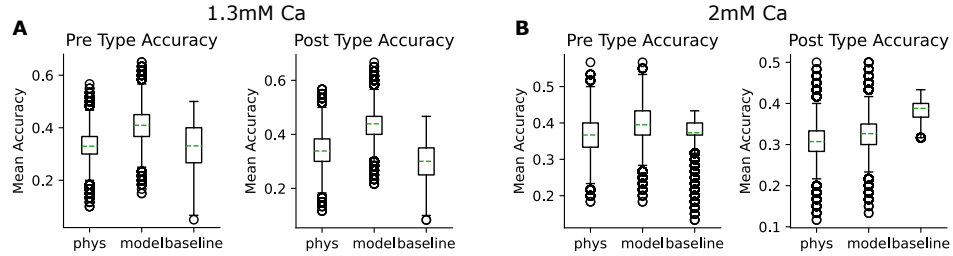


Figure 2.10: Supervised accuracy by extracellular calcium concentration matched for dataset size. The green lines indicate mean predictive accuracy across Bootstrapped shuffles, the box boundaries show interquartile range, and outlier values beyond the whiskers are plotted as circles. When matched for size (67 connections each) supervised accuracy is above baseline for model based prediction in 1.3 mM Ca (**A**) while neither representation supports accuracy above baseline in 2 mM Ca (**B**). Results are shown for 100,000 Bootstrap samples. *Related to Figure 5.*

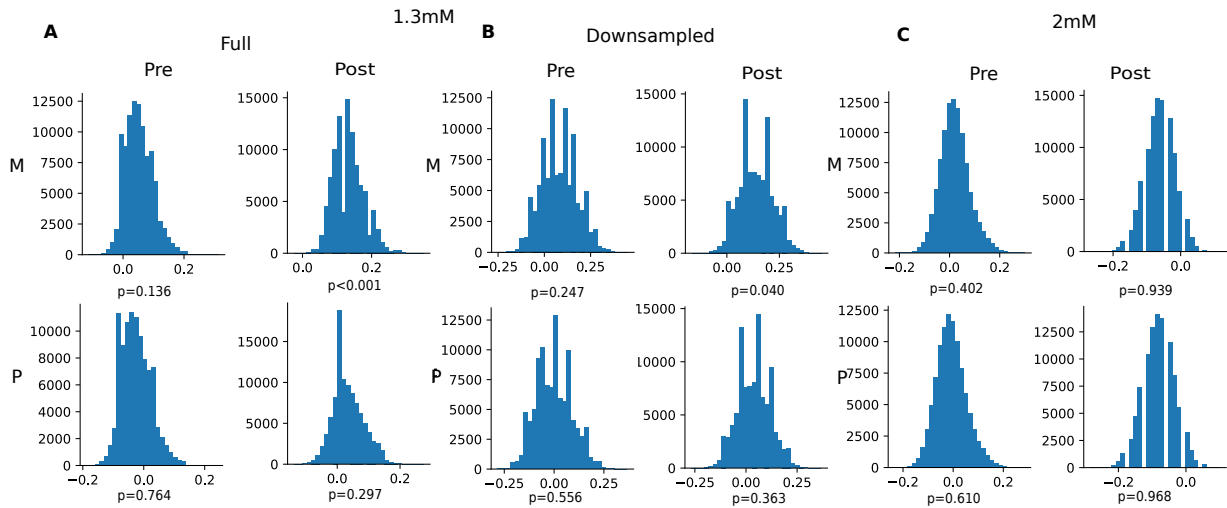


Figure 2.11: Results by dataset and representation for **A**: full sized 1.3 mM ca (85 pairs), **B**: downsampled 1.3mM Ca (67 pairs), and **C**: full sized 2 mM Ca (67 pairs) of Bootstrap testing for case when supervised model outperforms baseline with 100,000 shuffles. *Related to Figure 5.*

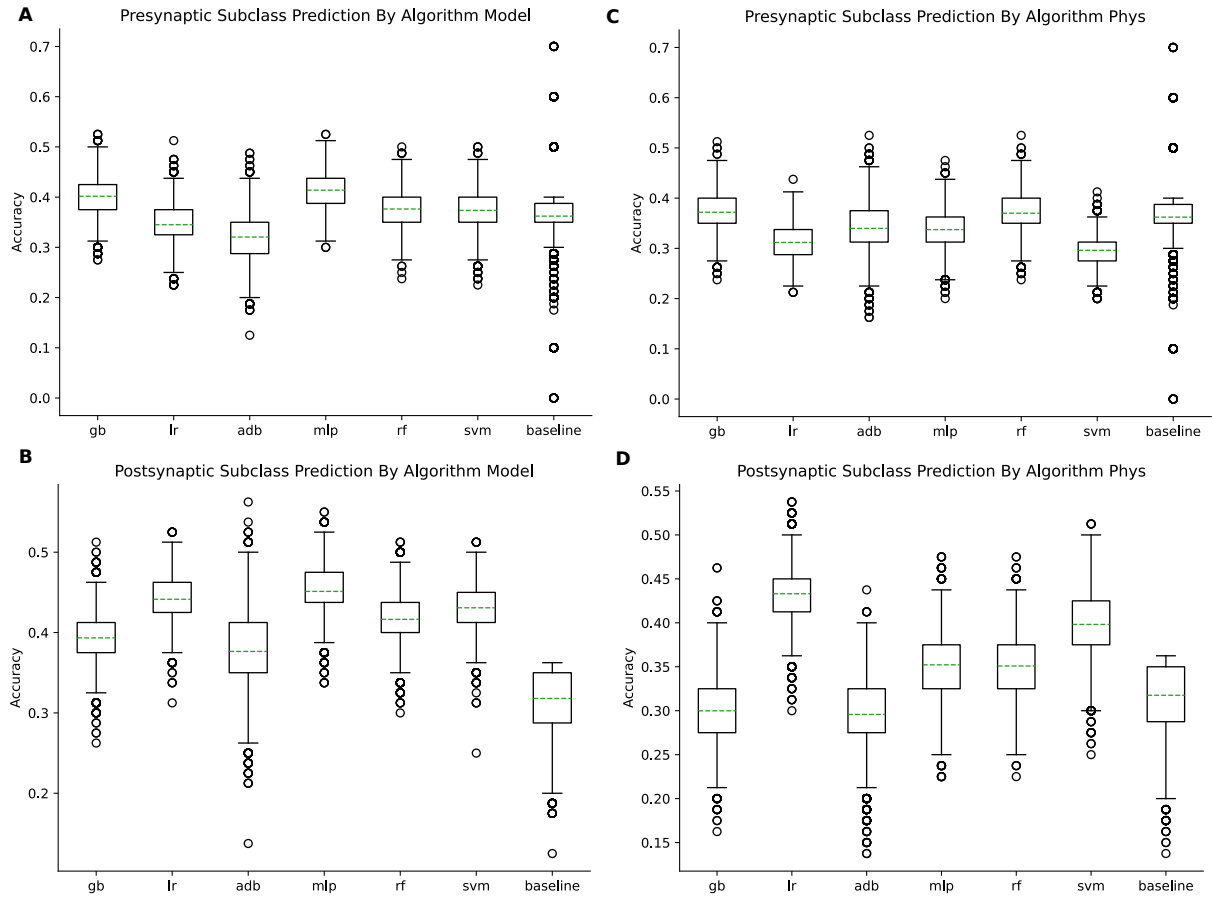


Figure 2.12: Boxplots show supervised out-of-sample testing accuracy in prediction of transgenically defined pre-synaptic and postsynaptic subclass for different algorithms across 10,000 Bootstrap iterations. The green line represents the mean predictive accuracy while box lines denote upper and lower quartiles of the data and circles represent outliers. The algorithms are as follows: ‘gb’=gradient boosting classifier, ‘lr’=logistic regression, ‘adb’=Adaboost classifier, ‘mlp’=multi-layer perceptron, ‘rf’=random forest, ‘svm’=support vector machine. Finally, ‘baseline’ displays the accuracy which can be attained by always predicting the most frequently occurring label from the training set. **A:** shows accuracies for prediction of pre-synaptic subclass using the model representation. **B:** shows accuracies for prediction of post-synaptic subclass using the model representation. **C:** shows accuracies for prediction of pre-synaptic subclass using a representation based on traditional measures. **D:** shows accuracies for prediction of post-synaptic subclass using a representation based on traditional measures. **Related to Figure 5.**

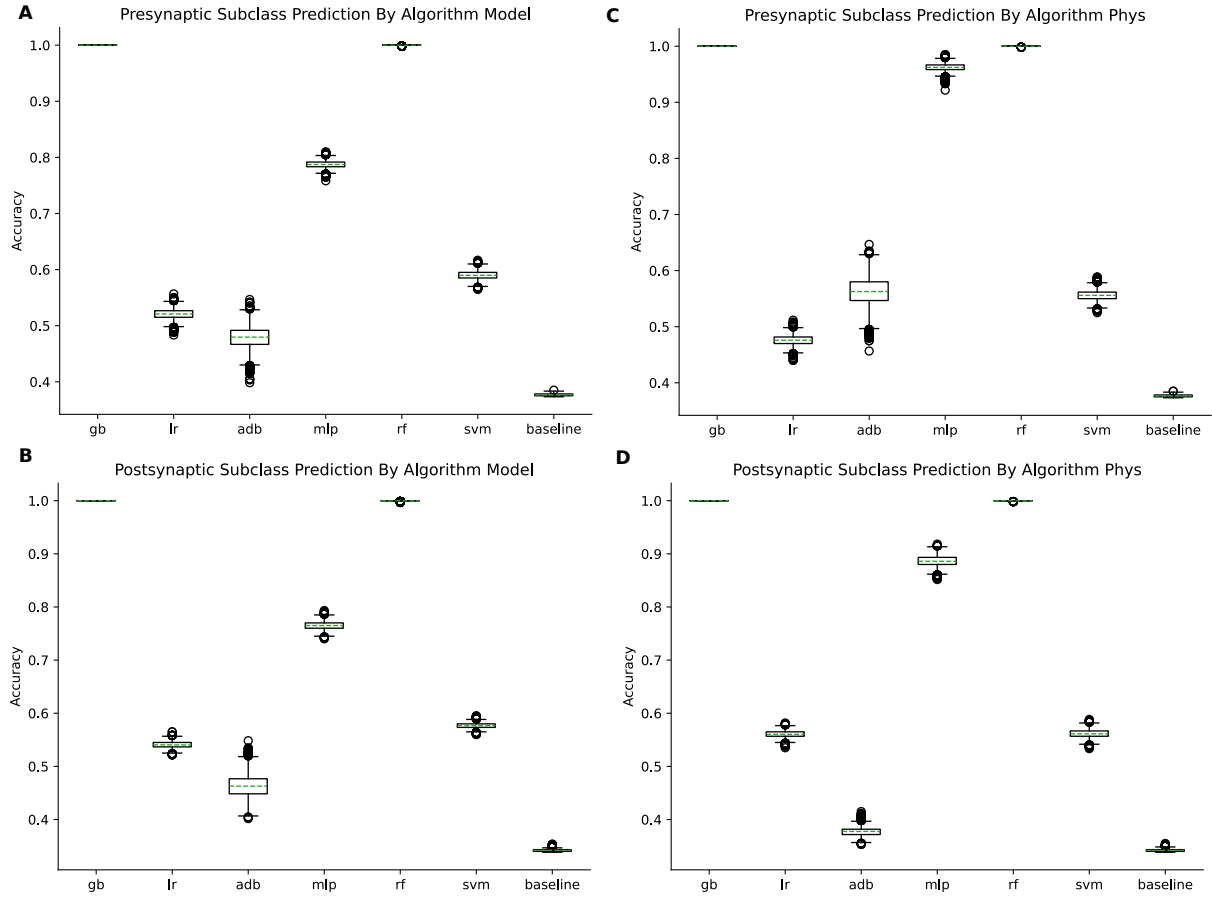


Figure 2.13: Boxplots show supervised within-sample training accuracy in prediction of transgenically defined pre-synaptic and postsynaptic subclass is shown for different algorithms across 10,000 Bootstrap iterations. This corresponds to the ability of the algorithms to predict identities of connections that they have previously seen. All other accuracy values in this work correspond to unseen "test" data. The green line represents the mean predictive accuracy while box lines denote upper and lower quartiles of the data and circles represent outliers. The algorithms are as follows: 'gb'=gradient boosting classifier, 'lr'=logistic regression, 'adb'=Adaboost classifier, 'mlp'=multi-layer perceptron, 'rf'=random forest, 'svm'=support vector machine. Finally, 'baseline' displays the accuracy which can be attained by always predicting the most frequently occurring label from the training set. **A:** shows accuracies for prediction of pre-synaptic subclass using the model representation. **B:** shows accuracies for prediction of post-synaptic subclass using the model representation. **C:** shows accuracies for prediction of pre-synaptic subclass using a representation based on traditional measures. **D:** shows accuracies for prediction of post-synaptic subclass using a representation based on traditional measures. *Related to Figure 5.*

PART II

CORTICAL WAVES

Is the ocean composed of water or of waves or of both?

*Some of my fellow passengers on the Atlantic were
emphatically of the opinion that it is composed of
waves;*

*but I think the ordinary unprejudiced answer would be
that it is composed of water.*

SIR ARTHUR EDDINGTON, NEW PATHWAYS IN
SCIENCE (1935)

BIOPHYSICAL MECHANISMS UNDERLYING THE CORTICAL RESPONSE TO ELECTRICAL MICROSTIMULATION

ABSTRACT

Electrical stimulation is a widely used tool for neuroscience research and has revolutionized clinical treatment of nervous system disorders. Yet how stimulation interacts with the biophysical properties of the different components of brain tissues to produce diverse neural responses is poorly understood. We present a unifying biophysical theory governing the dynamical interplay of excitatory and inhibitory responses to microstimulation in cor-

tex. Three key predictions emerged from this framework, each of which was validated with patch-clamp electrophysiology. First, microstimulation drove axial currents along both dendrites and axons that produced biphasic potentials at cell bodies. However, optical axotomy demonstrated that evoked action potentials were always mediated by axons. Second, differences in axonal cable properties led to early recruitment of parvalbumin-positive interneurons, which partially inhibited the circuit-level activation. Third, a synaptically-mediated spatial expansion of excited tissue could nonetheless occur upon synchronous recruitment of pyramidal cells. This spatial expansion is consistent with a fast and transient wave curtailed by the combined effect of refractoriness, inhibition and de-synchronization. Together, our findings identify the key biophysical determinants of microstimulation responses and their dynamical interactions, offering a predictive computational framework for precision engineering of this technique for future clinical and research purposes.

3.1 INTRODUCTION

Patterned electrical stimulation is now part of the standard of care for numerous neurological conditions [124]. Stimulation in the microampere range applied to the neocortex (henceforth referred to as microstimulation) has emerged as a promising approach for sensory restoration [62], epilepsy, and depression [125, 126], as well as for understanding the cortical functional organization [127, 128], connectivity [129, 130], attention [131], learning [132, 133] and core physiology [134, 135]. While clinical deep brain stimulation (DBS) [124] targets deep, mostly inhibitory feedforward structures, cortical microstimulation operates within nested recurrent excitatory-inhibitory networks [126]. This recurrence supports complex and sometimes counterintuitive phenomena such as balanced amplification [136] and travelling waves [21]. Furthermore, microstimulation causes activity that is highly synchronous across cells, potentially tapping into specific propagation mechanisms [137], oscillations [138], and travelling waves [139]. Whether these dynamics place cortical microstimulation in a qualitatively different regime than DBS remains largely unexplored.

Despite its relevance for human therapy and despite its long history in basic science, the biophysical mechanisms underlying the responses to microstimulation remain uncertain [126, 129]. Very weak ($1-10\ \mu\text{A}$) stimulation evokes action potentials in axons neighboring the electrode [140], leading to a sparse activation of cell bodies [69]. Higher stimulation intensities produce a sequence of excitation-then-inhibition [141–144], a perturbation that can travel laterally for over a millimeter. This secondary spread of activity underlies most

of the behavioral and systems-level effects of microstimulation [141, 145] and is thought to result from the complex recurrent organization of excitatory and inhibitory cells in the neocortex [146]. Elucidating the precise nature of this secondary expansion is pivotal for future engineering of behavioral, system-level, and therapeutic effects.

A first step toward this goal is investigating the physics of how different neuronal processes are recruited by an electric field. While in theory both axial currents along cables (axons or dendrites) and transverse currents across membranes can elicit action potentials [147], a convergence of prior modeling and experimental studies - each with their own limitations - supports the argument that axons are the primary locus of activation. Early modeling studies inferred activation properties based on the morphology of giant squid axons [148], while more recent modelling has leveraged biophysically detailed neurons to study network scale responses to stimulation [65, 149]. However, modelling has been constrained by the lack of temporally precise data about the spatiotemporal structure of the initial short latency response to stimulation. Experimental evidence remains indirect [140] or comes from outside of cortex [150, 151] and has not been able to fully disambiguate axons, somas, or dendrites [69, 152, 153]. This lack of direct evidence is a product of structural limitations: first, in the initial critical milliseconds of the response to electrical stimulation, extracellular recordings can be obscured by artifacts from the stimulation itself and second, the dynamics at play are fast compared to the temporal resolution of techniques such as two-photon calcium imaging [69, 154–156].

While the initial short duration response to microstimulation is characterized by net excitation, it depends on the properties of both excitatory and inhibitory neurons. There

is a diversity of axonal morphologies across cell types which may in turn imply type specific electrical recruitment biases. Determining whether the basic biophysics predict such a bias and characterizing its strength is key to understanding the initial conditions that trigger secondary spreading of activity in cortical networks [136, 138].

Perturbations to neural circuits, whether through electrical stimulation, localized insult, optogenetics, or strong synaptic input, can trigger laterally spreading activity [139]. Like a forest fire ignited by a stroke of lightning, a perturbation can spread from one neuron to the next and may not necessarily stop spreading. Multiple hypotheses have been put forward to explain how activity can spread over millimeters of tissue but not the whole brain. These include inhomogeneous connection strengths [139, 141, 157–161], excitatory delays [162], subnetworks without closed loops [163], and the propagation of early net excitation that is stabilized by slow mechanisms [161, 164]. Understanding the spatiotemporal structure of the initial response to microstimulation is essential to determining whether it can tap in to propagation mechanisms involving synchrony or travelling waves for which precise timing may underlie functions such as sensory representation [15–19], reward encoding [20], and perceptual gating [21]. We are therefore left with three questions for elucidating how microstimulation induces transient expansion: 1) What are the biophysical principles underlying AP generation in an external electric field? 2) How do morphological differences between excitatory and inhibitory cells shape recruitment bias? 3) How and under which conditions can an initial perturbation in both excitation and inhibition trigger transient expansion in cortical circuits with realistic connectivity?

3.2 RESULTS

3.2.1 TWO MODES OF MICROSTIMULATION-EVOKED ACTIVATION PRODUCE DISTINCT SOMATIC SIGNATURES

Microstimulation can trigger action potentials either through a direct depolarization of cell bodies [68] or through depolarization of dendrites and axons (cables; Fig. 3.1a). We modelled each of these modes of activation separately by considering two types of current that co-occur in neurons [147]. In the transverse current model, cell body depolarization arises directly from stimulation-induced currents across the cell membrane. In the axial current model, the curvature of the electric field around axons and dendrites produces current that flows along the cables. This elevates the membrane potential (V_m) near the stimulating electrode, causing a depolarization that then flows down cables ortho- and anti-dromically (Fig. 3.1b). The relative size of each effect depends on the shape of the field, the morphology of the neurons, and the electrical properties of their membranes.

To define the specific responses associated with each mode of activation, we performed computational simulations of intracellular recording from cell bodies near a simulated electrode. Transverse currents (Fig. 3.1c-f, top) produce an initial depolarization that is stronger for cell bodies closer to the electrode. In time, this initial depolarization relaxes back to the resting V_m (Fig. 3.1d, top). The peak potential occurs immediately after the end of the stimulation pulse. Its distance dependence is a function of the falloff of

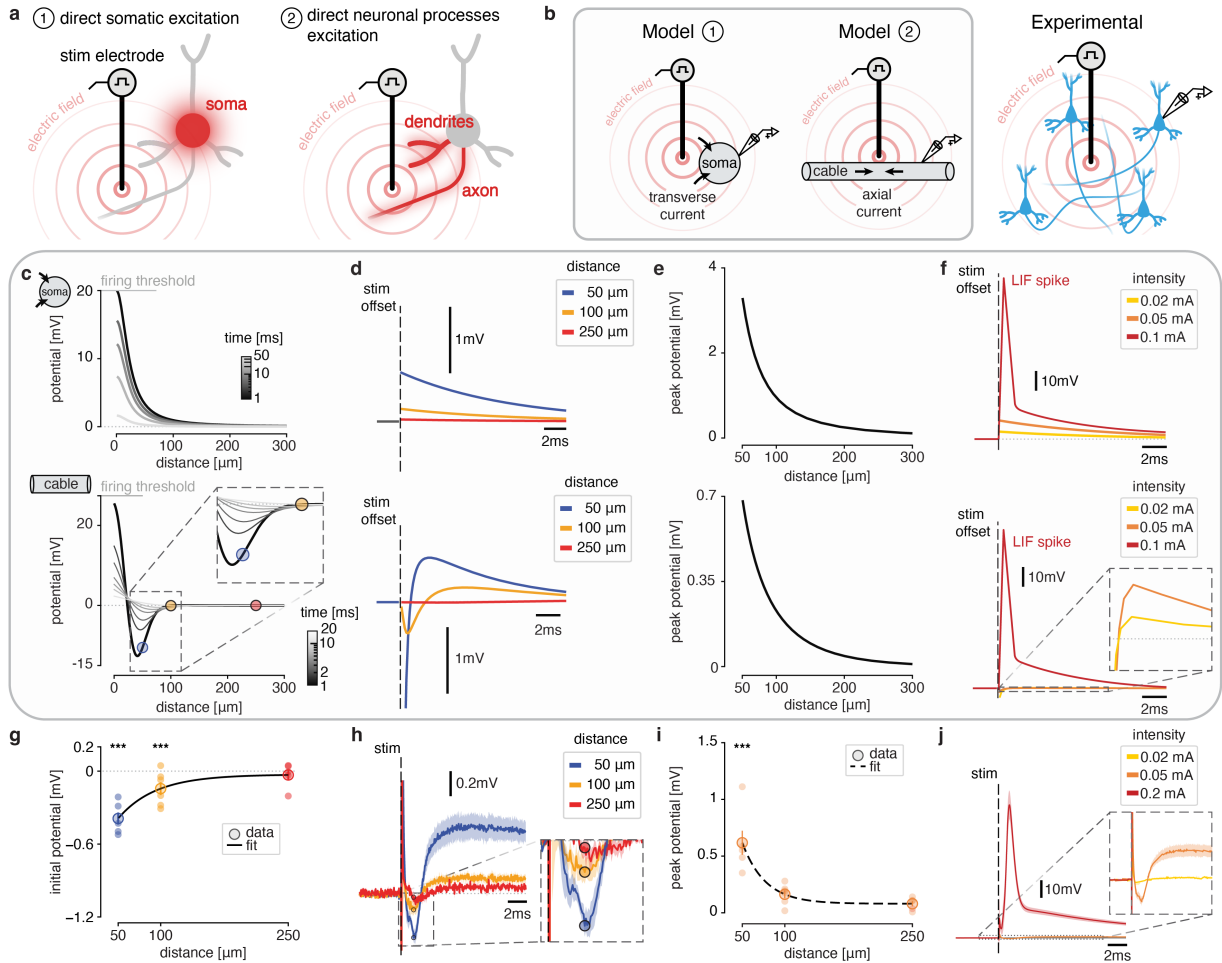


Figure 3.1: Near-electrode patch clamp distinguishes somatic from cable effects of microstimulation **a** Schematics representing the hypotheses that responses from electrical microstimulation could be the result of direct somatic excitation (left) or strong axial currents in neuronal processes (right). **b** Schematic of models and experiments of microstimulation. An electrode creates an electric field, which can create a strong transverse current at the soma (Model 1, left) or a strong axial currents in dendrites and axons of a cell further away (Model 2, right). **c** Transmembrane potential as a function of distance from the electrode for Model 1 (top) and Model 2 (bottom), different lines correspond to different times after the stimulation (colorbar). **d** Timecourse of evoked potentials at three distances (50, 100, 150 μm, as indicated in c, bottom) for Model 1 (top) and Model 2 (bottom). **e** Peak potential as a function of distance for Model 1 (top) and Model 2 (bottom). **f** Evoked potentials observed at a distance of 50 microns while increasing the intensity (20, 50, 100 μA) for Model 1 (top) and Model 2 (bottom). **g** Response potential 1.5 ms after stimulation offset (as indicated in f) for pyramidal cells at 50 μm (n = 6), 100 μm (n = 6), and 250 μm (n = 5) from the bipolar electrode. Error bars indicate s.e.m. **h** Stimulation-evoked responses (0.05 mA) in pyramidal cells at 50 μm (n = 6), 100 μm (n = 6), and 250 μm (n = 5) from the bipolar electrode. Inset shows the initial response at stimulation offset. Shaded region indicates s.e.m. **i** Peak evoked potentials for pyramidal cells situated 50 μm (n = 6), 100 μm (n = 6), and 250 μm (n = 5) from the bipolar electrode. Fitted exponential plus baseline ($Ae^{d/\lambda_{eff}} + B$) gave $\lambda_{eff} = 26.9 \mu m$. Error bars indicate s.e.m. **j** Evoked responses at peri-threshold intensities (0.02, 0.05 and 0.1 mA) for pyramidal cells at 50 μm from the bipolar electrode (n = 5). Shaded region indicates s.e.m. ***: p < 0.001

the electric field and hence a power law of separation between the recording site and the stimulating electrode (Fig. 3.1e, top). Increasing the electric field strength can increase this initial depolarization to reach the threshold for AP initiation. For pure transverse currents, a stimulation that is just below the threshold for AP initiation produces a large somatic depolarization (Fig. 3.1f, top).

In contrast, axial currents produce qualitatively different phenomena (Fig. 3.1c-f, bottom). Stimulation induces an initial depolarization that is high close to the electrode but flanked by hyperpolarization, reflecting the axial motion of charge. With time, both this central depolarization and the hyperpolarization at the flanks relax and broaden along the cable according to passive membrane dynamics [69]. When recorded at some distance from the electrode along the path of the cable, this produces potentials with a biphasic time course (Fig. 3.1d, bottom). These transients show smaller amplitudes farther from the electrode. This distance-dependence arises from power dissipation during passive propagation along the cable, such that the falloff with distance (Fig. 3.1e, bottom) scales with the electrotonic constant of the cable. Unlike in the transverse current mode, these transients can be observed even while the stimulation-induced extracellular potential around the cell is negligible. This mode of activation can initiate APs within the excitable cable (dendrite or soma). As a consequence, near-threshold stimulation can produce somatic potentials of very weak amplitudes since these are passively propagated from the site of AP initiation to the soma (Fig. 3.1f, bottom). Thus, two dynamical features of somatic potentials can distinguish the cell-body and cable mode of activation: biphasic subthreshold transients (Fig. 3.1d) and the abrupt production of APs (Fig. 3.1f).

3.2.2 AXIAL CURRENT FLOW ACCOUNTS FOR MICROSTIMULATION-EVOKED POTENTIALS

We can now contrast the specific predictions of these two modes of activation with whole-cell patch-clamp recordings near the microstimulation electrode. To isolate the direct (i.e. non-synaptic) effect, we recorded responses to stimulation in L2/3 cortical pyramidal neurons (PyrNs) in acute slices from primary visual cortex of adult mice in the presence of GABAergic and glutamatergic blockers. The initial potential (measured immediately after the capacitive artifact) showed negative deflections with an amplitude that fell off exponentially with distance (Fig. 3.1g). These initial negative deflections were followed by a slower depolarization, consistent with the biphasic transients predicted by the axial mode of activation (Fig. 3.1h). The peak potential reached by these transients also fell off exponentially as a function of distance with an effective space constant of $\sim 27 \mu\text{m}$ at $50 \mu\text{A}$ (Fig. 3.1i), commensurate with the effective space constant in basal dendrites [165]. Also consistent with cable activation, increases in intensity produced an abrupt transition from subthreshold transients to APs (Fig. 3.1j).

3.2.3 AXOTOMY REVEALS RELATIVE CONTRIBUTION OF DENDRITES AND AXONS

To disentangle the contributions of axons and dendrites in the responses to microstimulation, we performed axotomy on L2/3 PNs that were located at $50 \mu\text{m}$ from the stimulating

electrode, using two photon microscopy to sever the axon while leaving the initial segment intact. Responses to stimulation at different intensities were recorded using patch-clamp electrophysiology before and after axotomy (Fig. 3.2a). In both cases, synaptic transmission was blocked, as above. Axotomy abolished spiking across all stimulation intensities above rheobase (Fig. 3.2b-c). This indicates that dendrites do not produce APs in response to microstimulation at and around rheobase, consistent with their higher activation threshold [166]. For subthreshold stimulation intensities, the small biphasic transient recorded at the soma was unaffected by axotomy (Fig. 3.2d). This indicates that axial currents in dendrites are the main cause of observed somatic voltage changes. Since the peak somatic potentials reached by these dendritic effects are about 1-2 mV at stimulation intensities that trigger APs when the axon is intact, increasing the intensity by an order of magnitude (1-2 mA) would be required for these dendrite-mediated potentials to reach an amplitude sufficient to trigger an AP from the cell body. Thus, microstimulation elicits action potentials primarily via axons near the electrode, while dendritic axial currents are only capable of driving cell body activation at very high stimulation intensities (Fig. 3.2e).

3.2.4 DIFFERENCES IN AXONAL CABLE PROPERTIES PREDICT A RECRUITMENT BIAS ACROSS CELL TYPES

Different cell types show statistical differences in the cable properties of their axons, which affect the likelihood of their recruitment by microstimulation. Previous computational and experimental studies have established that myelination [151], cable diameter [148], and

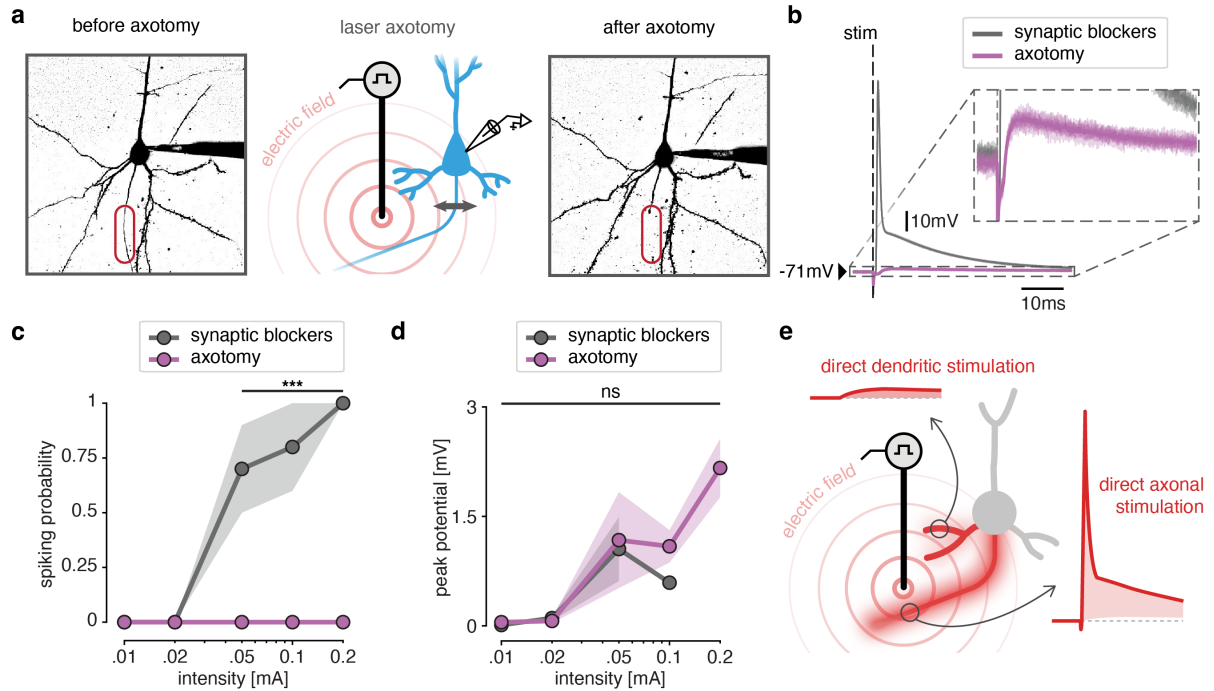


Figure 3.2: Axotomy distinguishes dendritic and axonal contributions of microstimulation
a Example two photon images of a pyramidal cell before (left) and after (right) axotomy. The red box shows the cut axonal part. **b** Evoked responses (0.1 mA) for pyramidal cells ($n = 5$) at $50 \mu\text{m}$ from the bipolar electrode, before and after axotomy. Shaded region indicates s.e.m. **c** Spiking probability from microstimulation (0.01 to 0.2 mA) for pyramidal cells ($n = 5$) at $50 \mu\text{m}$ from the bipolar electrode, before and after axotomy. Linear Mixed Model with pairwise comparison, $p < 0.0001$. Shaded region indicates s.e.m. **d** Peak responses from microstimulation (0.01 to 0.2 mA) for pyramidal cells ($n = 5$) at $50 \mu\text{m}$ from the bipolar electrode, before and after axotomy. Shaded region indicates s.e.m. Linear Mixed Model with pairwise comparison, $p = 0.75$. **e** Schematic depicting the relative contributions of direct stimulation of dendrites (top) and axon (right) to the recorded response at the soma. ***: $p < 0.001$, ns: non-significant

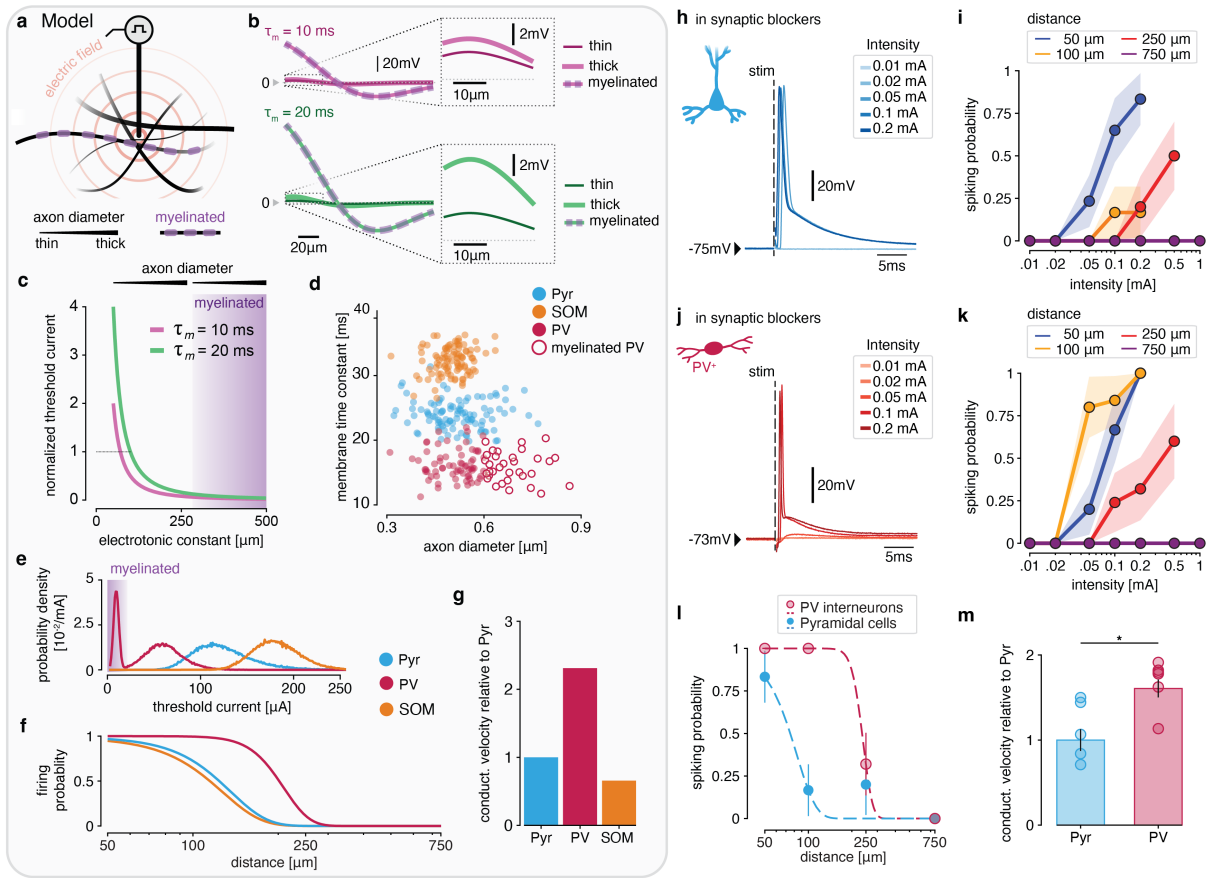


Figure 3.3: Cable properties predict differential recruitment of cortical cell types confirmed in slice experiments. **a** Schematic illustration of the model. A uniform cable under external stimulation can have varying thickness, myelination and membrane time constant. **b** The potential in the absence of firing threshold (free potential) along a uniform cable is shown, as measured immediately after a short stimulation that induced axial currents. Top, three cable length constants are considered, 80 μm (thin dark pink line), 100 μm (thick pink line) and 500 μm (thick pink dashed line). The inset expands the region between 0 and 30 μm. Bottom, similar types of cables with twice as large membrane time constant. **c** Theoretical relationship between electrotonic constant, membrane time constant and stimulation threshold. **d** Modeled cable properties across cell types. Based on electrophysiology and electron microscopy studies, we modeled different membrane time constant, axon diameters and myelination for SOM, Pyr and PV cells. **e** As a result of the heterogeneous cable properties, different axons are excited at different threshold intensities. The distribution of these thresholds is shown for the three different populations. **f** Firing probability for a cell body situated a fixed distance from the electrode as a function of this distance. **g** Average conduction velocity arising from the different cable properties across cell types. **h** Example evoked responses in synaptic blockers of a pyramidal cell for different stimulation intensities. The stimulating electrode is placed 50 μm away from the cell. **i** Spiking probability of evoked responses in synaptic blockers for pyramidal cells at different distances. Linear Mixed Model with distance ($p=0.0009$) and intensity ($p=2.08e-6$) as fixed effect. **j** Example evoked responses in synaptic blockers of a PV cell for different stimulation intensities. The stimulating electrode is placed 50 μm away from the cell. **k** Spiking probability of evoked responses in synaptic blockers for PV cells at different distances. Linear Mixed Model with distance ($p=1.06e-5$) and intensity ($p=2.75e-9$) as fixed effect. **l** Spiking probability of evoked responses in synaptic blockers for PV and pyramidal cells. Linear Mixed Model with distance ($p=2.46e-8$) and intensity ($p=0.007$) as fixed effect. **m** Conduction velocity of pyramidal cells and PV interneurons. Welch t -test. $p=0.02$ Error bars are sem. *: $p < 0.05$.

membrane resistivity [167] affect excitability.

The likelihood of recruiting a cell type with microstimulation depends on the cable properties of their axons, which may vary systematically across cell types. Detailed computational studies have predicted that membrane resistivity [167] and cable diameter [148] are key determinants of axonal excitability. Experimental studies have shown a direct effect of myelination in peripheral fibers [151]. Using a simple cable model, we derive the dependence of the activation threshold (I_T) on the observable properties of the axon (Fig. 3.3a-c; see Methods):

$$I_T = I_\infty \left(1 + \frac{\tau}{\lambda^2 T C} \right) \quad (3.1)$$

where T is the stimulating pulse duration, λ is the electrotonic length of the cable, τ is the membrane time constant and C captures the curvature of the field along the axon. The factor $\tau/\lambda^2 C$ determines the stimulation duration associated with a threshold that is double the rheobase (I_∞) and is called the chronaxie. This chronaxie equation recapitulates the facts that myelination, and conduction velocity (approximately λ/τ [168]) both increase excitability by increasing the electrotonic length. The seminal scaling laws of Stoney et al. (1968) [68] are also captured since I_∞ scales with the distance between the axon and the electrode d^p with the value of $p = 2$ expected for bipolar stimulation, the threshold is proportional to the squared distance, d^2 . It also predicts that a shorter membrane time constant translates to greater excitability. This expression was derived for monophasic pulses, but it can be generalized to biphasic pulses (Methods, Fig. 3.7). While this expression has a limited direct application because we do not know the curvature of the

electric field on the axon, it provides a compact description of the main biophysical elements regulating excitability: τ and λ .

These cable properties vary across cell types, such that parvalbumin-positive (PV) inhibitory neurons tend to be more myelinated, have slightly larger diameters [169–171] and have shorter membrane time constants [172] than pyramidal cells (Pyr). Somatostatin-positive (SST) neurons also show distinct properties, with slightly smaller diameters [173] and longer membrane time constants [172] than pyramidal cells. These differences in cable properties, however, are small and statistical. To estimate the expected differences in recruitment, we simulated a population of axons that had heterogeneous cable properties with statistical differences across cell types matching those reported in the electron microscopy and electrophysiology literature [169–173] (Fig. 3.3d). From the distribution of cable properties and our simple excitability model, we derived the distribution of threshold currents for axons at a fixed distance from the electrode. PV cells stood out as the most excitable cells, while SST cells had slightly lower excitability than Pyr cells. To better relate to the experimental observations, we assumed that when recording from a cell body some distance away from the electrode, there is a fixed distance-dependent probability that this cell has an axon close to the electrode. We used this model to predict the firing probability of cell bodies recorded at various distances from the electrode (Fig. 3.3f; see Methods). Furthermore, we illustrated that these different cable properties are associated with different conduction velocities (Fig. 3.3g). Together, observed cable and membrane properties across cell types predict that PV cells will be more readily recruited by microstimulation.

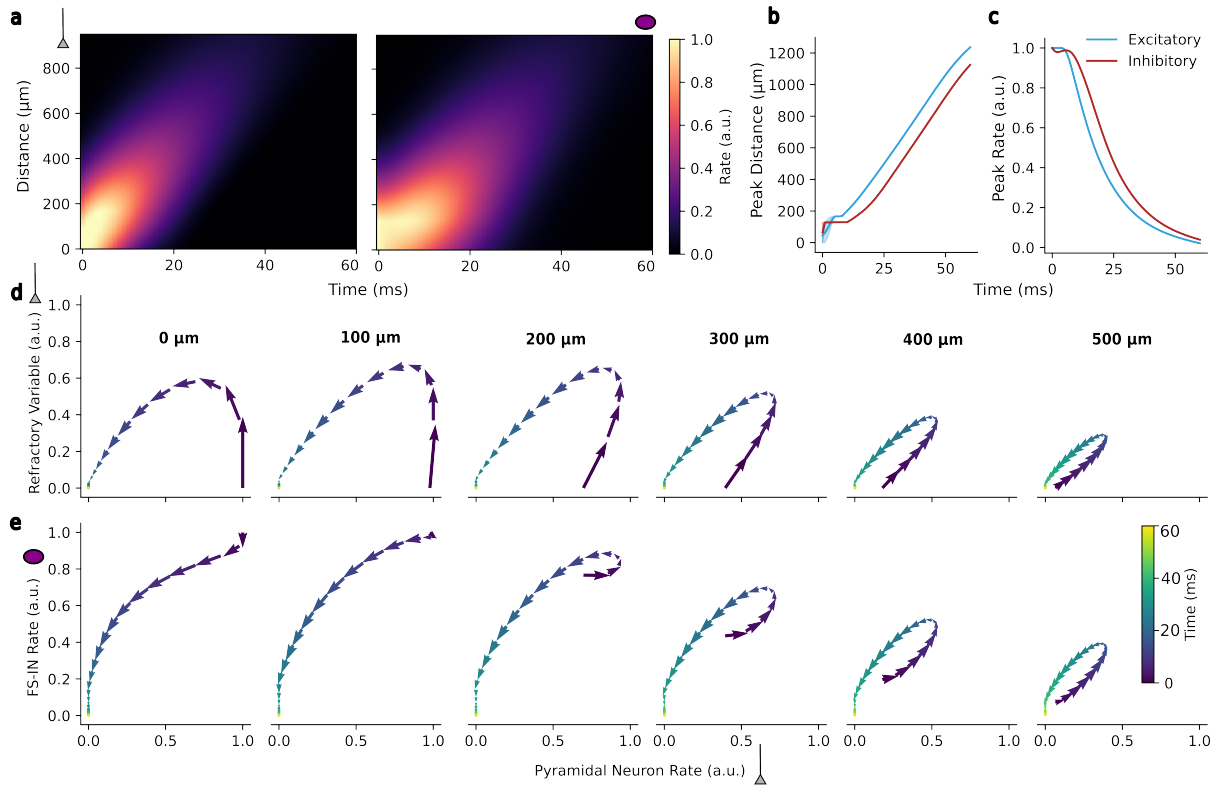


Figure 3.4: Strong stimulation induces a transient travelling wave pulse in a field model with refractory terms. *a* Heatmaps showing the evolution of the rate of a modeled excitatory population (left) and inhibitory population (right) over distance from the electrode and time following a representative microstimulation event. A decaying wave pulse can be seen propagation outward from the site of the initial perturbation. *b* Location of the highest rate of firing activity for excitatory and inhibitory neurons in the field model as a function of time. Both peaks travel away from the initial stimulation site over time. Shading indicates where the range of locations at maximal firing rate if there is more than a single highest firing rate location. This monotonic increase in peak excitatory distance corresponds to an outward expansion of the activity triggered by the stimulation *c* The highest rate at any location in each population as a function of time, this corresponds to the rate at the location shown in *b*. This monotonic decline in excitatory activity corresponds to a transient network response *d* Trajectories show the relationship between excitatory population rate and the refractory term over time at different distances from the electrode. Close to the electrode excitatory activity is already maximized by the stimulation event which drives an increase in the refractory variable that in turn reduces excitatory activity. Far from the electrode excitatory activity from locations active earlier in the response drives an increase in excitatory firing rate which again drives an increase in the refractory variable that eventually reduces excitatory rate. *e* Trajectories show the relationship between excitatory population rate and inhibitory rate over time at different distances from the electrode. Close to the electrode excitatory and inhibitory activity are initially maximized by the stimulation event, over time the activity of both populations declines with excitation declining faster. Far from the electrode excitatory activity from locations active earlier in the response drives an increase in excitatory and inhibitory firing rate which recurrently drives further increase until the combined inhibitory rate and refractory term drive down the rate in both populations.

3.2.5 PV INs ARE MORE READILY RECRUITED BY MICROSTIMULATION

To test the predicted difference in recruitment between PV INs and PyrNs, we performed patch-clamp recordings from these cell types in response to microstimulation (from 10 μ A to 1.0 mA) at distances ranging from 50 to 750 μ m while synaptic transmission was blocked. Responses of both PV INs and PyrNs showed increased spiking probability with increasing stimulation intensity at all distances. However, at distances greater than 50 μ m, the spiking probability of PyrNs remained low even at the highest stimulation intensity, indicating reduced recruitment of neurons further from the stimulating electrode (Fig. 3.3h-i). In comparison, PV INs required lower stimulation intensities than PyrNs to trigger spiking and exhibited higher spiking probabilities at greater distances (Fig. 3.3j-k). At a stimulation intensity of 200 μ A, PV INs exhibited higher spiking probabilities at distances up to 250 μ m than PyrNs (Fig. 3.3l), demonstrating enhanced recruitment of PV INs, consistent with model predictions (Fig. 3.3f). To further test that cable properties underlie these excitability differences, we measured the conduction velocity of PyrNs and PV INs. Our model predicted that a larger cable electrotonic length leads to more excitability and faster conduction. Consistently, PV INs exhibited higher conduction velocity than PyrNs (Fig. 3.3m). Our slice experiments thus support the conclusion that cable properties lead to preferential recruitment of PV INs.

3.2.6 TRANSIENT EXPANSION IS CONSISTENT WITH TRANSIENT TRAVELLING WAVES IN A TWO POPULATION FIELD MODEL WITH RE-FRACTORINESS

When microstimulation evokes action potentials in cells, by definition it triggers synaptic transmission to connected cells. This can result in additional “second order” firing [126]. Within highly recurrent cortical networks, second order firing can in turn induce a cascade of additional firing events. We sought to understand the constraints imposed on the occurrence of this secondary response by the relative connectivity between excitatory and inhibitory cells. In order to isolate mechanisms, we began with the simplest possible model: a one-dimensional field model where excitatory and inhibitory cells connected across space based on a Gaussian function (with a sigma of $115\ \mu\text{m}$) that was inspired by the relationship between connection probability and lateral inter-somatic distance in cortex [8]. In this model, the spatial scale of the overall response was either limited to a single initial round of second order firing, in which case its spatial scale was heavily constrained by the $115\ \mu\text{m}$ connectivity (Fig. 3.9a-c), or triggered runaway recurrent excitation (Fig. 3.9d-f). Neither the limited scope of the spatial response in the first case, nor the runaway activity of the second case, were consistent with previously reported in vivo responses [19]. Therefore, we sought a mechanism to explain the increase in the spatial scale of the response without runaway excitation, a response we refer to as transient expansion.

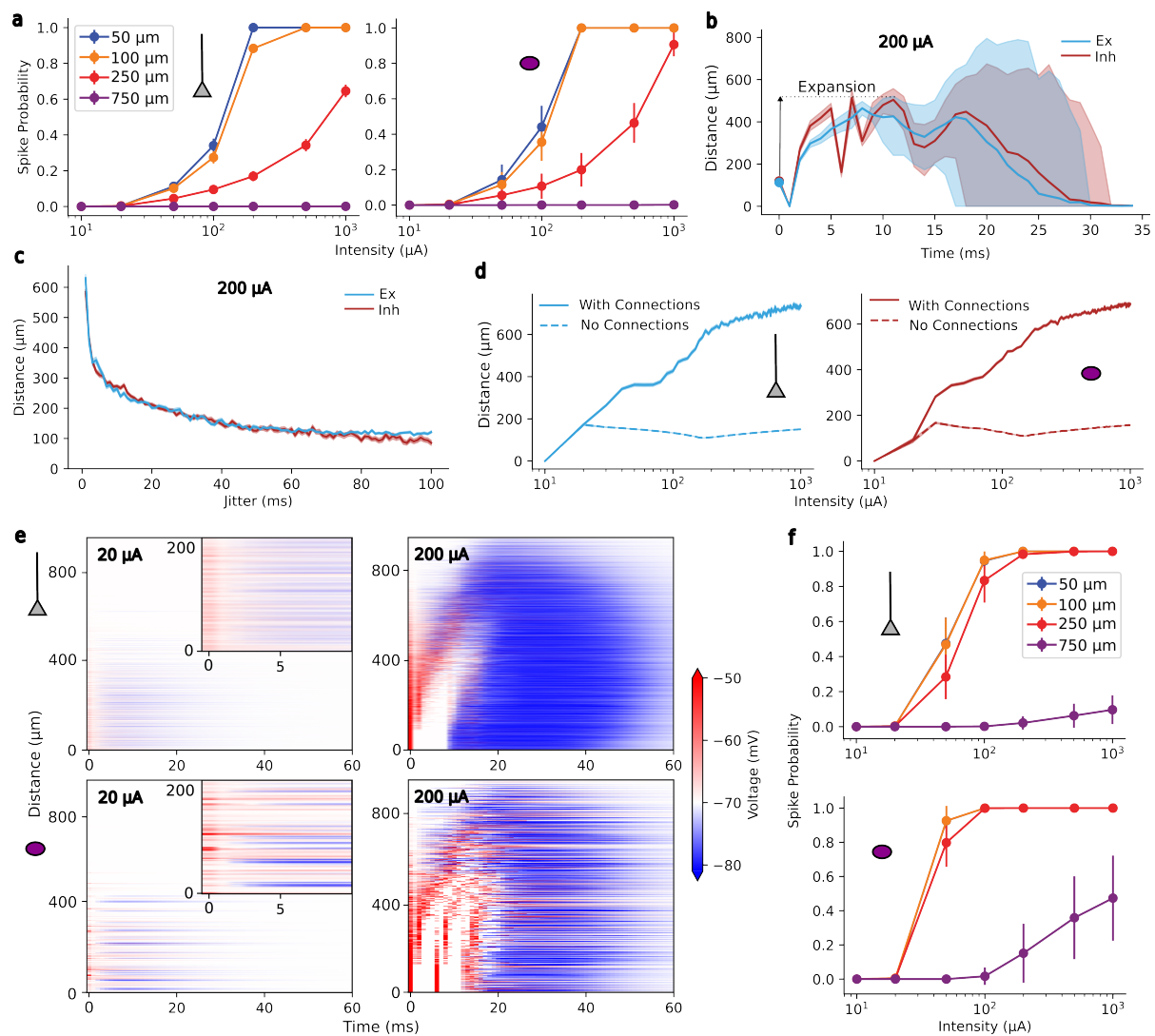


Figure 3.5: Spiking neural network model of the cortical response to electrical microstimulation. We simulated the response to stimulation in a spiking neural network model containing two classes of neurons: strictly excitatory neurons, and strictly inhibitory neurons. Neurons of each type were randomly assigned point locations in 1D space, defined with respect to the point of stimulation and connections were assigned based on class specific distance dependent connection probabilities from [174]. Other network parameters were set within biologically plausible ranges (see section 3.4.4). **a-d** And **f** show the model response to 200 electrical stimulations at each intensity indicated. **a** Mean firing probabilities for excitatory (left) and inhibitory (right) model neurons at different distances from the stimulating electrode in the absence of synaptic connectivity, bars indicate SD. **b** The mean distance from the point of stimulation of neurons spiking at each point in time is shown for excitatory (blue) and inhibitory (red) neurons in response to a 200 μ A pulse. The center line indicates the mean across repetitions while the shading indicates 95% of the data. The arrow inset highlights the synaptically mediated increase in mean distance of the response (i.e. transient expansion). **c** The mean distance from the point of stimulation at the peak of the response of excitatory (blue) and inhibitory (red) neurons when jitter is added to spread the firing associated with microstim over a much longer time course. As jitter increases, mean firing distance at peak decreases highlighting the importance of temporally synchronous stimulation for transient expansion. Solid lines indicate the mean over 100 repetitions and shading indicates SEM. **d** Mean distance from the point of stimulation of excitatory (left) and inhibitory (right) neurons spiking at the peak of the response. Solid lines show the model response with synaptic connections while dashed lines indicate the response to only the direct effect of stimulation without synaptic connections, shading indicates SEM. **e** Heatmaps showing the thresholded voltage response to stimulation of neurons at different distances over time. Each row corresponds to a single simulated neuron. Excitatory (top row) and inhibitory (bottom row) neurons are shown for example runs at 20 μ A (left column) and 200 μ A (right column). The insets in the left column show magnified sections of the same plots. **f** Mean firing probabilities for excitatory (top) and inhibitory (bottom) model neurons at different distances from the stimulating electrode in the presence of synaptic connectivity, bars indicate SD. The model predicts that above a minimum intensity threshold synaptic connectivity amplifies the response to electrical stimulation, allowing for spiking in more distant neurons as illustrated by the differences between **b** and **f**.

Past work has identified that in excitable media, the combination of a fast excitatory mechanism that facilitates propagation with a slower mechanism that stabilizes activity can give rise to travelling wave pulses [85]. In structured neural networks, these types of non-periodic travelling waves can propagate spatially without triggering runaway recurrent excitation [139, 161]. However, in models where the stabilization mechanism was inspired by spike frequency adaption or synaptic depression, the wave was slow relative to the initial short latency microstimulation response. For this reason, we introduced refractoriness into the model [164, 175] through a term that captured the difference in the relative refractory

rates of the PyrN and PV IN populations (i.e. we subjected the PyrN population to stronger refractoriness).

In simulations of the expanded model, PyrNs initially briefly overpowered FS INs which resulted in spatial propagation of activity (Fig. 3.4a-b). However, because PyrNs were unable to sustain the initial high rate of firing over time, FS INs stabilized the response over a longer time horizon (Fig. 3.4c-e). This process played out recursively: at each point in time, net excitatory activity at the wave front drove the creation of a neighboring peak in the next time step. The result was a non-periodic travelling wave that propagated away from the point of stimulation over time (Fig. 3.4b, 3.4d, and 3.4e).

Travelling pulses in excitable media are not always transient: a fire could potentially travel an infinite distance across an infinitely large dry forest. In our model, the travelling wave was transient because at each time step net excitation was only strong enough to produce a peak that was weaker than the one which came before (i.e. the wave was damped by the relative weights) (Fig. 3.4c). We conclude that transient expansion can occur naturally in a simple model of E & I networks when stimulation induces a sufficiently high level of initial firing even without the stabilization provided by an inhibitory surround.

3.2.7 TRANSIENT EXPANSION IN SPIKING MODELS DESPITE PREFERENTIAL RECRUITMENT OF INHIBITORY CELLS

The combined importance of refractoriness and high initial firing rate in transient expansion suggests that it is intimately tied to the concept of spiking synchrony. Because field models

do not produce spikes and their treatment of refractory effects is approximate [176], we developed a spiking neural network constrained by experimental observations. The model consisted of a strictly excitatory population (11,700 neurons) and a strictly inhibitory population (1300 neurons) with all neurons uniformly distributed across a one-dimensional space. We initialized the model with firing probabilities that were dependent on the distance of each neuron from the electrode. Population specific differences in recruitment were incorporated using a phenomenological model that was calibrated to be consistent with our earlier empirical observations (Fig. 3.5a, see section 3.4.4 for full model details).

Next, we introduced synaptic connectivity to the model. We set connection probabilities in a pair-specific manner (E-E, E-I, I-E, and I-I) based on the distance between neurons [158, 174, 177]. Despite the network having been initialized with a higher fraction of PV IN firing than PyrN firing, activity expanded spatially, dramatically increasing the overall scale of the response (Fig. 3.5b). This expansion was transient; over time the network reverted to a state with little to no activity. We observed this expansion at moderate to high ($>20 \mu\text{A}$) stimulus intensities below which the impact of connectivity on response scale was limited (Fig. 3.5d). For intensities greater than approximately $20 \mu\text{A}$ and below $100 \mu\text{A}$, this difference likely corresponded to primarily to second order firing. At higher intensities, the spatial scale of the response increased as a function of intensity, indicating that the response was shaped by an intensity dependent travelling wave.

By default, microstimulation in the model induces an initial volley of highly synchronous spikes. In order to test whether such synchrony was necessary to evoke transient travelling waves, we performed further simulations where we deliberately desynchronized inputs. In

these simulations we introduced jitter by uniformly distributing the times at which individual cells received microstimulation input over defined input windows without changing the total input. Strikingly, the jitter had a potent effect on the spatial scale of the response (Fig. 3.5c). Spreading microstimulation input over a 5 to 10 ms window approximately halved the mean distance from the electrode of cells that fired during the peak of the response. Decreasing the synchrony further by spreading spikes over long windows on the order of 100 ms ultimately abolished transient expansion. This shows that microstimulation can naturally induce a travelling wave of synchronous spikes across more than one synaptic connection making the evoked dynamics a form of unstable synfire chain [137].

What made this expansion transient? As in the field model, inhibitory cells in the spiking model fired at higher frequencies than excitatory neurons [178]. As a result, when a large number of both excitatory and inhibitory neurons were driven together and fired in synchrony they exerted an initial net excitatory effect but over time they desynchronized until they produced net inhibition. Because neurons closest to the electrode fired first, this sequence of net excitation followed by net inhibition was first seen close to the electrode and expanded outward over time (Fig. 3.5e). Once net inhibition expanded far enough to prevent further spatial expansion, overall activity declined until all neurons had returned to rest. This effect was intensity dependent in that at 20 μA , most direct responses were below firing threshold and the firing events which did exist were sparse, spatially dispersed, and predominantly inhibitory. There was thus no meaningful secondary sequence of synaptic activity. In contrast, 200 μA stimulation produced robust firing of both neuron types close to the electrode and a wave of excitatory activity was seen propagating outward from the

electrode during the first 20 ms following stimulation. The full intensity-dependent extent is shown in Figure 3.5f. Altogether, the propagating synchrony model can explain both the spatial extent and the transient nature of stimulation driven expansion.

3.2.8 TRANSIENT EXPANSION IN ACUTE CORTICAL SLICES

To test the predictions of the models, we evaluated transient expansion in neocortical slices by comparing the activity of PyrNs and PV INs with synaptic transmission intact. Subthreshold responses and spiking probability of PyrNs increased with stimulation intensity (Fig. 3.6a-c), more steeply near the electrode. Similarly, PV INs exhibited increases in both subthreshold responses and spiking probability with increasing stimulation intensity (Fig. 3.6d-f). These responses were stronger, and spiking probability was higher than in PyrNs. Consistent with the sequence of a rapid one or two spike phase of PyrN activity followed by a longer phase of PV IN dominance in our models, most PyrNs responses included one or two spikes (Fig. 3.11a-d) whereas at high stimulation intensities PV INs produced many spikes (Fig. 3.11j) and had firing rates as high as ~ 500 Hz (Fig. 3.11k).

To quantify transient expansion, we compared the spatial extent of activity of the PyrNs and PV INs before and after blocking synaptic transmission. Responses of PyrNs located 750 μm from the electrode were abolished after synaptic blockade (Fig. 3.6j). In PyrNs closer to the electrode and in PV INs, responses were reduced when synaptic transmission was blocked (Fig. 3.12). Additionally, the onset of responses to stimulation was delayed at distances greater than 100 μm (Fig. 3.6g), indicating transient expansion

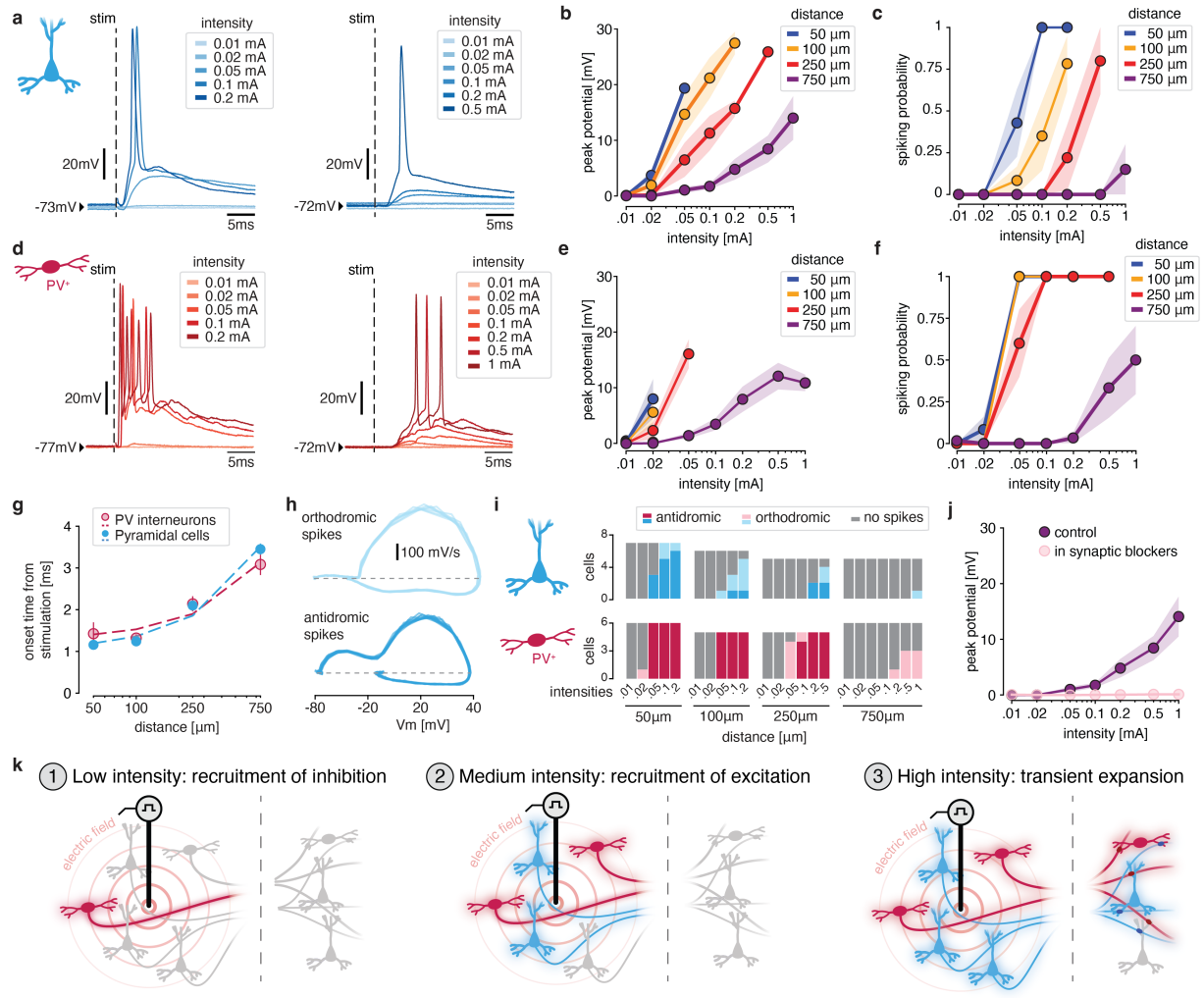


Figure 3.6: Transient spatial amplification by imbalance of E-I. **a** Example of evoked responses for a pyramidal cell situated 50 μm and 250 μm from the stimulating electrode. **b** Peak response of evoked responses from pyramidal cells at different distances. Linear Mixed Model with distance ($p=1.37\text{e-}6$) and intensity ($p\leq 2\text{e-}16$) as fixed effect. **c** Spiking probability of evoked responses from pyramidal cells at different distances. Linear Mixed Model with distance ($p=2.62\text{e-}11$) and intensity ($p=1.5\text{e-}11$) as fixed effect. **d** Example of evoked responses for a PV IN situated 100 μm and 750 μm from the stimulating electrode. **e** Peak response of evoked responses from PV cells at different distances. Linear Mixed Model with distance ($p=0.002$) and intensity ($p=1.81\text{e-}6$) as fixed effect. **f** Spiking probability of evoked responses from PV cells at different distances. Linear Mixed Model with distance ($p=6.77\text{e-}8$) and intensity ($p < 2\text{e-}16$) as fixed effect. **g** Onset time of evoked responses from PV and pyramidal cells. Linear Mixed Model with distance ($p = 1.61\text{e-}13$) and cell type ($p = 0.834$) as fixed effect. **h** Example phase plot of orthodromic and antidromic spikes. **i** Number of orthodromic and antidromic spikes depending on stimulation intensity and stimulation electrode distances. **j** Peak with and without drug for pyramidal cells situated at 750 μm from the stimulating electrode. Linear Mixed Model with drug condition (control vs. synaptic blockers, $p=3.54\text{e-}12$), intensity ($p=0.955$) and their interaction ($p=3.09\text{e-}8$). Error bars are sem. **k** Schematic demonstrating the recruitment of interneurons at low stimulation intensity (left), followed by the recruitment of pyramidal cells at medium stimulation intensity, up to a transient expansion with cells at greater distances being synaptically recruited at high stimulation intensity.

driven by synaptic transmission. Both of these observations were consistent with our model predictions (Fig. 3.5f). To further dissect spiking activity near the electrode compared to that at greater distances, we analyzed spike shapes and voltage trajectories via phase plots. Spikes induced by synaptic transmission exhibited a relatively slow depolarization up to AP initiation (orthodromic spike). In contrast, direct axonal activation produced a rapid depolarization that propagated back to the soma (antidromic, Fig. 3.6h). PyrNs mostly show antidromic spikes near the electrode and orthodromic spikes gradually become dominant further away from the electrode. In comparison, PV INs show the same transition but the transition from mostly antidromic to mostly orthodromic occurred farther.

3.3 DISCUSSION

Understanding the mechanisms underlying neural responses to microstimulation is critical for its effective deployment in both clinical and basic science contexts. A critical insight of our study is that distinct cable properties cause differential recruitment of specific excitatory and inhibitory populations. We showed that low intensities recruit mainly PV-positive axons in mouse cortex. At moderate intensities, microstimulation recruits some excitatory neurons but not enough to initiate the cascade necessary for transient spatial expansion of the response. High intensities directly activate most excitatory and inhibitory axons close to the electrode and create conditions of distinctly elevated local synchrony. This synchrony gives rise to a transient travelling wave. Because the wave front de-synchronizes as it propagates, the evoked response is transient in nature and propagates over an intensity

dependent distance.

If microstimulation can induce travelling waves in cortex, then it may be useful as an experimental perturbation to study the biological role of travelling waves. Sensory stimulation has been proposed to evoke cortical travelling waves that share information between spatially distant neurons [15, 16, 20, 21, 179] allowing for the integration of information from distinct, topologically organized receptive fields [12, 14] and acting as a record of the recent past [13]. Whether the waves evoked by microstimulation operate on the same pathways proposed to carry these types of information remains an open question. However, past work does support the idea that microstimulation can deliver sensory information to cortex [19] and that it is most effective in doing so when stimulation is structured to evoke naturalistic cortical activity [62]. Because our account is spatially and temporally precise, mechanistic, and computationally tractable, it could potentially enable the use of microstimulation as a general purpose tool for the study of evoked cortical travelling waves under conditions of both health and disease.

While careful optimization of microstimulation timing patterns shows promise in opening new therapeutic frontiers [62], it also requires exploration of a large space of stimulation parameters. Notably, trains of microstimulation evoke responses that interact with each other and become entangled with other history dependent dynamics such as short term plasticity [1, 8]. Studying dynamics in such a nuanced system entails fundamental tradeoffs between system complexity and mechanistic isolation that are exacerbated by the precision limits of large scale recording modalities: In high density electrical recordings, data are briefly obscured by artifacts from the electrical stimulation itself. Further, data recorded

with optical methods may lack the temporal resolution to precisely identify moving fronts of single spikes. In the face of this complexity we aimed to fully constrain the simplest case: single pulses. Our results lay a foundation for understanding the response to complex patterned stimulation in terms of the interactions of individual pulse responses with precisely defined spatiotemporal dynamics.

Can our model make predictions about the response of human tissue to cortical stimulation? Human neurons are larger than their rodent equivalents [153, 180] which implies that their axons have higher conduction velocities. This should result in axonal recruitment at lower stimulation intensities. Further, limited evidence suggests that, outside of Layer 1, GABAergic axons may be less myelinated in humans than in rodents [171]. Unless pyramidal cells express similarly lower levels of myelination, this would imply a weaker preference for the recruitment of inhibition and therefore responses with larger spatial scale. To this point, human cortical neurons appear to connect with other neurons at slightly longer average distances than in mice [174], implying that the mechanism of transient expansion may reach greater spatial scales. Taken together, these observations suggest that stimulation of human cortical tissue should evoke mechanistically similar responses on a larger spatial scale.

This study elucidates the biophysical mechanisms underlying diverse responses to cortical microstimulation paving the way for advancements in therapeutic interventions and in our understanding of cortical dynamics. Our findings show that a solid understanding of the underlying mechanisms can be leveraged to create predictive models of responses to stimulation. We hope that this will allow for the design of specific activation paradigms

tailored to clinical and research applications. This principled approach contrasts with the current reliance on trial and error, offering a more systematic and effective means of optimizing microstimulation protocols. Continued refinement in our understanding of these mechanisms is critical both for advancing our understanding of cortical systems and ultimately for improving outcomes in cases of neurological and neuropsychiatric disorders.

3.4 METHODS

3.4.1 EX VIVO EXPERIMENTS

Animal procedures All animal procedures were carried out in compliance with NIH guidelines for animal research and were approved by the Massachusetts Institute of Technology Committee on Animal Care. Adult (2-6 months of age) C57BL/6 (Charles River Laboratories) or PV-tdTomato (Jax #027395) of both sexes were used. All animals were kept in conventional social housing with food and water ad libitum on a 12-h light/dark cycle.

Slice preparation Sucrose-based artificial cerebrospinal fluid (aCSF) was used for slicing; it was composed of (in mM): 1.0 ascorbate, 1.1 calcium chloride (223506, Sigma), 9.0 glucose (G7021, Sigma), 4.1 magnesium chloride (63068, Sigma), 2.7 potassium chloride (60128, Sigma), 26.2 sodium bicarbonate (S5761, Sigma), 61.1 sodium chloride (S7653, Sigma), 1.2 sodium phosphate (S9638, Sigma), 3 sodium pyruvate (P2256, Sigma), 90 su-

crose (S1888, Sigma), with an osmolality of 295-302. The sucrose solution was partially frozen and kept cold during slicing. Artificial cerebrospinal fluid (aCSF), used for recovery and recording, contained (in mM): 1.0 ascorbate, 1.2 calcium chloride, 11.3 glucose, 1.2 magnesium chloride, 1.3 potassium chloride, 24.4 sodium bicarbonate, 122.3 sodium chloride, 1.3 sodium phosphate, 3.0 sodium pyruvate, with an osmolality of 300-303. All solutions were saturated with carbogen, consisting of 95% O₂ and 5% CO₂. Mice were deeply anesthetized with 5% isoflurane in O₂ prior to decapitation. The brain was rapidly extracted, blocked, and mounted in cold sucrose solution. A vibratome (Leica VT1200S) was used to cut 300 μ m thick slices, which were transferred to aCSF for 40-60 min at 36°C. Slices were then maintained at room temperature.

Patch-clamp recording L2/3 pyramidal cells and parvalbumin positive interneurons in temporal association cortex (TeA) located 165 to 310 μ m from pial surface were recorded in aCSF at 34-36°C. The intracellular pipette solution contained (in mM): 134 potassium D-gluconate (G4500, Sigma), 6.0 potassium chloride (60128, Sigma), 10 HEPES (H3375, Sigma), 4.0 sodium chloride (S7653, Sigma), 4.0 adenosine 5'-triphosphate magnesium (A9187, Sigma), 0.3 Guanosine 5'-triphosphate trisodium salt (G8877, Sigma), 14 phosphocreatine di(tris) (sc-212557, Santa Cruz), and 0.1 Alexa 488 (A10436, Thermo Fisher). Synaptic blockers were added to the aCSF at the following concentrations: 50 μ M D-AP5 (0106, Tocris), 20 μ M DNQX (2312, Tocris), 20 μ M Gabazine (HB0901, HelloBio) and 5 μ M CGP52432 (1246, Tocris). Cells were visualized using an Olympus BX-61 microscope equipped with Dodt optics and a 60x water-immersion objective (0.9 NA, Olympus).

Whole-cell current-clamp recording was performed with a Dagan BVC-700A amplifier. Thin-wall borosilicate patch pipettes (PG52151-4, WPI) were pulled to resistances of 4-10 M Ω . Pipette capacitance was compensated after gigaseal formation. Series resistance was bridge balanced throughout the experiment (11 - 30 M Ω). The liquid junction potential was not corrected. Current signals were digitized at 20 kHz. A two-photon laser scanning system (Ultima, Bruker) equipped with a Mai Tai DeepSee laser tuned to 1000 nm was used to identify tdTomato-expressing parvalbumin cells (PV-tdTomato mouse line) prior to patching. Once a cell was selected, a bipolar stimulating electrode (PI2ST30.01B3, MicroProbes; tip separation of 75 μ m and impedance of 10k Ω), was placed at a target distance from the soma (50, 100, 250 or 750 μ m). Pyramidal cells were situated at 63.8 ± 13 μ m (n = 6), 107.9 ± 13 μ m (n = 6), 272.2 ± 24 μ m (n = 5), 727.0 ± 43 μ m (n = 5) from the bipolar electrode. PV cells were situated at 60.8 ± 14 μ m (n = 6), 95.8 ± 20.1 μ m (n = 5), 248.5 ± 17 μ m (n = 5), 749 ± 23 μ m (n = 6) from the bipolar electrode. Somatic current injections were performed to measure input resistance (pyramidal: 30 to 120 M Ω , PV: 38-105 M Ω) and spiking properties. Electrode stimulation protocols consisted of ten sweeps. Each sweep lasted 2 seconds and was composed of a 100 ms baseline followed by a single 90 μ s-pulse with a stimulation intensity ranging from 0.01 to 1 mA, delivered by a stimulation isolator (ISO-Flex, AMPI). The order of the stimulation intensities was randomized, and each sweep was separated by 5 seconds. Each protocol was run without and with synaptic blockers. For axotomy experiments, the axon was severed proximal to the first axonal bifurcation with repeated laser (820 nm) linescans. The above protocols were then run with synaptic blockers.

Patch-clamp analysis For each stimulation intensity and cell, only responses within the 100 ms time-window following the onset of stimulation were analyzed. All voltage recordings were baseline-subtracted using the average voltage of the 100 ms time-window before stimulation. A threshold of three-standard deviations above the baseline was used as the method to detect evoked responses. The peak response was extracted from the 2 kHz-filtered average of baseline-subtracted trials without spikes. The spiking probability was computed as the proportion of trials with spikes among all trials. Response onset was defined as the first time point for which the baseline-subtracted voltage was three times higher than the standard deviation of the baseline. The time to peak was defined as the time between the onset of stimulation and the peak of the response. Antidromic spikes were differentiated from orthodromic spikes based on their onset threshold. For each cell, antidromic spike threshold was at least 10 mV lower than for orthodromic spike.

Except for Figure 3.3l, an exponential function was used as follows:

$$f(x) = Ae^{-Bx} \quad (3.2)$$

with A and B as free parameters to be fitted. For Figure 3l, a Gaussian function convolved with a rectangular function was used as follows:

$$h(x) = \frac{1}{2} \left[\operatorname{erf} \frac{x+A}{\sigma\sqrt{2}} - \operatorname{erf} \frac{x-A}{\sigma\sqrt{2}} \right] \quad (3.3)$$

where $\operatorname{erf}(\cdot)$ is the error function. a and σ are free parameters.

Statistics Statistical analyses were performed using R Studio (Foundation for Statistical Computing, Vienna, Austria). Linear Mixed Model (*lmer* function from *lmerTest* library) and Tukey-corrected pairwise comparisons were used to test significance of results. Cells were considered as random effects.

3.4.2 ANALYTICAL METHODS

Electrical stimulation of an excitable cable Electrical stimulation changes the electric field around neurons. We model the external field $V_e(x, t)$ along a dendrite or axon such that t is time and x is the position along a cable [78, 148, 181]. This external field changes the difference in potential between the inside and the outside the cell $V(x, t)$, leading to transverse and axial currents in the cable:

$$\tau \frac{\partial V}{\partial t} = -V + \lambda^2 \frac{\partial^2 V}{\partial x^2}. \quad (3.4)$$

Using this cable equation and assuming that the axial currents dominate the flow of charges inside the cable for short inputs ($-V \ll \lambda^2 \frac{\partial^2 V}{\partial x^2}$), a stimulation of short duration T will lead to a membrane potential that results from axial current term:

$$V_{\text{stim}} \equiv V(x, T) = \frac{\lambda^2 T}{\tau} \frac{\partial^2 V_e}{\partial x^2}. \quad (3.5)$$

The initial depolarization is expected to scale with the electrotonic length constant λ of the cable, the duration of the stimulation, and to be inversely proportional to the

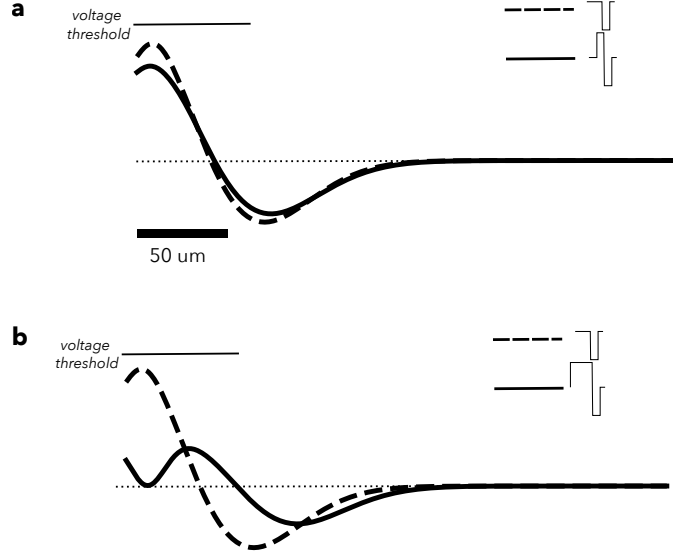


Figure 3.7: Application of the theory to biphasic pulses. **a** The axonal depolarization immediately after the stimulation is shown as a function of distance along the axon. The depolarization after a monopolar pulse (cathodal) is shown with a dashed line and that after a biphasic pulse (anodal then cathodal) is shown with a full line. **b** Same as a but increasing the time of the anodal pulse.

membrane time constant, τ .

If the cable is excitable, and if the curvature of the field has a minimum around at x^* , and if we can assume Ohmic conversion between current and the field, $V_e = -R_e I F(x)/d^p$ where R_e is the resistance of the electrode and $F(x)$ captures the shape of the field with p the dominant term of its multipolar expansion, and the minus sign is added to capture cathodal stimulation. Then the smallest current for which the cable will elicit an action potential is given by $V(x^*, t) = V_T$ such that

$$V_T = -\frac{\lambda^2 T R_e I_T C}{\tau d^p}, \quad (3.6)$$

where V_T is the difference between the resting potential and the threshold potential and $C = \left. \frac{\partial^2 F}{\partial x^2} \right|_{x=x^*}$ is the curvature of the field at x^* . Hence

$$I_T = -\frac{\tau V_T d^p}{\lambda^2 C T R_e}. \quad (3.7)$$

This equation describes how a cable excitability to external stimulation will be greater when there is a greater electrotonic space constant, a longer duration of stimulation, a smaller curvature of the field, a smaller height of the threshold, and a smaller membrane time constant.

Biphasic stimulation Since wider cables will have larger electrotonic length constant and faster propagation, this implies that the thicker cables will be more excitable [78], and thus that the faster cables will be more excitable [150]. Furthermore, since the electronic length constant of myelinated fibers is much larger than that of unmyelinated fibers, this description is consistent with the observation that myelinated fibers are more excitable [151].

Throughout the manuscript, the theory is exposed for monopolar pulses. It is fairly simple to expand the predictions for biphasic pulses (Fig. 3.7). The logic is to separate the problem in two steps. First one calculates the response to the first pulse using Eq. 3.5 with V_e of the chosen sign (anodal or cathodal) to capture the response to the first pulse. Then to capture the second pulse one uses Eq. 3.5 again but by replacing V_e by the difference between V_{stim} from the first pulse and V_e from the external stimulation with the

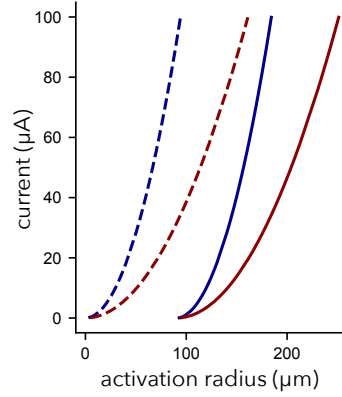


Figure 3.8: Relationship between activation radius of axon and soma. Average activation radius of axons based on distribution of axonal myelination/diameters (Fig. 3.3) of PV-positive (red, dashed line) and pyramidal cells (blue, dashed line) as a function of increasing stimulation intensity. The activation radius of somas of the corresponding cell types is shown with the full line. The somewhat unorthodox choice of the dependent variable on the abscissa was made to echo Stoney et al. (1968) [68].

appropriate sign (anodal or cathodal).

Passive cable propagation After the stimulation, in case no action potential was emitted, we consider the passive dynamics of the potential $V(x, t)$ that results from axial and transverse currents. The dynamics are captured by the cable equation (Eq. 3.4), which can be solved using a Green's function

$$V(x, t) = \iint \Phi(\tilde{x} - \tilde{x}', \tilde{t} - \tilde{t}') V_{\text{stim}}(\tilde{x}') \delta(\tilde{t}') d\tilde{x}' d\tilde{t}'. \quad (3.8)$$

For an infinite cable, the Green's function $\Phi(x, t) = \Theta(t) \exp(-t/\tau - x^2/4(t/\tau)\lambda^2) / \sqrt{4\pi t/\tau}$. The resulting diffusion integral can be easily integrated numerically.

Strength-duration curves Long input durations are thought to trigger axons through a distinct mechanism since the rheobase I_∞ associated with long inputs is independent of axonal properties like propagation speed [150]. We therefore assume that the response to a sustained stimulation is dominated by transverse currents. Accordingly, we consider a stationary solution to the cable equation assuming a localized input at $x = 0$ and an infinite cable is given by Ref. [182]

$$V(x) = R_e I d^{-p} e^{-|x/\lambda|} \quad (3.9)$$

where R_e relates to the steady-state resistance of electrode and external media. Solving for I that reaches a threshold at $x = 0$, we find the rheobase:

$$I_\infty = \frac{V_T d^p}{R_e}. \quad (3.10)$$

Together, we obtain an equation that interpolates between the short T and the large T as

$$I_T = I_\infty \left(1 + \frac{T_C}{T} \right) \quad (3.11)$$

such that at small T we get $I_T \approx I_\infty T_C/T$ and at large T we get $I_T = I_\infty$. The chronaxie T_C is

$$T_C = \frac{\tau}{\lambda^2 C}. \quad (3.12)$$

where C is the curvature of the field at its minimum.

Eqs. 3.11 and 3.12 describe how the chronaxie is proportional to the inverse of the electrotonic length. Since the speed of passive propagation is also proportional to the electrotonic length constant, this implies that the chronaxie is inversely proportional to the speed of propagation, as observed in [150, 183]. Furthermore, as myelination increases considerably the length constant, this implies that the chronaxie of myelinated fibers will be smaller, as observed experimentally [151].

Activated radius with homogeneous cables In a homogeneous population, each cable is captured by the same electrotonic length, membrane time constant. Hence, for a given stimulation current I , the axons will be activated within a volume defined by the distance d for which the AP threshold is reached (as in Eq. 3.7 but varying distance for a fixed current):

$$a(I)^p = \frac{\lambda^2 C I T R_e}{\tau V_T}, \quad (3.13)$$

where a is the distance d for which $V = V_T$. This defines the radius of the volume of activated axons.

Heterogeneity of axonal cable properties For each cell type considered, we modeled axon-to-axon variations in diameters, myelination, and membrane time constant. The diameters, D , were sampled from a log-normal distribution

$$p(D) = \frac{1}{sD\sqrt{2\pi}} \exp\left(-\frac{\log^2(D - \mu_D)}{2s^2}\right) \quad (3.14)$$

where s is the scale parameter and μ_D is the location parameter. These parameters (s, μ_D) were $(0.1 \mu\text{m}, 0.1)$ for pyramidal cells, $(0.1 \mu\text{m}, 0.15)$ for PV-positive cells, and $(0.05 \mu\text{m}, 0.1)$ for SST-positive cells. The membrane time constants were sampled from a Gaussian distribution. The mean and standard deviation (μ_τ, σ_τ) were $(20, 3)$ ms for pyramidal cells, $(10, 3)$ ms for PV-positive cells, and $(30, 3)$ ms for SST-positive cells. For PV-positive cells, we modeled proximal myelination for diameters larger than 0.6 by increasing λ^2 five-fold. Given that myelinated fibers are estimated to have 10-100 times larger electrotonic length [184], this is likely an underestimation.

Activated radius with heterogeneous cables This heterogeneity of cable properties implies a heterogeneity of excitability according to Eq. 3.7. Using the resistivity of the membrane (ρ_m) and the cytoplasm (ρ_i) we can use the definition of the electrotonic length $\lambda^2 = D\rho_m/4\rho_i$ in Eq. 3.7 to make the relationship explicit such that the axon will fire if the stimulating current I is sufficiently large to satisfy

$$\frac{I}{d^p} > \alpha \frac{\tau}{DT} \quad (3.15)$$

where $\alpha = \frac{4V_T\rho_i}{R_e\rho_m} \left(\frac{\partial^2 F}{\partial x^2} \right)^{-1}$ is a constant that captures the curvature of the field on the path of the axon, the electrode resistance, the separation between resting potential and the action potential threshold as well as the resistivities of the membrane and cytoplasm. Establishing the value of this constant with precision is difficult because of the need to know electrode resistance and the path of the axon across the electric field. In effect, not

knowing this constant implies that we cannot predict precisely the correct scale of the current. Since our work is focused on the relative activation of different cell types, and since we can calibrate to the activation threshold, we have chosen an arbitrary value of $\alpha = .3$ in our simulations. However, we note that it is possible that this constant shows variations across cell types, especially if the axons have a different average orientation across cell types (affecting the curvature factor).

At any given stimulation current, and using Eq. 3.15 across 50 000 samples of the diameters and time constants, we can compute the probability of an axon being activated as a function of distance by calculating the fraction of axons that are above threshold at each distance such that the radius $a(I) = \text{argmin}_d |P(D > \alpha \tau d^p / IT) - 0.3|$, is the distance for which the probability that an axon is activated decays to a value of 0.3t. From this numerical quantity, we can establish the firing radius by calculating the distance at which the firing probability reaches a value of one third. Together, this produces an estimate of the radius of activated axons, $a(I)$ (Fig. 3.8, dashed lines, using $p = 2$). Together, Eq. 3.15 and can be used to compute the ratio of directly recruited excitation over inhibition for axons

$$\frac{r_E(t = 0, x = d)}{r_I(t = 0, x = d)} = \frac{P(D > \alpha_E \tau_E d^p / IT)}{P(D > \alpha_I \tau_I d^p / IT)} \quad (3.16)$$

where we have used indices for excitation and inhibition, allowing α to depend on cell types although this is assumed to be cell type invariant in the simulations $\alpha_E = \alpha_I = \alpha$.

Spread of antidromic activation Given our estimate of the activated radius for axons, we model the resulting activation of somas distributed in space. For this we use a heuristic approach: we consider that the probability that a cell body has its axon passing in the vicinity of the electrode follows a Gaussian probability distribution over distance with standard deviation σ_{ax} . Given that the activation profile of axons has a relatively sharp threshold at the activation radius, we thus model the activation probability of somas as the convolution between a rectangular and Gaussian function (Eq. 3.3). In order to ensure that the probability of the somatic activation is one when the soma is within the activated volume of axons, we chose $A = a(I) + 2\sigma_{ax}$. This activated volume of somas can be used to produce the intensity dependent activation radius (Fig. 3.8 and the cell-type dependent activation profile (Fig. 3.3f).

Speed of axonal propagation Based on Ref. [168], we model the velocity of action potential propagation along an axon having membrane time constant τ and electrotonic length λ as $v \propto \lambda/\tau$.

3.4.3 FIELD MODEL

Refractory Field Model With Spatial Dependence We consider a field model of microstimulation with a small number of parameters to study whether and how a transient propagating front can be produced from initial conditions analogous to microstimulation that preferentially activates inhibitory neurons. We consider linear dynamics with spa-

tially dependent connectivity, the rates of excitation and inhibition $r_E(x, t)$ and $r_I(x, t)$ are defined as follows:

$$\tau_E \frac{\partial r_E}{\partial t} = -r_E - R_{ref} + K * (w_{EE} r_E - w_{IE} r_I) \quad (3.17)$$

$$\tau_I \frac{\partial r_I}{\partial t} = -r_I + K * (w_{EI} r_E - w_{II} r_I) \quad (3.18)$$

where $*$ denotes spatial convolution and we clip rates below zero (preventing negative rates) and above 1 (defined as the maximum possible firing rate at each location). Weights are defined as $w_{pre,post}$, for fig. 3.4 we set $w_{EE} = 2.8$, $w_{EI} = 2.0$, $w_{IE} = 0.27$, $w_{II} = 0.8$, and $\tau_E = \tau_I = 10$ ms. The connectivity kernel $K(x - x')$ maps activity at x' to activity at x and is chosen as a Gaussian with standard deviation $\sigma_{conn} = 115 \mu\text{m}$:

$$K(x - x') = \frac{1}{\sigma_{conn} \sqrt{2\pi}} \exp\left(-\frac{(x - x')^2}{2\sigma_{conn}^2}\right). \quad (3.19)$$

$R_{ref}(x, t)$ captures decreased firing ability of the excitatory population due to refractoriness. Because the inhibitory fast spiking interneuron population is expected to recover more rapidly from stimulation we adapt the approach of [175] by defining $R_{ref}(x, t)$ to capture the difference in the refractoriness of the two populations this term evolves with the following dynamics:

$$\tau_{ref} \frac{\partial R_{ref}}{\partial t} = \beta r_E - R_{ref} \quad (3.20)$$

where β describes the strength of the refractory effect and τ_{ref} describes its time course.

For fig. 3.4 we set $\tau_{ref} = 3$ ms and $\beta = 0.65$. In order to demonstrate the impact of refractoriness wave generation and propagation we show how the model behaves when R_{ref} is fixed to a value of 0 in fig. 3.9a-c (decaying, limited expansion) and 3.9d-f (non-decaying expansion). In order to remain in a decaying activity regime where activity eventually returned to rest after the stimulation response fig. 3.9a-c weight parameters driving recurrent excitation had to be weakened and those for inhibition were strengthened to $w_{EE} = 1.8$, $w_{EI} = 2.4$, $w_{IE} = 0.8$, $w_{II} = 0.8$. In contrast, 3.9d-f used the same weight values as 3.4 and illustrates how expansion of activity in the model results in non-decaying activity in the absence of refractoriness. 3.9g-i present the responses from fig. 3.4 again for ease of comparison.

Initial Conditions In order to simulate the network response to stimulation, we describe the direct effect of stimulation in cortex using a computationally tractable phenomenological equation calibrated for consistency with our empirical observations (Figs. 3.3i and 3.3k). Consistent with our results in Fig. 3.2 and existing literature [140] we model the direct response to microstimulation in terms of the axonal response to stimulation. We describe the probability that any given neuron has a portion of its axon sufficiently close to the electric field to be driven above its firing threshold $V_T = 20$ mV as the product of a normalized zero centered Gaussian function $g(x|\mu = 0, \sigma_{stim})$ describing the spatial distribution of axons and the log of input intensity I . This allows us to define state vectors of initial firing rates $\mathbf{r_E}(t = 0)$ and $\mathbf{r_I}(t = 0)$:

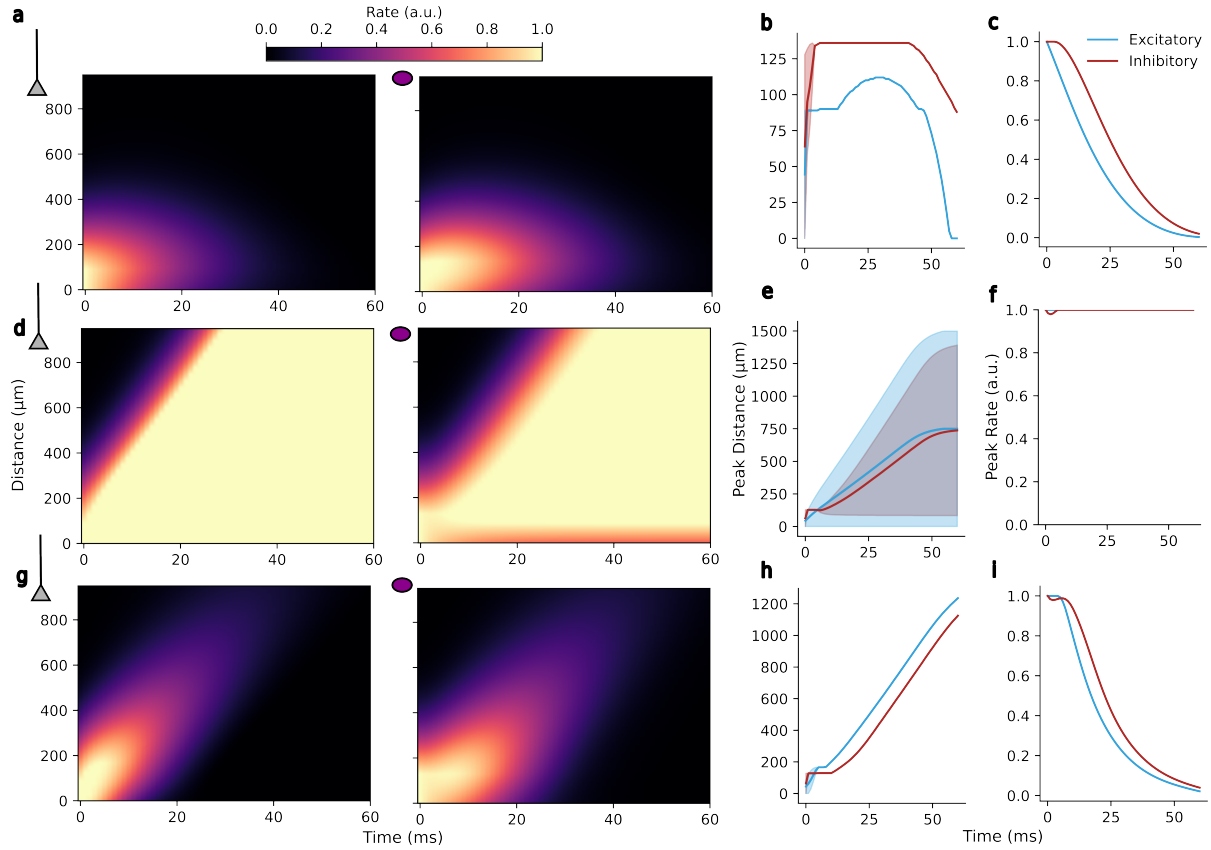


Figure 3.9: Comparison of field models with and without refractory variable. **a-c** Correspond to a field model without a refractory term calibrated for expansion with decaying activity. **a** Heatmaps for excitatory neuron (left) and inhibitory neuron (right) firing rate over space and time in response to a microstimulation event at $t = 0$ ms. **b** The location of the highest activity is shown for each point in time in terms of distance from the stimulating electrode for excitatory neurons (blue) and inhibitory neurons (red). In cases where the firing rate is maximal at multiple points in space, the solid line indicates the midpoint of these locations and all maximal firing rate locations are indicated with shading. **c** The rate at the location with the highest firing rate at each point in time is shown in terms of distance from the stimulating electrode for excitatory neurons (blue) and inhibitory neurons (red). **d-f** Correspond to a field model without a refractory term calibrated for non-decaying expansion. **d** shows heatmaps as in **a** but for the non-decaying case, **e** shows peak locations as in **b** but for the non-decaying case, and **f** shows peak amplitudes as in **c** but for the non-decaying case. **g-i** Show the equivalent plots to **d-f** for the equivalent model configuration but with a refractory variable added to the model.

$$\mathbf{r}_{\mathbf{E}}(t = 0) = \min \left(1.0, \gamma_E \ln \left(1 - \frac{I - I_{min}}{I_{min}} \right) g(\mathbf{x}_{\mathbf{E}} \mid 0, \sigma_{stim}^2) \right) \quad (3.21)$$

$$\mathbf{r}_{\mathbf{I}}(t = 0) = \min \left(1.0, \gamma_I \ln \left(1 - \frac{I - I_{min}}{I_{min}} \right) g(\mathbf{x}_{\mathbf{I}} \mid 0, \sigma_{stim}^2) \right) \quad (3.22)$$

Where $\sigma_{stim} = 210 \mu\text{m}$, $\gamma_E = 245$, and $\gamma_I = 270$ are type-specific scaling parameters and a threshold is applied preventing probability values above 1. This representation also makes it possible to consider cases where class specific activation probabilities differ in terms of intensity or spatial profile as a result of differences in spatial structure or axon diameter [80, 140] by varying the ratio of γ_I and γ_E or σ_E^{stim} and σ_I^{stim} . We note that this function reconciles the long range, sparse activation observed in [69] and the localized, intensity dependent scaling of activation observed in [68]. We further augment this equation to adjust the minimum required intensity to trigger nontrivial firing activity by adding a minimum intensity threshold $I_{min} = 19 \mu\text{A}$.

Simulation With the initial conditions defined, we numerically generated the values in figs. 3.4 and 3.9 by combining the forward Euler method with an exponential integrator for the leak and leak like terms. These simulations were performed using Numpy version 2.0.2, Pytorch version 2.11.0 (CPU), and plotted with Matplotlib version 3.10.0.

3.4.4 SPIKING NEURAL NETWORK MODEL

The results in fig. 3.5 describe the stimulus response of a one dimensional spiking model where neuron locations are defined along a single lateral dimension over distances up to $2000\ \mu m$ from a point source electrode. We assume a Dalean network [159, 185–187] where neurons are either strictly excitatory or inhibitory.

Initialization We begin by considering the case of direct electric stimulation delivered via a single pulse from a single point in the cortical slice. First, we define excitatory and inhibitory neuronal populations of $n_E = 11700$ and $n_I = 1300$ neurons (a 9 to 1 ratio) with time dependent membrane voltages described by the state vectors $\mathbf{v}_E(t)$ and $\mathbf{v}_I(t)$. Each neuron is assigned a position denoted by the state vectors $\mathbf{x}_E$ and $\mathbf{x}_I$. These positions are drawn from a uniform distribution spanning a single spatial dimension defined relative to the position of the stimulating electrode.

Initially, we define the membrane voltages of the slice neurons not directly activated by the stimulation to be near rest.¹ To allow for a computational implementation compatible with backward optimization and to represent a subthreshold component of the response to electrical stimulation, stimulation driven firing rates were transformed into a direct change in membrane potential by applying a Relaxed Bernoulli function $\mathcal{RB}$ (as defined for Pytorch version 2.10 [188]) with temperature $\zeta = 1.1$ to the rate over the full duration T of a short stimulation pulse. We sample from this function to generate activity vectors

¹For brevity, we define the resting membrane potentials as 0 (i.e. $v_{rest} = 0$) this does not change our results in the general case and is re-scaled -70mV later.

u_E and u_I which we then scale to membrane thresholds with an additional scaling factor $\eta = 1.2$ to set initial membrane voltages $\mathbf{v}_E(t = 0)$ and $\mathbf{v}_I(t = 0)$:

$$\mathbf{u}_E \stackrel{\text{i.i.d}}{\sim} \mathcal{RB}(\mathbf{r}_E(t = 0)T) \quad (3.23)$$

$$\mathbf{u}_I \stackrel{\text{i.i.d}}{\sim} \mathcal{RB}(\mathbf{r}_I(t = 0)T) \quad (3.24)$$

$$\mathbf{v}_E(t = 0) = \eta V_T \mathbf{u}_E \quad (3.25)$$

$$\mathbf{v}_I(t = 0) = \eta V_T \mathbf{u}_I \quad (3.26)$$

Where $\mathbf{r}_E(t = 0)$ is defined as in equation 3.21 and $\mathbf{r}_I(t = 0)$ is defined as in equation 3.22

Transition rules Following the initial electrical stimulation, neuron voltages evolve according to standard leaky-integrate and fire dynamics [182] vectorized for efficient simulation [76]:

$$\tau_{mem} \frac{d\mathbf{v}_E}{dt}(t) = -\mathbf{v}_E(t) + R_E \mathbf{I}_E(t) \quad (3.27)$$

$$\tau_{mem} \frac{d\mathbf{v_I}}{dt}(t) = -\mathbf{v_I}(t) + R_I \mathbf{I_I}(t) \quad (3.28)$$

$$\tau_{syn} \frac{d\mathbf{I_E}}{dt}(t) = -\mathbf{I_E}(t) + \mathbf{S_E}(t) \mathbf{W_{EE}} - \mathbf{S_I}(t) \mathbf{W_{IE}} \quad (3.29)$$

$$\tau_{syn} \frac{d\mathbf{I_I}}{dt}(t) = -\mathbf{I_I}(t) + \mathbf{S_E}(t) \mathbf{W_{EI}} - \mathbf{S_I}(t) \mathbf{W_{II}} \quad (3.30)$$

Where $\tau_{mem} = 10$ ms, $\tau_{syn} = 5$ ms, and $\mathbf{S_E}$ and $\mathbf{S_I}$ are vectors containing the spike trains of excitatory and inhibitory neurons from the circuit. We treat these outputs as discrete spikes arising from the Heaviside step function such that for an arbitrary neuron:

$$S(t) = \Theta(V(t) - V_T) \quad (3.31)$$

where Θ denotes the Heaviside step function. Each spike is also accompanied by a voltage reset to the rest potential. After firing, neurons are then gated by an absolute refractory period of 5ms for inhibitory neurons and 8ms for excitatory neurons during which they cannot spike again.

We define the weight matrices $\mathbf{W_{EE}}$, $\mathbf{W_{EI}}$, $\mathbf{W_{IE}}$, $\mathbf{W_{II}}$ to be strictly Dalean² sparse matrices with distance dependent class specific connectivity matrices $\mathbf{C^{EE}}$, $\mathbf{C^{EI}}$, $\mathbf{C^{IE}}$, and $\mathbf{C^{II}}$ defined as follows:

²Neurons can be either strictly excitatory or strictly inhibitory

$$\mathbf{C}_{ij}^{\mathbf{EE}} \stackrel{\text{i.i.d}}{\sim} \mathcal{B}(\min(1.0, A_{EE}g((\mathbf{x}_{\mathbf{E}})_i - (\mathbf{x}_{\mathbf{E}})_j \mid 0, \sigma_{syn}^2))) \quad (3.32)$$

$$\mathbf{C}_{ij}^{\mathbf{EI}} \stackrel{\text{i.i.d}}{\sim} \mathcal{B}(\min(1.0, A_{EI}g((\mathbf{x}_{\mathbf{E}})_i - (\mathbf{x}_{\mathbf{I}})_j \mid 0, \sigma_{syn}^2))) \quad (3.33)$$

$$\mathbf{C}_{ij}^{\mathbf{IE}} \stackrel{\text{i.i.d}}{\sim} \mathcal{B}(\min(1.0, A_{IE}g((\mathbf{x}_{\mathbf{I}})_i - (\mathbf{x}_{\mathbf{E}})_j \mid 0, \sigma_{syn}^2))) \quad (3.34)$$

$$\mathbf{C}_{ij}^{\mathbf{II}} \stackrel{\text{i.i.d}}{\sim} \mathcal{B}(\min(1.0, A_{II}g((\mathbf{x}_{\mathbf{I}})_i - (\mathbf{x}_{\mathbf{I}})_j \mid 0, \sigma_{syn}^2))) \quad (3.35)$$

Where $g(\mathbf{x} \mid \mu = 0, \sigma_{syn}^2)$ again denotes a normalized Gaussian function centered at zero that we scale with class specific scaling parameters $A_{EE} = 16$, $A_{IE} = 38$, $A_{EI} = 30$ and $A_{II} = 39$ respectively and $\sigma_{syn} = 115\mu m$. These values are threshold to a 0 to 1 range after which we sample from Bernoulli distributions $\mathcal{B}$ to define connections (i.e. nonzero entries in the weight matrices).

We note that while our treatment considers a single spatial dimension for simplicity, the distance values for both connectivity and the direct effect of stimulation are simple Euclidean distances. This means that arbitrarily shaped networks in any number of spatial dimensions can be defined and encoded into this matrix form without incurring additional computational cost at run time (for an equivalent number of neurons). We sample synaptic efficacies (connection weights) from neuron type specific half-normal distributions:

$$\mathbf{E}_{\mathbf{EE}} \stackrel{\text{i.i.d}}{\sim} \beta HalfNorm(\sigma_E^{eff}) \quad (3.36)$$

$$\mathbf{E}_{\mathbf{EI}} \stackrel{\text{i.i.d}}{\sim} \beta \text{HalfNorm}(\sigma_E^{eff}) \quad (3.37)$$

$$\mathbf{E}_{\mathbf{IE}} \stackrel{\text{i.i.d}}{\sim} \beta \text{HalfNorm}(\sigma_I^{eff}) \quad (3.38)$$

$$\mathbf{E}_{\mathbf{II}} \stackrel{\text{i.i.d}}{\sim} \beta \text{HalfNorm}(\sigma_I^{eff}) \quad (3.39)$$

Where $\mathbf{E}_{\mathbf{EE}}$, $\mathbf{E}_{\mathbf{EI}}$, $\mathbf{E}_{\mathbf{IE}}$, and $\mathbf{E}_{\mathbf{II}}$ are matrices of efficacies (strengths) for all possible connections, $\text{HalfNorm}()$ is a half-normal sampling distribution parameterized by $\sigma_E^{eff} = 0.4$ and $\sigma_I^{eff} = 0.6$ respectively. These distributions are intended to correspond to observed efficacy distributions from [174], with inhibitory connections being more powerful. Finally, β is a scaling parameter that allows the strength of the weight matrix to be increased to compensate for cases when the number of simulated neurons is low relative to the spatial scale. We set $\xi = 3$ in our one dimensional simulations to compensate for the low number of neurons relative to distance. Finally, we define $\mathbf{W}_{\mathbf{EE}}$, $\mathbf{W}_{\mathbf{EI}}$, $\mathbf{W}_{\mathbf{IE}}$, $\mathbf{W}_{\mathbf{II}}$ by taking the Hadamard products $\odot$ of corresponding connectivity and efficacy matrices:

$$\mathbf{W}_{\mathbf{EE}} = \mathbf{E}_{\mathbf{EE}} \odot \mathbf{C}^{\mathbf{EE}} \quad (3.40)$$

$$\mathbf{W}_{\mathbf{EI}} = \mathbf{E}_{\mathbf{EI}} \odot \mathbf{C}^{\mathbf{EI}} \quad (3.41)$$

$$\mathbf{W}_{\text{IE}} = \mathbf{E}_{\text{IE}} \odot \mathbf{C}^{\text{IE}} \quad (3.42)$$

$$\mathbf{W}_{\text{II}} = \mathbf{E}_{\text{II}} \odot \mathbf{C}^{\text{II}} \quad (3.43)$$

Generalization to Multiple Spatial Dimensions As we have noted above, equivalent connection probabilities to those in equations 3.32-3.35 can be calculated for networks with arbitrary spatial dimensions and geometries so long as the distance between neurons can be calculated. Similarly, the direct response to stimulation defined in equations 3.23-3.26 can be calculated as long as the distance between neurons and the electrode can be calculated. In fig. 3.10 and supplementary video 1, we do this to simulate a two-dimensional network defined in a 2mm by 2mm space with a point electrode in the center by calculating Euclidean distances between neurons and between neurons and the electrode. That network contained 105,300 excitatory neurons and 11700 inhibitory neurons which were uniformly distributed along both spatial dimensions. Because this neuron count was higher than in the one dimensional model we did not need to scale connection strength and set $\xi = 1$.

Simulation We numerically generated the values in fig. 3.5, 3.10, and supplementary video 1 by combining the forward Euler method with an exponential integrator for the leak and leak like terms. In order to perform these simulations, we developed a new package for simulation of dynamics in spatially structured spiking neural networks built

with partial inspiration from [76], and [158] using Pytorch version 2.7.0 (GPU), Numpy version 1.26.4, and Scipy version 1.11.4. Figures were plotted with Matplotlib version 3.10.7. All spiking network simulations were performed using an NVIDIA L40S graphics processing unit (GPU) running CUDA version 12.8. A reference upper bound for run time is that for a multi-dimensional network with a size of 400,000 neurons, 1000 ms of stimulus response simulation generated in 1 ms time steps required approximately four minutes of GPU run time.

3.4.5 TABLE OF SYMBOLS, SECTIONS 3.4.1 AND 3.4.2

symbol	quantity	value	units
$f(x)$	function	function	a.u.
$h(x)$	function	function	a.u.
I	current	variable	μA
I_T	current at activation threshold	Variable	μA
λ^2	electrotonic length of a cable	Variable	μm
τ	membrane time constant	Variable	ms
τ_E	membrane time constant (excitatory cells)	Variable	ms
τ_I	membrane time constant (inhibitory cells)	Variable	ms
C	curvature of field along an axon	Variable	a.u.
T	duration of a short stimulating pulse	Variable	μs
I_∞	current at rheobase	Variable	μA
d	distance between the axon and the electrode	Variable	μm
p	dominant term in electrical multipolar expansion	2 (bipolar)	a.u.
A	free parameter	variable	a.u.
a	activated radius	variable	μm
B	free parameter	variable	a.u.
σ	fitted parameter for gaussian function	variable	a.u.
$\text{erf}(\cdot)$	error function	function	NA
V_e	voltage of external field	variable	mV
V	potential difference between inside and outside of a cell	function	mV
V_T	voltage threshold	20.0	mV
x	distance	variable	μm
t	time	variable	ms
V_{stim}	membrane potential from axial current following short stimulation	variable	mV
x^*	location along cable at which the field has minimum curvature	variable	μm
R_e	electrode resistance	variable	m Ω
$\Phi(x, t)$	Green's function	function	a.u.
$\delta(\cdot)$	Dirac delta function	function	a.u.
$\Theta(\cdot)$	Heaviside step function	function	a.u.
$G(x \mu, \sigma^2)$	non-normalized Gaussian function	function	a.u.
T_C	chronaxie	variable	μs
D	axon diameter	variable	μm
s	scale parameter	PyrNs: 0.1, PV INs: 0.1, SOM: 0.05	μm
μ_D	location parameter	PyrNs: 0.1, PV INs: 0.15, SOM: 0.1	a.u.
μ_τ	mean of membrane time constant sampling distribution	PyrNs: 20.0, PV INs: 10.0, SOM: 30.0	ms
σ_τ	SD of membrane time constant sampling distribution	PyrNs: 3.0, PV INs: 3.0, SOM: 3.0	ms
ρ_m	membrane resistivity	variable	Ωm
ρ_i	cytoplasm resistivity	variable	Ωm
$\alpha, \alpha_E, \alpha_I$	constant	0.3	a.u.
σ_{ax}	standard deviation of axon proximity distribution	variable	a.u.
$P_{ax}(x)$	probability of axon firing at x	function	a.u.
$\rho(I)$	activation radius of cell bodies for stimulus intensity I	variable	μm
v	velocity of action potential propagation	variable	ms ⁻¹

Table 3.1

3.4.6 TABLE OF SYMBOLS, SECTION 3.4.4

symbol	quantity	value	units
n_E	number of cells in excitatory population	variable	a.u.
n_I	number of cells in inhibitory population	variable	a.u.
$\mathbf{v}_E(t)$	state vector of excitatory neuron voltages	vector	mV
$\mathbf{v}_I(t)$	state vector of inhibitory neuron voltages	vector	mV
$\mathbf{x}_E$	state vector of excitatory neuron locations	vector	μm
$\mathbf{x}_I$	state vector of inhibitory neuron locations	vector	μm
$\mathbf{r}_E(t=0)$	vector of stimulus driven excitatory neuron firing rates	vector	Hz
$\mathbf{r}_I(t=0)$	vector of stimulus driven inhibitory neuron firing rates	vector	Hz
$g(x \mu=0, \sigma)$	normalized zero centered Gaussian function	function	a.u.
γ_E	scaling parameter	245.0	a.u.
γ_I	scaling parameter	270.0	a.u.
σ_{stim}	free parameter describing the spatial distribution of axons	210.0	μm
I_{min}	minimum required intensity to trigger nontrivial firing activity	19.0	mV
$\mathcal{RB}(P)$	Relaxed Bernoulli	distribution	a.u.
ζ	temperature of Relaxed Bernoulli	1.1	a.u.
η	scaling factor	1.2	a.u.
$\mathbf{u}_E$	vector of unscaled excitatory response samples	vector	a.u.
$\mathbf{u}_I$	vector of unscaled inhibitory response samples	vector	a.u.
τ_{syn}	synaptic time constant	5.0	ms
τ_{mem}	membrane time constant	5.0	ms
R_E	excitatory synaptic input resistance	variable	$m\Omega$
R_I	inhibitory synaptic input resistance	variable	$m\Omega$
$\mathbf{I}_E(t)$	vector of synaptic input currents to excitatory neurons at t	vector	mA
$\mathbf{I}_I(t)$	vector of synaptic input currents to inhibitory neurons at t	vector	mA
$\mathbf{S}_E(t)$	binary vector of excitatory neuron spikes at t	vector	binary
$\mathbf{S}_I(t)$	binary vector of inhibitory neuron spikes at t	vector	binary
$\mathbf{W}_{EE}$	weight matrix for excitatory neuron input to excitatory neurons	matrix	μA
$\mathbf{W}_{EI}$	weight matrix for excitatory neuron input to inhibitory neurons	matrix	μA
$\mathbf{W}_{IE}$	weight matrix for inhibitory neuron input to excitatory neurons	matrix	μA
$\mathbf{W}_{II}$	weight matrix for inhibitory neuron input to inhibitory neurons	matrix	μA
$\mathbf{C}^{EE}$	binary connectivity matrix for excitatory neuron input to excitatory neurons	matrix	binary
$\mathbf{C}^{EI}$	binary connectivity matrix for excitatory neuron input to inhibitory neurons	matrix	binary
$\mathbf{C}^{IE}$	binary connectivity matrix for inhibitory neuron input to excitatory neurons	matrix	binary
$\mathbf{C}^{II}$	binary connectivity matrix for inhibitory neuron input to inhibitory neurons	matrix	binary
A_{EE}	scaling parameter	16.0	a.u.
A_{EI}	scaling parameter	30.0	a.u.
A_{IE}	scaling parameter	38.0	a.u.
A_{II}	scaling parameter	39.0	a.u.
σ_{syn}^2	free parameter describing distance dependence of connectivity	115.0	μm
$\mathcal{B}(P)$	Bernoulli distribution	distribution	a.u.
$\mathbf{E}_{EE}$	efficacy matrix for excitatory neuron input to excitatory neurons	matrix	μA
$\mathbf{E}_{EI}$	efficacy matrix for excitatory neuron input to inhibitory neurons	matrix	μA
$\mathbf{E}_{IE}$	efficacy matrix for inhibitory neuron input to excitatory neurons	matrix	μA
$\mathbf{E}_{II}$	efficacy matrix for inhibitory neuron input to inhibitory neurons	matrix	μA
ξ	efficacy multiplier	3 (1D), 1 (3D)	a.u.
$HalfNorm(\sigma)$	half normal distribution	3 distribution	a.u.
σ_E^{eff}	parameter describing distribution of excitatory synaptic efficacies	0.4	a.u.
σ_I^{eff}	parameter describing distribution of inhibitory synaptic efficacies	0.6	a.u.
$\odot$	Hadamard product	operation	a.u.

Table 3.2

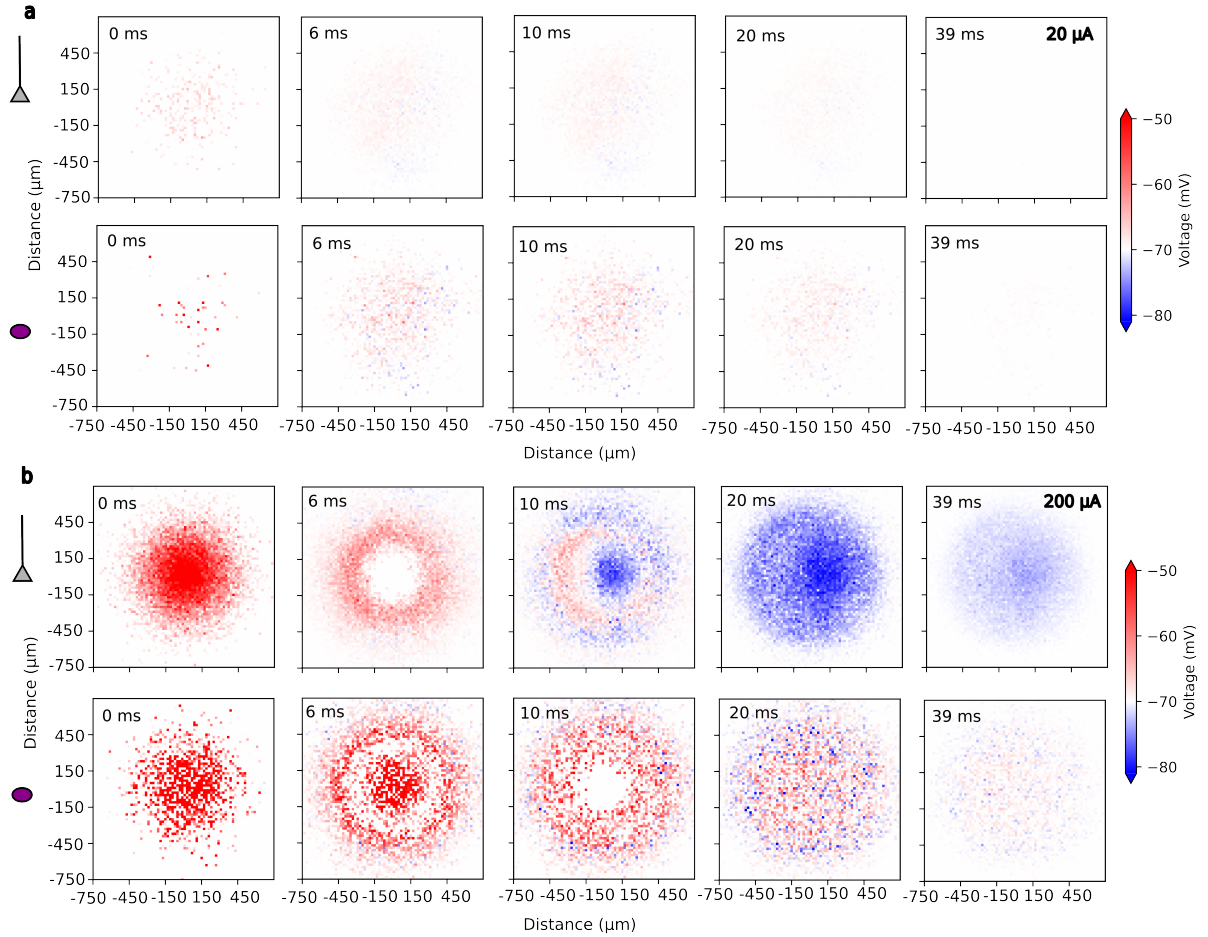


Figure 3.10: Time steps from example pulse response of 2D spiking model corresponding to supplemental animation. **a** Response to $20 \mu\text{A}$ stimulation of excitatory neurons (top row) and inhibitory neurons (bottom row). Mean voltage is shown by bin with red values indicating excitation and blue indicating inhibition. Microstimulation induced activation is sparse at 0 ms, and secondary firing induces somewhat more widely spread activity, cells then leak back to rest. **b** Response to $200 \mu\text{A}$ stimulation of excitatory neurons (top row) and inhibitory neurons (bottom row). Microstimulation induces strong synchronous activation at 0 ms, this activity triggers secondary firing which forms a structured wave front that propagates outward while decaying in amplitude. Concurrently, excitatory activity drives inhibitory neuron activity which forces activity to decay near the original stimulation site. Between 10ms and 20ms the amplitude of the wave front declines to such a degree that it can no longer drive outward expansion and the combination of refractoriness and inhibitory neuron activity forces overall activity to decay back to rest.

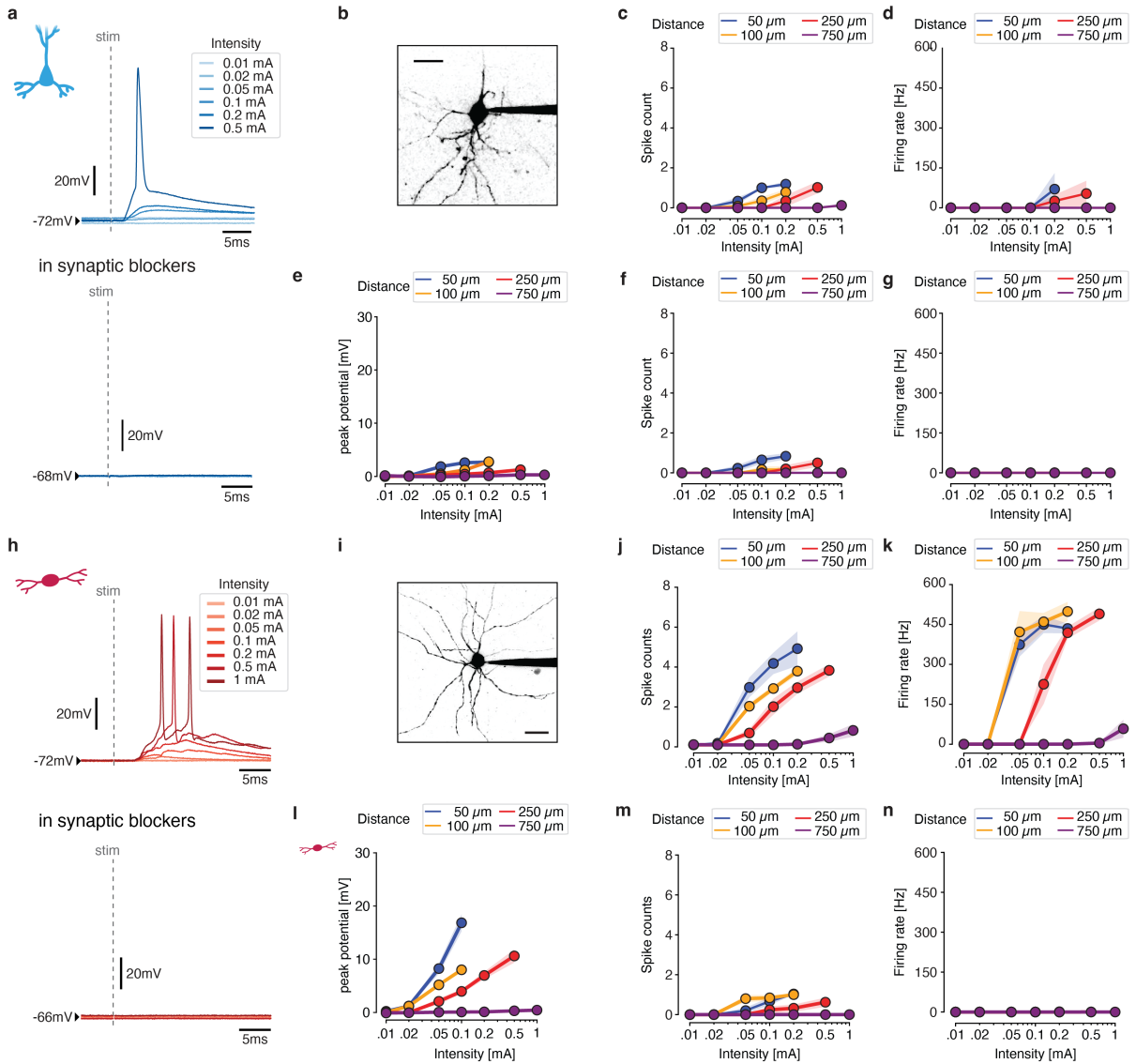


Figure 3.11: Characterization of evoked responses for pyramidal and PV cells. **a** Example evoked responses in control condition (top) and in synaptic blockers (bottom). The stimulating electrode is placed 250 μm away from the cell. **b** Example pyramidal cell stimulated in a. **c** Spike count of evoked responses from pyramidal cells for different distances. Linear Mixed Model with distance ($p=3.98e-5$) and stimulation intensity ($p<2e-16$) as fixed effect. **d** Firing rate of evoked responses from pyramidal cells at different distances. Linear Mixed Model with distance ($p=0.240$) and intensity ($p=1.51e-7$) as fixed effect. **e** Peak response of evoked responses in synaptic blockers from pyramidal cells at different distances. Linear Mixed Model with distance ($p=0.002$) and intensity ($p=0.008$) as fixed effect. **f** Spike count of evoked responses in synaptic blockers from pyramidal cells for different distances. Linear Mixed Model with distance ($p=0.007$) and stimulation intensity ($p<2e-16$) as fixed effect. **g** Firing rate of evoked responses from pyramidal cells at different distances. **h** Example evoked responses in control condition (top) and in synaptic blockers (bottom). The stimulating electrode is placed 750 μm away from the cell. **i** Example PV+ cell stimulated in h. **j** Spike count of evoked responses from PV cells for different distances. Linear Mixed Model with distance ($p=5.36e-8$) and stimulation intensity ($p<2e-16$) as fixed effect. **k** Firing rate of evoked responses from PV cells at different distances. Linear Mixed Model with distance ($p=1.47e-12$) and intensity ($p<2e-16$) as fixed effect. **l** Peak response of evoked responses in synaptic blockers from PV cells at different distances. Linear Mixed Model with distance ($p=0.0003$) and intensity ($p=0.0002$) as fixed effect. **m** Spike count of evoked responses in synaptic blockers from PV cells for different distances. Linear Mixed Model with distance ($p=9.03e-6$) and stimulation intensity ($p<2e-16$) as fixed effect. **n** Firing rate of evoked responses in synaptic blockers from PV cells at different distances. Linear Mixed Model with distance ($p=0.203$) and intensity ($p=0.356$) as fixed effect. Error bars indicate s.e.m.

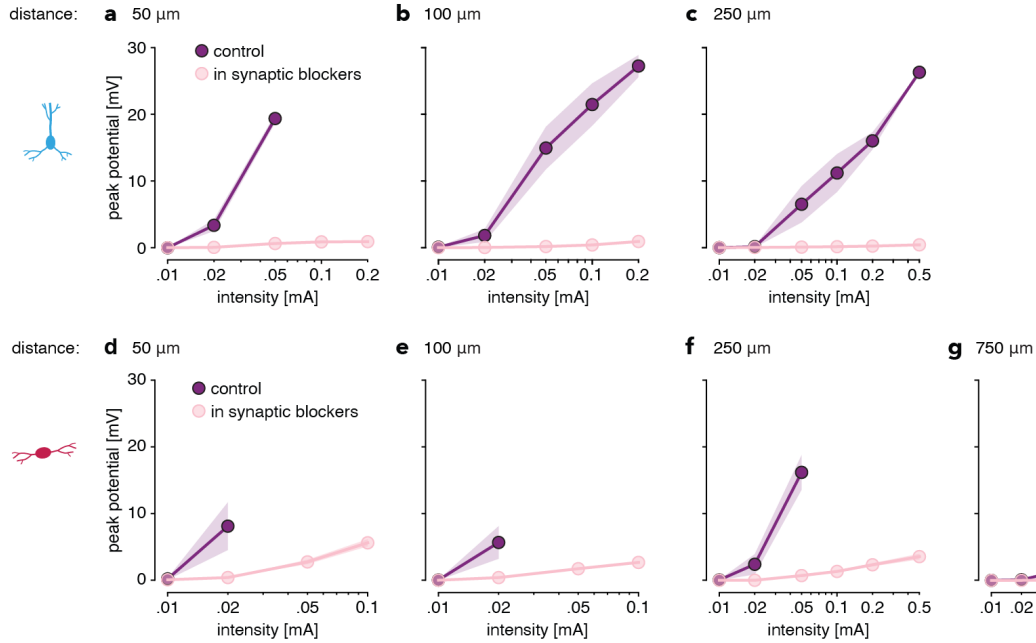


Figure 3.12: Comparison of the effect of synaptic blockers on peak potential for different distances. **a** Peak responses in control condition (violet) and in synaptic blockers (pink) from pyramidal cells. The stimulating electrode is placed 50 μm away from the cell. Linear Mixed Model with intensity ($p=0.427$), the experimental condition (control vs. in synaptic blockers, $p=1.35\text{e-}15$) and their interaction ($p=5.24\text{e-}14$). **b** Peak responses in control condition (violet) and in synaptic blockers (pink) from pyramidal cells. The stimulating electrode is placed 100 μm away from the cell. Linear Mixed Model with intensity ($p=0.642$), the experimental condition (control vs. in synaptic blockers, $p=2.20\text{e-}15$) and their interaction ($p=5.96\text{e-}12$). **c** Peak responses in control condition (violet) and in synaptic blockers (pink) from pyramidal cells. The stimulating electrode is placed 250 μm away from the cell. Linear Mixed Model with intensity ($p=0.765$), the experimental condition (control vs. in synaptic blockers, $p=1.06\text{e-}14$) and their interaction ($p=5.23\text{e-}11$). **d** Peak responses in control condition (violet) and in synaptic blockers (pink) from PV cells. The stimulating electrode is placed 50 μm away from the cell. Linear Mixed Model with intensity ($p=0.0902$), the experimental condition (control vs. in synaptic blockers, $p=0.0103$) and their interaction ($p=0.0191$). **e** Peak responses in control condition (violet) and in synaptic blockers (pink) from PV cells. The stimulating electrode is placed 100 μm away from the cell. Linear Mixed Model with intensity ($p=0.368$), the experimental condition (control vs. in synaptic blockers, $p=0.016$) and their interaction ($p=0.025$). **f** Peak responses in control condition (violet) and in synaptic blockers (pink) from PV cells. The stimulating electrode is placed 250 μm away from the cell. Linear Mixed Model with intensity ($p=0.0479$), the experimental condition (control vs. in synaptic blockers, $p=1.88\text{e-}9$) and their interaction ($p=3.05\text{e-}8$). **g** Peak responses in control condition (violet) and in synaptic blockers (pink) from PV cells. The stimulating electrode is placed 750 μm away from the cell. Linear Mixed Model with intensity ($p=0.889$), the experimental condition (control vs. in synaptic blockers, $p<2\text{e-}16$) and their interaction ($p=1.17\text{e-}10$).

STRUCTURED CONNECTIVITY MAY ENCODE LOCAL PERTURBATIONS IN CORTICAL TRAVELLING WAVES

4.1 INTRODUCTION

In the previous chapter we established that the network response to cortical microstimulation could be seen as a transient travelling wave. This is consistent with previous work that has described the response to stimulation in terms of travelling wave pulses in simplified field models and one dimensional spiking network models [139, 161]. Within neuroscience, travelling waves have also been observed in the context of naturalistic sensory evoked re-

sponses. Experimentally, travelling waves (broadly) have been associated with putative functions in sensory processing where they are thought to be evoked by sensory input and share information between spatially distant neurons [15–18, 163], in the gating of perceptual sensitivity [21], and encoding of reward history [20]. Computationally, travelling wave based propagation of information has improved the performance of recurrent artificial neural networks by allowing for the integration of information from distinct, topologically organized receptive fields [14], acting as a record of the recent past [13, 86], and supporting prediction of future sequences of sensory input [12].

The broad potential relevance of travelling waves to sensory processing, perceptual gating, and representations of past and future events evokes a natural question about microstimulation evoked transient travelling wave pulses: are these responses evoked because electrical stimulation taps into a natural information carrying pathway of the brain? Waves are ubiquitous in physical, biological, and engineered systems [189] but many, such as large, periodic, whole brain oscillations, do not fit the profile of the fast, possibly transient, travelling wave pulses we observed as a consequence of electrical microstimulation [190]. Not all waves are the same, and while intracortical microstimulation can convey sensory information [19, 62], determining whether microstimulation evoked waves carry the same behaviorally relevant sensory information as sensory evoked waves is an empirical question that is beyond the scope of this thesis. However, by extending the spatially structured spiking neural network model from chapter 3.4.4, it is possible to ask whether simulated synaptic input alone can evoke transient travelling waves.

4.2 EXTENDED STRUCTURED CORTICAL SPIKING NEURAL NETWORK MODEL

In 3.4.4 we defined neural connectivity based on Gaussian connection probability functions of inter-somatic distance. In order to isolate propagation mechanisms, these distributions assumed that the relationship between distance and connection probability was the same for all cell type pairs and $\sigma_{syn} = 115 \mu m$ was used for all type pairs, loosely inspired by rodent distributions in [8]. For this chapter, instead of assuming that all cell types connect based on the same function of distance, we redefined connectivity to incorporate the specific cell-type specific connection probability distributions from [8] (Fig. 4.1b). This is equivalent to redefining equations 3.32, 3.33, 3.34, and 3.32 as

$$\mathbf{C}_{ij}^{\mathbf{EE}} \stackrel{\text{i.i.d}}{\sim} \mathcal{B}(\min(1.0, A_{EEg}((\mathbf{x}_{\mathbf{E}})_i - (\mathbf{x}_{\mathbf{E}})_j \mid 0, (\sigma_{ee}^{syn})^2))) \quad (4.1)$$

$$\mathbf{C}_{ij}^{\mathbf{EI}} \stackrel{\text{i.i.d}}{\sim} \mathcal{B}(\min(1.0, A_{EIg}((\mathbf{x}_{\mathbf{E}})_i - (\mathbf{x}_{\mathbf{I}})_j \mid 0, (\sigma_{ei}^{syn})^2))) \quad (4.2)$$

$$\mathbf{C}_{ij}^{\mathbf{IE}} \stackrel{\text{i.i.d}}{\sim} \mathcal{B}(\min(1.0, A_{IEg}((\mathbf{x}_{\mathbf{I}})_i - (\mathbf{x}_{\mathbf{E}})_j \mid 0, (\sigma_{ie}^{syn})^2))) \quad (4.3)$$

$$\mathbf{C}_{ij}^{\mathbf{II}} \stackrel{\text{i.i.d}}{\sim} \mathcal{B}(\min(1.0, A_{IIg}((\mathbf{x}_{\mathbf{I}})_i - (\mathbf{x}_{\mathbf{I}})_j \mid 0, (\sigma_{ii}^{syn})^2))) \quad (4.4)$$

Where $\sigma_{ee}^{syn} = 130 \mu m$, $\sigma_{ei}^{syn} = 100 \mu m$, $\sigma_{ie}^{syn} = 100 \mu m$, and $\sigma_{ii}^{syn} = 130 \mu m$ respectively for mice. Because human data were present but necessarily more limited in the dataset we also inferred human distributions by adding $30 \mu m$ (approximating the mean difference

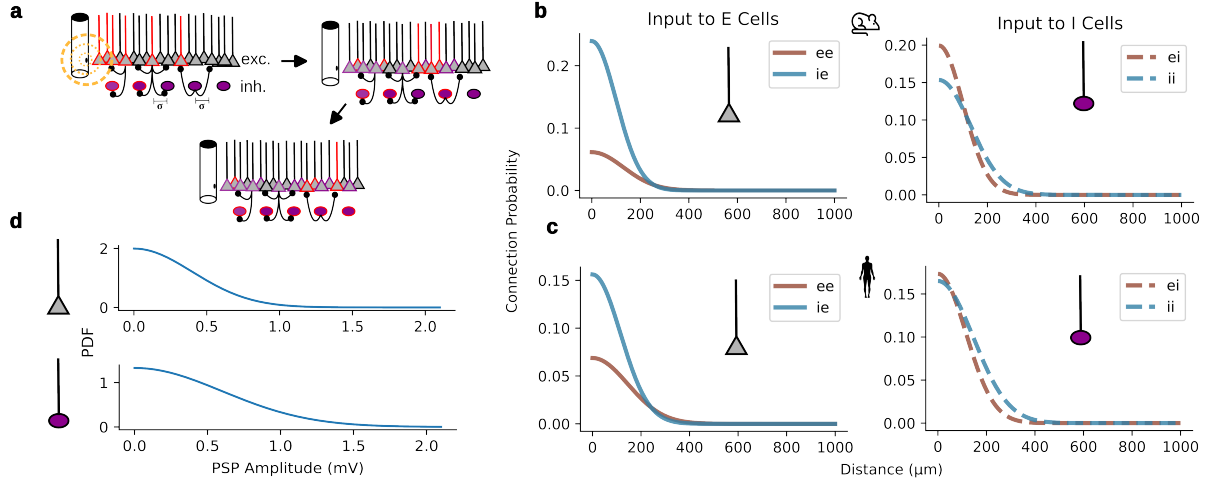


Figure 4.1: In a newly developed software package we leverage sparse tensor computing in Pytorch [188] to simulate leaky integrate and fire networks in three-dimensional rodent and human cortical tissue, with empirically constrained connectivity, in response to direct electrical stimulation. This allows for simulation of tissue volumes up to 4.8 mm^3 ($\leq 800,000$ neurons) on a single L40S GPU. **a:** Illustration of the propagation of activity in a structure neural network follow spatially localized perturbation. **b-c :** Gaussian distributions for the connection probability of different cell class pairs as a function of distance for rodents (**b**) and humans (**c**). In the human case we approximate the pair (e-e, e-i, i-e, ii) specific values based on the human class level fits and the class pair structure of the rodent data because class pair specific fits were not reported for humans. **d:** Class specific efficacy values (i.e. connection weights) for the connections are sampled from class specific half-normal distributions based on values from [8].

in pooled values between species) to the σ value of rodent cell type specific distributions.

This yielded human model values of $\sigma_{ee}^{syn} = 145 \text{ } \mu\text{m}$, $\sigma_{ei}^{syn} = 115 \text{ } \mu\text{m}$, $\sigma_{ie}^{syn} = 115 \text{ } \mu\text{m}$, $\sigma_{ii}^{syn} = 145 \text{ } \mu\text{m}$, and class specific scaling parameters $A_{EE} = 25$, $A_{IE} = 45$, $A_{EI} = 50$ and $A_{II} = 60$ (Fig. 4.1c).

For both species we also generalized our spiking network simulation to a three dimensional sheet of cortical tissue measuring $4 \text{ mm} \times 4 \text{ mm} \times 0.3 \text{ mm}$ that contained 320,000 excitatory cells and 80,000 inhibitory cells. We computed the Euclidean distance between all cell-pairs for construction of the connectivity matrices. Interestingly, in the cell-class spe-

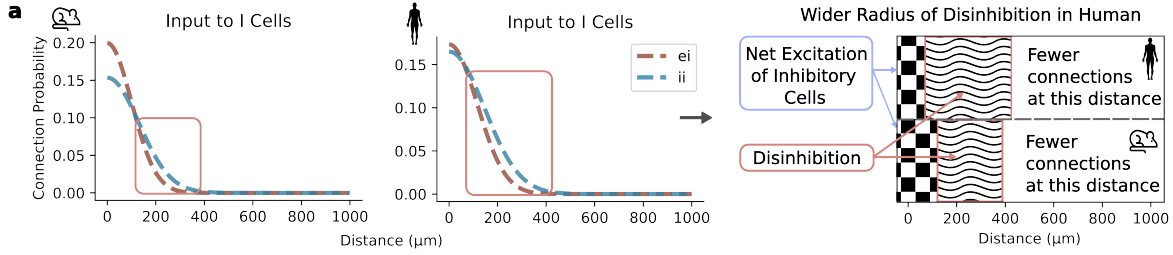


Figure 4.2: *Evoked waves encode stimulus intensity and propagate further in human. A: Pre-synaptic excitatory and inhibitory connections disinhibit activity far from the pre-synaptic cells by applying net inhibition to distant inhibitory cells. The region of expected disinhibition is indicated in rectangles overlayed on connection probabilities from Fig. 4.1B and 4.1C (right).*

cific distributions the inhibitory to inhibitory connection distribution was spatially wider than the connections from excitatory cells to inhibitory cells. As a result, it was possible for spatially localized firing activity to drive an increase in inhibition locally while disinhibiting cells further away (Fig. 4.2). This distal disinhibition was wider in humans suggesting that the inhibitory damping of activity which drove monotonic decay in the amplitude of excitatory travelling wave pulses in Fig. 3.4 should exert a weaker effect in human cortical networks. We set the absolute refractory period of excitatory neurons in the network to 20 ms and that for inhibitory neurons to 5 ms.

4.3 RESULTS

Having prepared three dimensional spiking neural networks with cell-class specific distance dependent connection probability we then introduced background synaptic input to the network as a substitute for the context of the wider brain. At each time step, we assigned each cell 100 simulated pre-synaptic background “neurons” represented by Poisson

processes with 5 Hz firing frequencies that were randomly assigned an excitatory (95 % chance) or inhibitory (5% chance) class identity and allocated synaptic efficacies from the same class specific efficacy distributions that were used to build the networks (Fig. 4.1 and section 3.4.4). For all simulations we allowed the network activity to equilibrate in the presence of this background input for 200 ms before introducing stimulation.

With the network configured, we simulated the effect of a single pulse of 200 μA electrical microstimulation at the center of the sheet (location 2 mm x 2 mm x 0.15 mm) in mouse and human networks built. In order to isolate the effect of the cell-type specific connectivity, we tested both the symmetric 115 μm connectivity from chapter 3 and the cell class specific connectivity distributions described above. Fig. 4.3 shows the progression of responses to this stimulation over time in terms of binned thresholded voltages presented in the two longer 4 mm dimensions. The brightest orange points indicate firing events, blue indicates hyperpolarization, and cells at rest are black. As seen in fig. 4.3a, the rodent network with symmetric connectivity had a response to stimulation that was slightly weaker than, but largely consistent with, the predictions of our one dimensional model without background input from 3.5. The human symmetric connectivity network response shown in fig. 4.3b was somewhat stronger than the rodent response and had a more pronounced post stimulation inhibitory effect. The responses of the networks that had cell-class specific distance dependent connectivity were more dramatic: In the mouse case, the overall propagation distance was larger than in was in the mode with symmetric connectivity and had a more pronounced initial excitatory pulse fig. 4.3c. In the human case, the impact of distal disinhibition was dramatic and allowed the evoked pulse to travel

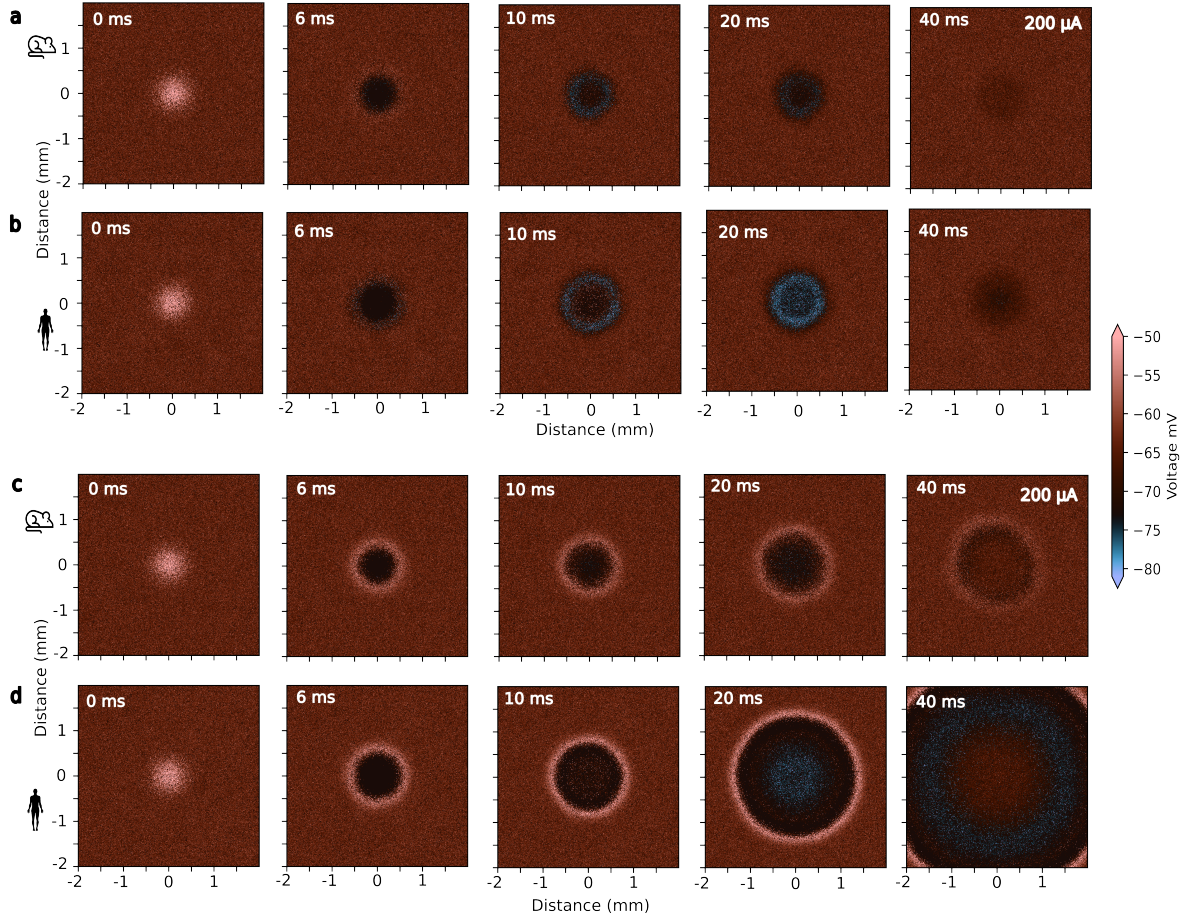


Figure 4.3: *Human and Rodent Microstimulation Responses. a* Response at different time points in the presence of background input following stimulation at 200 μA in a model with equal connectivity σ for all cell-type pairs calibrated to mouse cortical connectivity. *b* The same as **a** but calibrated to human cortical connectivity. *c* Response at different time points in the presence of background input following stimulation at 200 μA in a model with different cell-type pair specific connectivity σ values calibrated to mouse cortical connectivity. *d* The same as **c** but calibrated to human cortical connectivity.

to the boundaries of the simulated sheet. This predicted measurable response activity at distances close to 2 mm fig. 4.3d. Ultimately, our mouse models had similar responses in both connectivity regimes that were both consistent with our empirical response observations in Fig. 3.6. In contrast, cell-class specific connectivity had a large impact in the human case and yielded a testable prediction: If distal disinhibition arising from cell class specific distributions facilitates activity propagation in human cortical networks, then microstimulation evoked responses should propagate at least twice as far in human cortical tissue as in rodent tissue.

Once we had established large structured networks with background input and validated that they were capable of supporting microstimulation evoked travelling waves, we sought to determine whether synaptic input (such as that from sensory events) could evoke traveling waves.¹ In addition to the background input described above, we assigned 100 pre-synaptic neurons per cell to drive a receptive field with a radius of $135\ \mu\text{m}$ that was centered in the slice. We used Poisson processes to generate field specific inputs, and assigned them excitatory (95 % of inputs) and inhibitory (5 % of inputs) identities. We then sampled connection strengths from the same class specific efficacy distributions that were used to build the networks (Fig. 4.1 and section 3.4.4). First, the networks were allowed to equilibrate to the background input for 200 ms. Next, input from driving “neurons” was delivered for a 10 ms period at a firing frequency of either 10 Hz or 60 Hz. Background input was sustained for the entire simulation. Fig. 4.4 shows the cell class specific re-

¹We performed these simulations in the models with the empirically inspired cell-class specific distance dependent connection probabilities.

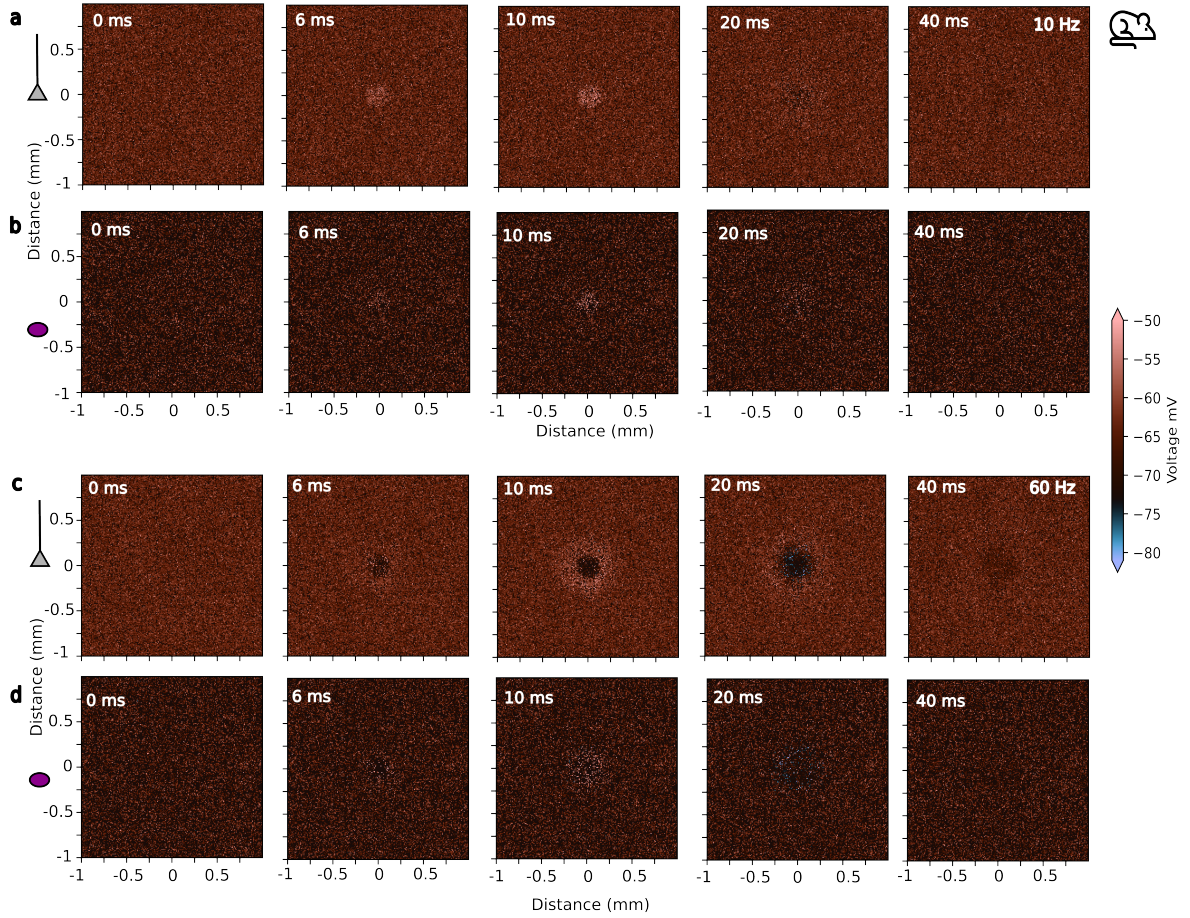


Figure 4.4: *Synaptically evoked waves in mouse SNN. **a-b** Response of network with cell-type pair specific cortical connectivity calibrated to mouse connectivity at different time points in the presence of background input following targeted synaptic input to central receptive field for 10 ms from $t=0$ ms to $t=10$ ms at 10 Hz. **a** the top row shows the excitatory neuron response and **b** the lower row shows the corresponding inhibitory neuron response. **c-d** show the same thing but for a synaptic input at 60 Hz.*

sponse of the rodent network to this localized pulse of synaptic drive. The 10 Hz response in Fig. 4.4a (excitatory cells) and Fig. 4.4b (inhibitory cells) both capture robust evoked firing in the target field, but no travelling waves were evoked. In contrast, when high frequency synaptic input was delivered at 60 Hz (Fig. 4.4c and 4.4d), a robust wave front was evoked which propagating outwards and triggering activity in cells that are spatially adjacent to the driven field. These differing responses suggest that wave formation may be conditionally gated behind a requirement of high frequency synaptic input.

We repeated our spatially localized synaptic drive simulation using the networks configured to simulate human cortical responses. Stimulation at 10 Hz had little effect on firing activity outside of the targeted field, though in the human case there was some faint depolarization of cells adjacent to the target field (Fig. 4.5a and 4.5b). At 60 Hz however, the synaptic stimulation evoked a travelling wave (Fig. 4.5c and 4.5d). This wave propagated to the boundaries of the sheet after which the network returned to its background driven equilibrium.

While the spatial scale of the synaptic responses to the 60 Hz drive was larger than that evoked by the 10 Hz drive, it was also evoked by a higher total level of synaptic input. In order to better understand the conditional gating of wave formation by temporal properties of synaptic drive, we performed another set of simulations. For these simulations, we balanced the total input rate over a 120 ms window. In one case we delivered six 10 ms pulses of synaptic drive separated by 10 ms intervals. In the other case, we delivered a single 10 ms 60 Hz pulse. Both of these protocols were delivered after 200 ms of background input. Fig. 4.6 shows the number of evoked spikes at different distances from the center

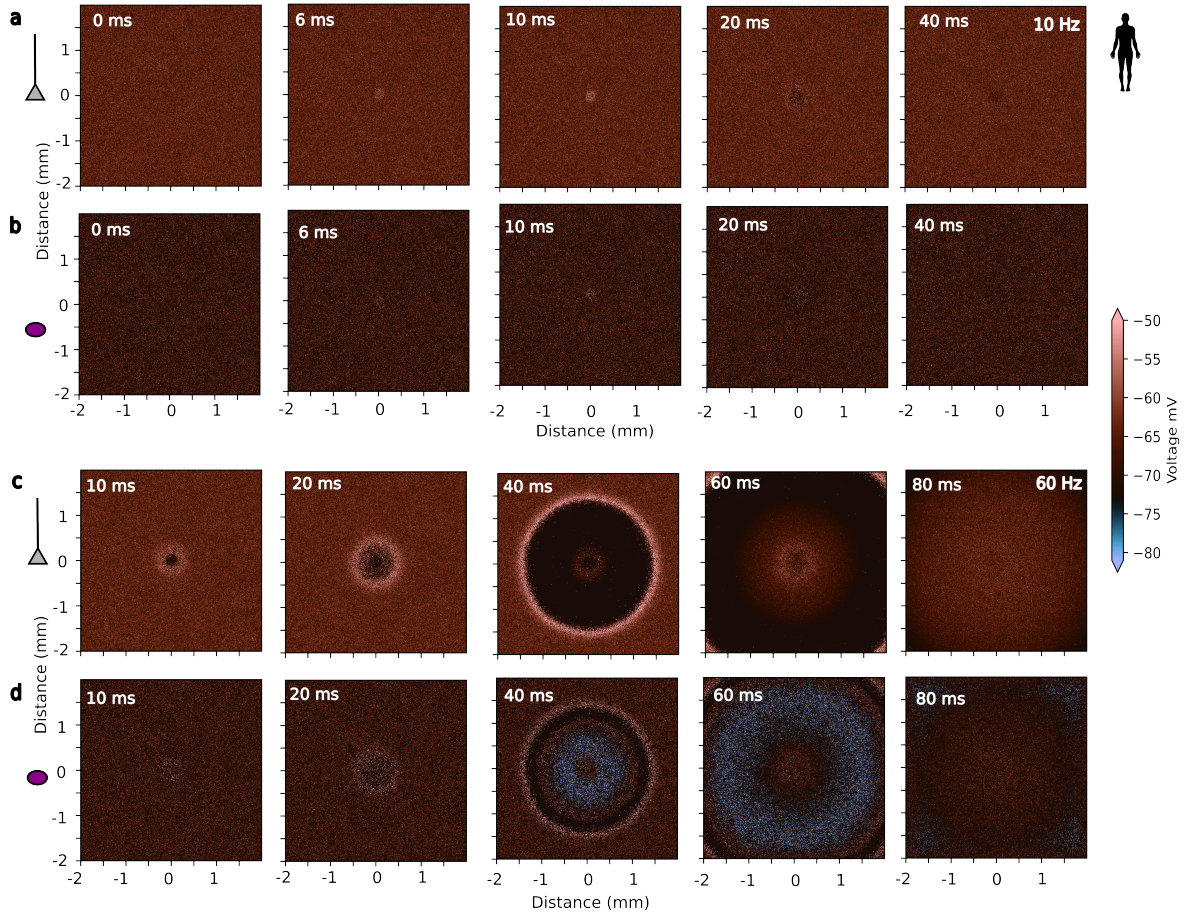


Figure 4.5: Synaptically evoked waves in human SNN. **a-b** Response of network with cell-type pair specific cortical connectivity calibrated to human connectivity at different time points in the presence of background input following targeted synaptic input to central receptive field for 10 ms from $t=0$ ms to $t=10$ ms at 10 Hz. **a** the top row shows the excitatory neuron response and **b** the lower row shows the corresponding inhibitory neuron response. **c-d** show the same thing but for a synaptic input at 60 Hz.

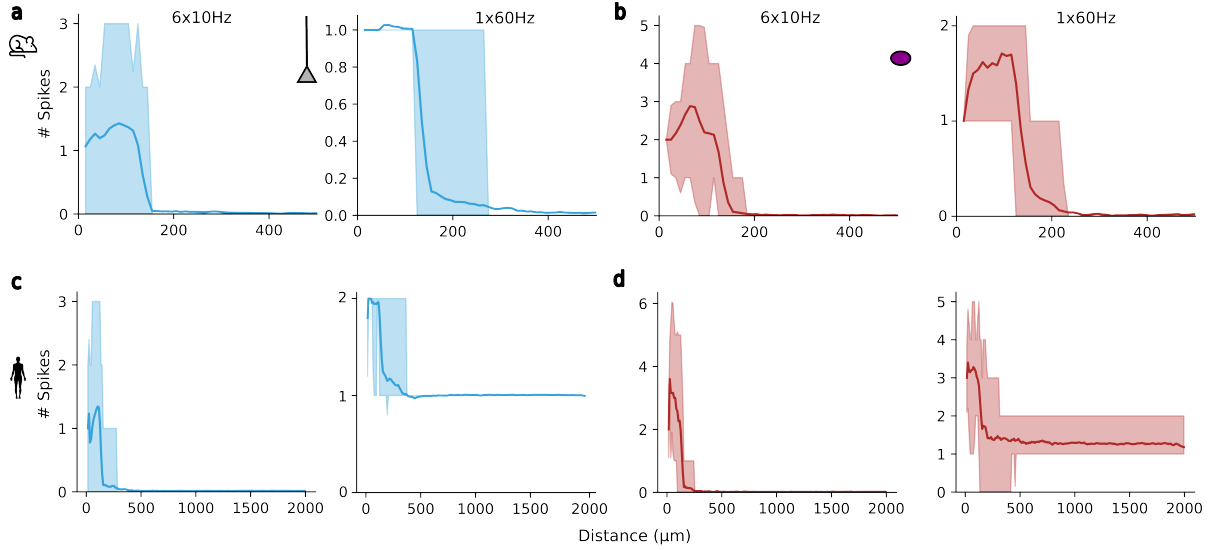


Figure 4.6: *Spatial response distribution of responses to different temporal patterns. **a** Excitatory neuron spike counts in a rodent SNN model in response to six 10 ms periods of 10 Hz synaptic drive with 10 ms inter drive intervals (left) or one 10 ms period of 60 Hz synaptic drive. The number of spikes is reported over a 120 ms window. The solid line is represents mean count and shading indicates the central 90 % of the response distribution. **b** is the same as **a** but for inhibitory neurons. **c** is the same as **a** but for an SNN model calibrated to human connectivity. **d** is the same as **b** but for an SNN model calibrated to human connectivity.*

of the receptive field. In mice, the train of stimulation at 10 Hz drove a higher total number of spikes in cells within the target field than the single 60 Hz pulse (Fig. 4.6a) and evoked very limited firing outside of the field. In contrast, the 60 Hz pulse almost never triggered multiple spikes in the target field (4.6b) but did trigger activity in cells over almost double the radius. In the human model the 10 Hz drive (4.6c) largely drove spiking within the target field, though a degree of mild surrounding activity was observable. At 60 Hz, almost every cell in the network fired at least a single spike and the mean number of spikes in the driven field was also higher than in the 10 Hz case, but the upper end of the firing count distribution was lower than that of the train of 10 Hz inputs.

4.4 DISCUSSION

What does the greater propagation distance of travelling waves in simulated human cortical networks mean? Events in the world are structured both spatially and temporally: while brains vary greatly in size across species, they must coordinate neuronal populations in response to shared environmental timescales. If sensory stimulation evokes cortical travelling waves that share information between spatially distant neurons, then these waves must travel farther in human cortex to coordinate fields corresponding to an equivalent fraction of sensory space in mice. Viewed through the lens of spatiotemporal integration, this is consistent with travelling waves in the larger human cortex amplifying the same spatially localized signals to a greater degree and sharing them with a dramatically greater number of neurons.

The frequency dependent gating of wave formation suggests that evoked waves could exist as a contextual flag to modify the context or salience of input and amplify it beyond the conventional post-synaptic targets of a receptive field. For example, burst or other high frequency codes indicating novelty [191] or error [45] could potentially be amplified by travelling waves to elevate perceptual sensitivity [21]. Waves could also trigger spiking following temporally coincident inputs in adjacent cells to coordinate timing dependent learning rules [87–92]. Specifically, novelty evoked waves could forcibly trigger spiking activity in spatially proximal cells. In cells near the receptive field where the wave originated, the wave would trigger a spike that followed feedforward synaptic input from before the

event. This could potentiate connections associated with earlier inputs that were temporally correlated with the unexpected event. Farther from the receptive field where the wave originated, the wave could trigger a spike up to tens or hundreds of milliseconds after the event. This could potentiate connections related to sequences that followed the unexpected event. When repeated many times, this process may allow for wave based coordination of local spike timing dependent plasticity rules and support self-supervised learning of spatially and temporally structured event sequences. Whether travelling waves perform these roles remains an open question that requires empirical validation. Nonetheless, travelling waves may play significant roles in sensory processing, memory, or the coordination of learning and should be given serious consideration as part of the neural code.

PART III

CONCLUSION

CONCLUSION

A comprehensive, quantitative, and empirically anchored understanding of the optimal neural code within cortical circuits is crucial for elucidating the fundamental language of the brain. This objective holds widespread significance to the refinement of existing theory and the advancement of future biomedical interventions, guided by an enhanced understanding of neural systems [192–194]. Additionally, this improved understanding may lay the groundwork for improving industrially pertinent technologies, such as neuromorphic (brain-like) computing [35, 46, 195, 196].

Our results in chapter 2 suggest the existence of 5 to 6 subtypes of STP in glutamatergic cortical synapses that overlap species, are not explainable by continuum hypotheses, and appear to have a genetic basis. The existence of 5 to 6 subtypes changes our understanding of STP in several ways: a plurality of facilitating types hints at the possibility of multiple high frequency coding schemes. Bimodal STP (with fast depression and longer lived facili-

tation) may support decoding of multi-level (≥ 3) multiplexing schemes [43] where cortical circuits transmit three or more different messages from the same axon using a combination of long bursts, short bursts, and isolated events. The diversity of STD we found may be attributed to differences in release probability or asynchronous release [109] which in turn hint at different numbers of pre-synaptic contacts which may serve to conserve energy at the synapse [110].

Understanding these dynamics will allow the field to revise theories of message separation between post-synaptic targets [26–29], detection of different firing rates [30–32], separation of different types of bursts and events [31], signal correlation [37] and decorrelation [38] and preservation of time delay information [33–36]. This understanding, will be relevant to theoretical accounts of working memory [39–41] and credit assignment for biological learning [34, 45, 46]. Ultimately, by presenting a basic dictionary of the functional motifs of information transmission in the brain, and elucidating their genetic correlates, we offer a basic taxonomy that may facilitate advancement in the understanding of several crucial functions of the brain.

How are synaptic dynamics functionally organized across GABAergic cells? In work outside of this thesis, that is now available as a preprint [3], we have applied the framework from chapter 2 [1] to GABAergic cortical synapses. Combined with the functional subtypes presented in chapter 2, this supports a complete understanding of canonical cortical circuit motifs. This understanding is a foundation for understanding how STP in cortical connections influences information processing. Our findings suggest that unlike in glutamatergic synapses, the dynamics of GABAergic connections are consistent with a

continuum of STP properties. This continuum nonetheless encodes genetic identities in the same way that the excitatory types do, suggesting that the GABAergic continuum is an organizational feature rather than a failure of machine learning. This implies that signal processing which involves conditional gating of specific patterns types may occur primarily in the excitatory connections of cortical circuits. Of these excitatory connections, the most diverse excitatory projections are likely those onto GABAergic cells. One possible account of these results is that larger excitatory cells have more physical resources for the complex processing required to encode multiple signals into their output. For example, complex dendritic processing may encode information received at dendrites into bursts, supporting the type of multiplexing quantified in [43] and leveraged in [45] and [46]). Under this paradigm, the distinct subtypes of synaptic dynamics in excitatory cortical connections may serve to separate the distinct signals encoded into their spiking output for different targets.

If they belong to a functional continuum, then GABAergic cells could decode inputs and perform information processing roles in cortical circuits by projecting decoded signals to specific cell targets and sub-cellular locations. For example, an interneuron could preferentially receive burst-coded information through facilitation and then perform different circuit level operations based on its projections. The interneuron could relay this signal forward to other inhibitory neurons which could ultimately result in suppression of inhibition. Alternately, if the interneuron relays the signal to a specific dendritic location on an excitatory cell, that may allow the excitatory cell to receive the specific signal component decoded by the interneuron. These dynamics might also act as a reservoir to encode tempo-

rally relevant information [35], for example in STP at mossy fiber-granule cell connections in cerebulum, five-types of STP yielded high precision sub-millisecond representations of time [34, 197]. If this logic is extended by adding an arbitrary number of STP types, each tuned to be responsive to patterns on a slightly different timescale then at the limit of many subtypes the population of connections is likely to become indistinguishable from a functional continuum.

MICROSTIMULATION AND CORTICAL WAVES

What does our model of microstimulation mean for the treatment of diseases involving dys-regulated cortical dynamics? Many applications require both inference (to understand the state of the brain) and stimulation (to change that state). For example, treating Epilepsy includes both detection of seizure onset or continuation and stimulation to suppress it [57]. As such, intracortical brain computer interfaces (iBCIs) must be able to infer important information from the brain while compensating for the effect of activity arising from stimulation. Moreover, in order to be usable, signal processing and control must be very rapid, and must have low computational and energy demands compatible with portable human applications. However, activity in the brain propagates through neural circuits that have complicated recurrent structures and nonlinear dynamics [8, 64] which can result in similarly complicated response patterns. As such, provided that the stimulus magnitude is sufficiently great, our ability to efficiently separate these overlapped dynamics is limited [66, 67]. Because the model we present in chapter 3 is mechanistic in nature and compu-

tationally efficient, it may support inference of the properties of stimulus evoked responses in cases where such responses could otherwise not be separated from internally generated dynamics.

The waves evoked in our model of microstimulation propagate according to the properties of waves in excitable media. If disease states in the brain include dysregulated activity that also obeys these properties, then it may be possible to cancel abnormal travelling waves directly through collisions between disease related pulses and electrically evoked pulses (because neither pulse can travel through the trailing shock left by the other). This could make it possible to avoid engineering complicated or global stimulation patterns for destructive interference of more classical waveforms. Alternately, it is possible that diseases such as Epilepsy are associated with changes in the properties that shape the dynamics of cortical wave formation and propagation [146]. If this is the case, then a mechanistic understanding of the synaptic and refractory conditions that give rise to healthy cortical travelling waves may support the identification of biophysical conditions under which they can become dysregulated. Such conditions could become new therapeutic targets facilitating the potential development of new clinical interventions.

When paired with precise mechanistic models, cortical microstimulation shows promise for expanded use in conditions such as Parkinson's disease. In Parkinson's, excess synchrony in beta oscillations is thought to be expressed in multiple areas including motor cortex and is thought to partially underlie disease symptoms [198, 199]. Microstimulation in motor cortex has been explored as a possible therapeutic intervention, but is partly limited by the difficulty of predicting stimulus responses in complex cortical circuits [200–

202], as a result, deep brain stimulation is the dominant therapeutic approach [203, 204]. Nonetheless, the potential use of therapeutic cortical microstimulation in Parkinson’s disease may be highly synergistic with emerging brain-spine interface based neuroprosthetic treatments [54] for the disease because these neuroprosthetics may already include the implantation of intracortical electrical interfaces. Therefore, it may be possible to deliver therapeutic stimulation without the requirement of an independent surgery if intracortical electrical interfaces are designed from the outset for both recording and microstimulation.

The potential of microstimulation as a tool to evoke cortical travelling waves has implications both for the restoration of sensory function in clinical populations [62] and for the fundamental study of information processing in the brain [15–19, 127, 128]. Whether the waves predicted by our models of electrical microstimulation are the same as those evoked by sensory stimulation remains an open question. However, past work has hinted that microstimulation can convey sensory information [19] and does so most effectively when the evoked activity mimics naturalistic activity [62]. If microstimulation can evoke naturalistic dynamics, then careful engineering of evoked waves could support improved sensory neuroprosthetics. Further, studying the information carrying properties of evoked waves may permit the generation of specific predictions that can be tested empirically through microstimulation.

How could synaptic dynamics decode the information contained in cortical travelling waves? In fig. 4.6 synaptic input to receptive fields generated repeated spiking activity when low to moderate frequency input was regularly delivered. In contrast, brief high frequency input triggered evoked waves that produced either a single spike or no spikes

in the vast majority of excitatory neurons within a radius much wider than the receptive field itself. Short-term plasticity operates on spike trains which means that it does not meaningfully shape the decoding of single spikes. However, in a topographically organized network, spikes evoked by travelling waves are not the only spikes: receptive fields that receive feedforward input will have their own firing activity. If a travelling wave propagates over a receptive field that is also driven by feedforward input within a coincident timing window, then cells in the receptive field will fire an elevated number of spikes.

Spike trains evoked by the coincidence of feedforward input and passing cortical waves could likely be detected by rapid time course short-term facilitation such as that expressed by the ‘R3’ rodent cluster and ‘H5’ human cluster from chapter 2 (fig. 2.4). Conversely, biphasic synapses where rapid depression is followed by long time course facilitation, such as the ‘R0’ rodent cluster and ‘H2’ human cluster from chapter 2 (fig. 2.4) could filter out short coincident events in favour of detecting firing that is driven by sustained feedforward input to a receptive field. Combining synapses with both of these types of STP could potentially allow higher cortical areas to de-multiplex input from topographically structured networks into regular feedforward receptive field activity and coincident events in weakly activated receptive fields. In this scenario, the coincident passage of travelling waves could elevate the perceptual salience of weak input to receptive fields. This provides a possible explanation for wave based modulation of perceptual sensitivity [21]. Alternately, multiplexed wave codes could potentially permit the transfer of coincident error or reward [20] signals to higher cortical areas.

The language of the brain is articulated by the rich dynamics of both individual synapses

and of sophisticated networks shaped by synaptic connectivity. In order to understand this language, we present a basic functional taxonomy of cortical synaptic dynamics in mice and humans, a detailed mechanistic account of the cortical response to microstimulation, and a speculative exploration of cortical travelling waves at the level of individual spike patterns. At the level of individual synapses, we found five functional groupings of synaptic dynamics in mice and six such groupings in humans which partially converged with transgenically defined cell types. These provide a lens into the basic motifs of information transmission in the brain. At the level of networks, we have identified the key biophysical determinants of the cortical response to microstimulation. This offers a predictive, mechanistic framework for precision engineering of patterned stimulation as a tool to treat disease and to understand the brain. Finally, we have generated a falsifiable prediction about how the structure of human cortical connectivity shapes the spatial extent of its response to microstimulation. Many questions remain about the link between dynamics that occur in individual synapses at the nanometer scale, dynamics that occur in structured networks at the millimeter scale, and the biological miracle of the brain's ability to process information. This thesis provides a cross-species account of how the structure of cortical connectivity shapes dynamics across networks and a set of tools by which those networks can be studied in the hope that it brings us a bit closer to the answers of these fundamental questions.

REFERENCES

1. Beninger, J., Rossbroich, J., Tóth, K. & Naud, R. Functional subtypes of synaptic dynamics in mouse and human. *Cell Reports* **43** (2024).
2. Rossbroich, J., Trotter, D., Beninger, J., Tóth, K. & Naud, R. Linear-nonlinear cascades capture synaptic dynamics. *PLoS computational biology* **17**, e1008013 (2021).
3. Poirier, J., Beninger, J., Tóth, K. & Naud, R. Functional Continuum of GABAergic Synaptic Dynamics Reflects Genetic Identities. *bioRxiv*, 2026–06 (2026).
4. Poirier, J., Beninger, J. & Naud, R. Computational protocol for modeling and analyzing synaptic dynamics using SRPlasticity. *STAR protocols* **6**, 103652 (2025).
5. Aristotle, A. *De anima (On the soul)* (Penguin London, 1986).
6. Heil, J. *Philosophy of mind: A contemporary introduction* (Routledge, 2019).
7. Hodgkin, A. L. & Huxley, A. F. Action potentials recorded from inside a nerve fibre. *Nature* **144**, 710–711 (1939).
8. Campagnola, L. *et al.* Local connectivity and synaptic dynamics in mouse and human neocortex. *Science* **375**, eabj5861 (2022).
9. Feng, T. The changes in the end-plate potential during and after prolonged stimulation. *Chin. J. Physiol* **13**, 79–107 (1941).
10. Eccles, J. C., Katz, B. & Kuffler, S. W. Nature of the” endplate potential” in curarized muscle. *Journal of Neurophysiology* **4**, 362–387 (1941).
11. Magleby, K. & Zengel, J. E. A dual effect of repetitive stimulation on post-tetanic potentiation of transmitter release at the frog neuromuscular junction. *The Journal of Physiology* **245**, 163–182 (1975).
12. Benigno, G. B., Budzinski, R. C., Davis, Z. W., Reynolds, J. H. & Muller, L. Waves traveling over a map of visual space can ignite short-term predictions of sensory input. *Nature Communications* **14**, 3409 (2023).
13. Keller, T. A., Muller, L., Sejnowski, T. & Welling, M. *Traveling waves encode the recent past and enhance sequence learning* in *International Conference on Learning Representations* **2024** (2024), 12763–12789.
14. Jacobs, M., Budzinski, R. C., Muller, L., Ba, D. E. & Keller, T. A. *Traveling waves integrate spatial information into spectral representations* in *Second Workshop on Representational Alignment at ICLR 2025* ().

15. Gonzales, D. L. *et al.* Touch-evoked traveling waves establish a translaminar spacetime code. *Science Advances* **11**, eadr4038 (2025).
16. Binguier, V., Chavane, F., Glaeser, L. & Frégnac, Y. Horizontal propagation of visual activity in the synaptic integration field of area 17 neurons. *Science* **283**, 695–699 (1999).
17. Aggarwal, A. *et al.* Visual evoked feedforward–feedback traveling waves organize neural activity across the cortical hierarchy in mice. *Nature communications* **13**, 4754 (2022).
18. Muller, L., Reynaud, A., Chavane, F. & Destexhe, A. The stimulus-evoked population response in visual cortex of awake monkey is a propagating wave. *Nature communications* **5**, 3675 (2014).
19. Hickman, J. L., Hughes, G., Sahai, E., Miles, M. & Denman, D. J. Neural population dynamics of direct electrical stimulation of neocortex. *Cell Reports* **45** (2026).
20. Zabehe, E., Foley, N. C., Jacobs, J. & Gottlieb, J. P. Beta traveling waves in monkey frontal and parietal areas encode recent reward history. *Nature Communications* **14**, 5428 (2023).
21. Davis, Z. W., Muller, L., Martinez-Trujillo, J., Sejnowski, T. & Reynolds, J. H. Spontaneous travelling cortical waves gate perception in behaving primates. *Nature* **587**, 432–436 (2020).
22. Stevens, C. F. & Wang, Y. Facilitation and depression at single central synapses. *Neuron* **14**, 795–802 (1995).
23. Lysetskiy, M., Földy, C. & Soltesz, I. Long- and short-term plasticity at mossy fiber synapses on mossy cells in the rat dentate gyrus. *Hippocampus* **15**, 691–696 (2005).
24. De Pasquale, R. & Sherman, S. M. Synaptic properties of corticocortical connections between the primary and secondary visual cortical areas in the mouse. *Journal of Neuroscience* **31**, 16494–16506 (2011).
25. Sherman, S. M. Thalamocortical interactions. *Current opinion in neurobiology* **22**, 575–579 (2012).
26. Crick, F. Function of the thalamic reticular complex: the searchlight hypothesis. *Proceedings of the National Academy of Sciences* **81**, 4586–4590 (1984).
27. Scanziani, M., Gähwiler, B. H. & Charpak, S. Target cell-specific modulation of transmitter release at terminals from a single axon. *Proceedings of the National Academy of Sciences* **95**, 12004–12009 (1998).
28. Markram, H., Wang, Y. & Tsodyks, M. Differential signaling via the same axon of neocortical pyramidal neurons. *Proceedings of the National Academy of Sciences* **95**, 5323–5328 (1998).
29. Reyes, A. *et al.* Target-cell-specific facilitation and depression in neocortical circuits. *Nature neuroscience* **1**, 279–285 (1998).
30. Abbott, L. F., Varela, J., Sen, K. & Nelson, S. Synaptic depression and cortical gain control. *Science* **275**, 221–224 (1997).
31. Fuhrmann, G., Segev, I., Markram, H. & Tsodyks, M. Coding of temporal information by activity-dependent synapses. *Journal of neurophysiology* (2002).
32. Klyachko, V. A. & Stevens, C. F. Excitatory and feed-forward inhibitory hippocampal synapses work synergistically as an adaptive filter of natural spike trains. *PLoS biology* **4**, e207 (2006).

33. Cook, D. L., Schwindt, P. C., Grande, L. A. & Spain, W. J. Synaptic depression in the localization of sound. *Nature* **421**, 66–70 (2003).
34. Barri, A., Wiechert, M., Jazayeri, M. & DiGregorio, D. Synaptic basis of a sub-second representation of time in a neural circuit model. *Nature Communications* **13**, 7902 (2022).
35. Aitken, K. & Mihalas, S. Neural population dynamics of computing with synaptic modulations. *Elife* **12**, e83035 (2023).
36. Wu, Y. K. & Zenke, F. Nonlinear transient amplification in recurrent neural networks with short-term plasticity. *Elife* **10**, e71263 (2021).
37. Mejías, J. F. & Torres, J. J. The role of synaptic facilitation in spike coincidence detection. *Journal of Computational Neuroscience* **24**, 222–234 (2008).
38. Goldman, M. S., Maldonado, P. & Abbott, L. Redundancy reduction and sustained firing with stochastic depressing synapses. *Journal of Neuroscience* **22**, 584–591 (2002).
39. Barak, O. & Tsodyks, M. Persistent activity in neural networks with dynamic synapses. *PLoS computational biology* **3**, e35 (2007).
40. Mejias, J. F. & Torres, J. J. Maximum memory capacity on neural networks with short-term synaptic depression and facilitation. *Neural computation* **21**, 851–871 (2009).
41. Mongillo, G., Barak, O. & Tsodyks, M. Synaptic theory of working memory. *Science* **319**, 1543–1546 (2008).
42. Izhikevich, E. M., Desai, N. S., Walcott, E. C. & Hoppensteadt, F. C. Bursts as a unit of neural information: selective communication via resonance. *Trends Neurosci.* **26**, 161–167 (2003).
43. Naud, R. & Sprekeler, H. Sparse bursts optimize information transmission in a multiplexed neural code. *Proceedings of the National Academy of Sciences* **115**, E6329–E6338 (2018).
44. Naud, R., Friedenberger, Z. & Toth, K. Silences, Spikes and Bursts: Three-Part Knot of the Neural Code. *arXiv e-prints*, arXiv–2302 (2023).
45. Payeur, A., Guerguiev, J., Zenke, F., Richards, B. A. & Naud, R. Burst-dependent synaptic plasticity can coordinate learning in hierarchical circuits. *Nature neuroscience* **24**, 1010–1019 (2021).
46. Greedy, W., Zhu, H. W., Pemberton, J., Mellor, J. & Ponte Costa, R. Single-phase deep learning in cortico-cortical networks. *Advances in Neural Information Processing Systems* **35**, 24213–24225 (2022).
47. Buonomano, D. V. & Merzenich, M. M. Temporal information transformed into a spatial code by a neural network with realistic properties. *Science*, 1028–1028 (1995).
48. Chabrol, F. P., Arenz, A., Wiechert, M. T., Margrie, T. W. & DiGregorio, D. A. Synaptic diversity enables temporal coding of coincident multisensory inputs in single neurons. *Nature neuroscience* **18**, 718–727 (2015).
49. Chamberland, S., Evstratova, A. & Tóth, K. Interplay between synchronization of multivesicular release and recruitment of additional release sites support short-term facilitation at hippocampal mossy fiber to CA3 pyramidal cells synapses. *Journal of Neuroscience* **34**, 11032–11047 (2014).

50. Savanthrapadian, S. *et al.* Synaptic properties of SOM-and CCK-expressing cells in dentate gyrus interneuron networks. *Journal of Neuroscience* **34**, 8197–8209 (2014).
51. Neubrandt, M. *et al.* Single bursts of individual granule cells functionally rearrange feedforward inhibition. *Journal of Neuroscience* **38**, 1711–1724 (2018).
52. Vyleta, N. P. & Jonas, P. Loose coupling between Ca²⁺ channels and release sensors at a plastic hippocampal synapse. *Science* **343**, 665–670 (2014).
53. Lorach, H. *et al.* Walking naturally after spinal cord injury using a brain–spine interface. *Nature* **618**, 126–133 (2023).
54. Milekovic, T. *et al.* A spinal cord neuroprosthesis for locomotor deficits due to Parkinson’s disease. *Nature Medicine* **29**, 2854–2865 (2023).
55. Patrick-Krueger, K. M., Burkhart, I. & Contreras-Vidal, J. L. The state of clinical trials of implantable brain–computer interfaces. *Nature Reviews Bioengineering*, 1–18 (2024).
56. Bickenbach, J., Officer, A., Shakespeare, T., von Groote, P., Organization, W. H., *et al.* *International perspectives on spinal cord injury* (World Health Organization, 2013).
57. Skarpaas, T. L. & Morrell, M. J. Intracranial stimulation therapy for epilepsy. *Neurotherapeutics* **6**, 238–243 (2009).
58. Herron, J. *et al.* The convergence of neuromodulation and brain–computer interfaces. *Nature Reviews Bioengineering*, 1–3 (2024).
59. Romeni, S., Valle, G., Mazzoni, A. & Micera, S. Tutorial: a computational framework for the design and optimization of peripheral neural interfaces. *Nature Protocols* **15**, 3129–3153 (2020).
60. Capogrosso, M. *et al.* A computational model for epidural electrical stimulation of spinal sensorimotor circuits. *Journal of Neuroscience* **33**, 19326–19340 (2013).
61. Wagner, F. B. *et al.* Targeted neurotechnology restores walking in humans with spinal cord injury. *Nature* **563**, 65–71 (2018).
62. Valle, G. *et al.* Tactile edges and motion via patterned microstimulation of the human somatosensory cortex. *Science* **387**, 315–322 (2025).
63. Van Hoof, R. *et al.* Optimal placement of high-channel visual prostheses in human retinotopic visual cortex. *Journal of neural engineering* (2025).
64. Shapson-Coe, A. *et al.* A petavoxel fragment of human cerebral cortex reconstructed at nanoscale resolution. *Science* **384**, eadk4858 (2024).
65. Kumaravelu, K. & Grill, W. M. Neural mechanisms of the temporal response of cortical neurons to intracortical microstimulation. *Brain stimulation* **17**, 365–381 (2024).
66. Zhou, A., Johnson, B. C. & Muller, R. Toward true closed-loop neuromodulation: artifact-free recording during stimulation. *Current opinion in neurobiology* **50**, 119–127 (2018).
67. Drebitz, E., Rausch, L.-P. & Kreiter, A. K. A novel approach for removing micro-stimulation artifacts and reconstruction of broad-band neuronal signals. *Journal of Neuroscience Methods* **332**, 108549 (2020).

68. Stoney Jr, S., Thompson, W. & Asanuma, H. Excitation of pyramidal tract cells by intracortical microstimulation: effective extent of stimulating current. *Journal of neurophysiology* **31**, 659–669 (1968).
69. Histed, M. H., Bonin, V. & Reid, R. C. Direct activation of sparse, distributed populations of cortical neurons by electrical microstimulation. *Neuron* **63**, 508–522 (2009).
70. Claar, L. D. *et al.* Cortico-thalamo-cortical interactions modulate electrically evoked EEG responses in mice. *ELife* **12**, RP84630 (2023).
71. Russo, S. *et al.* Thalamic feedback shapes brain responses evoked by cortical stimulation in mice and humans. *bioRxiv* (2024).
72. Borda, L., Gozzi, N., Preatoni, G., Valle, G. & Raspopovic, S. Automated calibration of somatosensory stimulation using reinforcement learning. *Journal of NeuroEngineering and Rehabilitation* **20**, 131 (2023).
73. Willett, F. R., Avansino, D. T., Hochberg, L. R., Henderson, J. M. & Shenoy, K. V. High-performance brain-to-text communication via handwriting. *Nature* **593**, 249–254 (2021).
74. Shaeri, M. *et al.* A 2.46-mm² Miniaturized Brain–Machine Interface (MiBMI) Enabling 31-Class Brain-to-Text Decoding. *IEEE Journal of Solid-State Circuits* (2024).
75. Liu, T. *et al.* Decoding finger velocity from cortical spike trains with recurrent spiking neural networks in 2024 *IEEE Biomedical Circuits and Systems Conference (BioCAS)* (2024), 1–5.
76. Neftci, E. O., Mostafa, H. & Zenke, F. Surrogate gradient learning in spiking neural networks: Bringing the power of gradient-based optimization to spiking neural networks. *IEEE Signal Processing Magazine* **36**, 51–63 (2019).
77. Sourmpis, C., Petersen, C. C., Gerstner, W. & Bellec, G. Biologically informed cortical models predict optogenetic perturbations. *bioRxiv*, 2024–09 (2024).
78. McNeal, D. Analysis of a model for excitation of myelinated nerve. *IEEE transactions on biomedical engineering* **23**, 329–337 (1976).
79. Sweeney, J. D., Ksienski, D. A. & Mortimer, J. T. A nerve cuff technique for selective excitation of peripheral nerve trunk regions. *IEEE Transactions on Biomedical Engineering* **37**, 706–715 (1990).
80. Rattay, F. Analysis of models for external stimulation of axons. *IEEE transactions on biomedical engineering*, 974–977 (1986).
81. Joucla, S. & Yvert, B. The “mirror” estimate: an intuitive predictor of membrane polarization during extracellular stimulation. *Biophysical journal* **96**, 3495–3508 (2009).
82. French, A. P. *Vibrations and Waves* (The M.I.T. Introductory Physics Series, Norton, 1971).
83. Keener, J. P. Waves in excitable media. *SIAM Journal on Applied Mathematics* **39**, 528–548 (1980).
84. Tyson, J. J. & Keener, J. P. Singular perturbation theory of traveling waves in excitable media (a review). *Physica D: Nonlinear Phenomena* **32**, 327–361 (1988).
85. FitzHugh, R. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical journal* **1**, 445–466 (1961).

86. Karuvally, A., Sejnowski, T. & Siegelmann, H. T. *Hidden Traveling Waves bind Working Memory Variables in Recurrent Neural Networks* in *International Conference on Machine Learning* (2024), 23266–23290.
87. Inglebert, Y., Aljadeff, J., Brunel, N. & Debanne, D. Synaptic plasticity rules with physiological calcium levels. *Proceedings of the National Academy of Sciences* **117**, 33639–33648 (2020).
88. Gerstner, W., Kempter, R., Van Hemmen, J. L. & Wagner, H. A neuronal learning rule for sub-millisecond temporal coding. *Nature* **383**, 76–78 (1996).
89. Yaeger, C. E., Mojica Soto-Albors, R., Liu, W., Beltramini, A. & Harnett, M. T. Plateau potentials are instructive signals for behavioral timescale synaptic plasticity in the neocortex. *bioRxiv*, 2025–11 (2025).
90. Caya-Bissonnette, L., Naud, R. & Beique, J.-C. Cellular substrate of eligibility traces. *BioRxiv*, 2023–06 (2023).
91. Caya-Bissonnette, L. & Béique, J.-C. Half a century legacy of long-term potentiation. *Current Biology* **34**, R640–R662 (2024).
92. Bittner, K. C., Milstein, A. D., Grienberger, C., Romani, S. & Magee, J. C. Behavioral time scale synaptic plasticity underlies CA1 place fields. *Science* **357**, 1033–1036 (2017).
93. Pala, A. & Petersen, C. C. In vivo measurement of cell-type-specific synaptic connectivity and synaptic transmission in layer 2/3 mouse barrel cortex. *Neuron* **85**, 68–75 (2015).
94. Ghanbari, A. *et al.* Modeling the short-term dynamics of in vivo excitatory spike transmission. *Journal of Neuroscience* **40**, 4185–4202 (2020).
95. Lynn, M. B. *et al.* A slow 5-HT_{1A}R-mediated recurrent inhibitory network in raphe computes contextual value through synaptic facilitation. *bioRxiv*, 2022–08 (2022).
96. Francioni, V. & Harnett, M. T. Rethinking single neuron electrical compartmentalization: dendritic contributions to network computation in vivo. *Neuroscience* **489**, 185–199 (2022).
97. Díaz-Quesada, M., Martini, F. J., Ferrati, G., Bureau, I. & Maravall, M. Diverse thalamocortical short-term plasticity elicited by ongoing stimulation. *Journal of Neuroscience* **34**, 515–526 (2014).
98. Lee, J. H., Campagnola, L., Seeman, S. C., Jarsky, T. & Mihalas, S. Functional synapse types via characterization of short-term synaptic plasticity. *bioRxiv*, 648725 (2019).
99. Testa-Silva, G. *et al.* High bandwidth synaptic communication and frequency tracking in human neocortex. *PLoS biology* **12**, e1002007 (2014).
100. Bugeon, S. *et al.* A transcriptomic axis predicts state modulation of cortical interneurons. *Nature* **607**, 330–338 (2022).
101. Tasic, B. *et al.* Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nature neuroscience* **19**, 335–346 (2016).
102. Tasic, B. *et al.* Shared and distinct transcriptomic cell types across neocortical areas. *Nature* **563**, 72–78 (2018).

103. Endres, S. C., Sandrock, C. & Focke, W. W. A simplicial homology algorithm for Lipschitz optimisation. *Journal of Global Optimization* **72**, 181–217 (2018).
104. Pozzorini, C. *et al.* Automated high-throughput characterization of single neurons by means of simplified spiking models. *PLoS Comp. Biol.* **11**, e1004275 (2015).
105. Ankerst, M., Breunig, M. M., Kriegel, H.-P. & Sander, J. OPTICS: Ordering points to identify the clustering structure. *ACM Sigmod record* **28**, 49–60 (1999).
106. Hodge, R. D. *et al.* Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573**, 61–68 (2019).
107. Varela, J. A. *et al.* A quantitative description of short-term plasticity at excitatory synapses in layer 2/3 of rat primary visual cortex. *Journal of Neuroscience* **17**, 7926–7940 (1997).
108. Oswald, A.-M. M., Lewis, J. E. & Maler, L. Dynamically interacting processes underlie synaptic plasticity in a feedback pathway. *Journal of neurophysiology* **87**, 2450–2463 (2002).
109. Evstratova, A., Chamberland, S., Faundez, V. & Tóth, K. Vesicles derived via AP-3-dependent recycling contribute to asynchronous release and influence information transfer. *Nature communications* **5**, 5530 (2014).
110. Harris, J. J., Jolivet, R. & Attwell, D. Synaptic energy use and supply. *Neuron* **75**, 762–777 (2012).
111. Chamberland, S. *et al.* Functional specialization of hippocampal somatostatin-expressing interneurons. *bioRxiv* (2023).
112. Costa, R. P., Sjöström, P. J. & Van Rossum, M. C. Probabilistic inference of short-term synaptic plasticity in neocortical microcircuits. *Frontiers in computational neuroscience* **7**, 75 (2013).
113. Kim, M.-H. *et al.* Target cell-specific synaptic dynamics of excitatory to inhibitory neuron connections in supragranular layers of human neocortex. *Elife* **12**, e81863 (2023).
114. Buonomano, D. & Carvalho, T. P. A novel learning rule for long-term plasticity of short-term synaptic plasticity enhances temporal processing. *Frontiers in Integrative Neuroscience*, 20 (2011).
115. Chamberland, S., Evstratova, A. & Tóth, K. Short-term facilitation at a detonator synapse requires the distinct contribution of multiple types of voltage-gated calcium channels. *Journal of Neuroscience* **37**, 4913–4927 (2017).
116. Van der Maaten, L. & Hinton, G. Visualizing data using t-SNE. *Journal of machine learning research* **9** (2008).
117. McInnes, L., Healy, J. & Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426* (2018).
118. Bykowska, O. *et al.* Model-based inference of synaptic transmission. *Frontiers in synaptic neuroscience* **11**, 21 (2019).
119. Hunter, J. D. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering* **9**, 90–95 (2007).
120. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585**, 357–362 (2020).

121. Van Rossum, G. & Drake, F. L. *Python 3 Reference Manual* ISBN: 1441412697 (CreateSpace, Scotts Valley, CA, 2009).
122. Virtanen, P. *et al.* SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* **17**, 261–272 (2020).
123. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011).
124. Li, J. *et al.* Differential synaptic depression mediates the therapeutic effect of deep brain stimulation. *Nature Neuroscience*, 1–13 (2025).
125. Bergey, G. K. *et al.* Long-term treatment with responsive brain stimulation in adults with refractory partial seizures. *Neurology* **84**, 810–817 (2015).
126. Hughes, C., Chen, X., Grill, W. & Kozai, T. D. Neural mechanisms underlying intracortical microstimulation for sensory restoration. *Nature Biomedical Engineering*, 1–17 (2026).
127. Penfield, W. & Boldrey, E. Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain: A journal of neurology* (1937).
128. Matyas, F. *et al.* Motor control by sensory cortex. *Science* **330**, 1240–1243 (2010).
129. Duffau, H. Stimulation mapping of white matter tracts to study brain functional connectivity. *Nature Reviews Neurology* **11**, 255–265 (2015).
130. Kara, P., Pezaris, J. S., Yurgenson, S. & Reid, R. C. The spatial receptive field of thalamic inputs to single cortical simple cells revealed by the interaction of visual and electrical stimulation. *Proceedings of the National Academy of Sciences* **99**, 16261–16266 (2002).
131. Moore, T. & Armstrong, K. M. Selective gating of visual signals by microstimulation of frontal cortex. *eng. Nature* **421**, 370–3 (2003).
132. Moser, E., Mathiesen, I. & Andersen, P. Association between brain temperature and dentate field potentials in exploring and swimming rats. *Science* **259**, 1324–1326 (1993).
133. Doron, G. *et al.* Perirhinal input to neocortical layer 1 controls learning. *Science* **370**, eaaz3136 (2020).
134. Stuart, G. J. & Sakmann, B. Active propagation of somatic action potentials into neocortical pyramidal cell dendrites. *Nature* **367**, 69–72 (1994).
135. Stevens, C. F. & Wang, Y. Facilitation and depression at single central synapses. *Neuron* **14**, 795–802 (1995).
136. Murphy, B. K. & Miller, K. D. Balanced amplification: a new mechanism of selective amplification of neural activity patterns. *Neuron* **61**, 635–648 (2009).
137. Diesmann, M., Gewaltig, M.-O. & Aertsen, A. Stable propagation of synchronous spiking in cortical neural networks. *Nature* **402**, 529–533 (1999).
138. Buzsáki, G. & Wang, X.-J. Mechanisms of gamma oscillations. *Annual review of neuroscience* **35**, 203–225 (2012).

139. Pinto, D. J. & Ermentrout, G. B. Spatially structured activity in synaptically coupled neuronal networks: I. Traveling fronts and pulses. *SIAM journal on Applied Mathematics* **62**, 206–225 (2001).
140. Nowak, L. & Bullier, J. Axons, but not cell bodies, are activated by electrical stimulation in cortical gray matter i. evidence from chronaxie measurements: I. Evidence from chronaxie measurements. *Experimental brain research* **118**, 477–488 (1998).
141. Logothetis, N. K. *et al.* The effects of electrical microstimulation on cortical signal propagation. *Nature neuroscience* **13**, 1283–1291 (2010).
142. Ferezou, I. *et al.* Spatiotemporal dynamics of cortical sensorimotor integration in behaving mice. *Neuron* **56**, 907–923 (2007).
143. Tanaka, Y. *et al.* Focal activation of neuronal circuits induced by microstimulation in the visual cortex. *Journal of Neural Engineering* **16**, 036007 (2019).
144. Butovas, S., Hormuzdi, S. G., Monyer, H. & Schwarz, C. Effects of electrically coupled inhibitory networks on local neuronal responses to intracortical microstimulation. *Journal of neurophysiology* **96**, 1227–1236 (2006).
145. Jankowska, E., Padel, Y. & Tanaka, R. The mode of activation of pyramidal tract cells by intracortical stimuli. *The Journal of physiology* **249**, 617–636 (1975).
146. Pinto, D. J., Patrick, S. L., Huang, W. C. & Connors, B. W. Initiation, propagation, and termination of epileptiform activity in rodent neocortex in vitro involve distinct mechanisms. *Journal of Neuroscience* **25**, 8131–8140 (2005).
147. Ranck, J. Extracellular stimulation. *Electrical stimulation research techniques*, 1–36 (1981).
148. McIntyre, C. C. & Grill, W. M. Selective microstimulation of central nervous system neurons. *Annals of biomedical engineering* **28**, 219–233 (2000).
149. Aberra, A. S., Peterchev, A. V. & Grill, W. M. Biophysically realistic neuron models for simulation of cortical stimulation. *Journal of neural engineering* **15**, 066023 (2018).
150. West, D. & Wolstencroft, J. Strength-duration characteristics of myelinated and non-myelinated bulbospinal axons in the cat spinal cord. *The Journal of Physiology* **337**, 37–50 (1983).
151. Li, C. & Bak, A. Excitability characteristics of the A-and C-fibers in a peripheral nerve. *Experimental neurology* **50**, 67–79 (1976).
152. Stuart, G., Spruston, N. & Häusser, M. Action potential initiation and backpropagation in the mammalian CNS. *Trends in Neuroscience* **20**, 125–131 (1997).
153. Beaulieu-Laroche, L. *et al.* Enhanced dendritic compartmentalization in human cortical neurons. *Cell* **175**, 643–651 (2018).
154. Dadarlat, M. C., Sun, Y. J. & Stryker, M. P. Activity-dependent recruitment of inhibition and excitation in the awake mammalian cortex during electrical stimulation. *Neuron* **112**, 821–834 (2024).
155. Dadarlat, M. C., Sun, Y. & Stryker, M. P. *Widespread activation of awake mouse cortex by electrical stimulation in 2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)* (2019), 1113–1117.

156. Hughes, C. & Kozai, T. Dynamic amplitude modulation of microstimulation evokes biomimetic onset and offset transients and reduces depression of evoked calcium responses in sensory cortices. *Brain stimulation* **16**, 939–965 (2023).
157. Tolias, A. S. *et al.* Mapping cortical activity elicited with electrical microstimulation using FMRI in the macaque. *Neuron* **48**, 901–911 (2005).
158. Huang, C. *et al.* Circuit models of low-dimensional shared variability in cortical networks. *Neuron* **101**, 337–348 (2019).
159. Wilson, H. R. & Cowan, J. D. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical journal* **12**, 1–24 (1972).
160. Vieira, P. G., Krause, M. R., Laamerad, P. & Pack, C. C. Brain stimulation preferentially influences long-range projections. *Science Advances* **11**, eadx2106 (2025).
161. Palkar, G., Wu, J.-y. & Ermentrout, B. The inhibitory control of traveling waves in cortical networks. *PLoS computational biology* **19**, e1010697 (2023).
162. Davis, Z. W. *et al.* Spontaneous traveling waves naturally emerge from horizontal fiber time delays and travel through locally asynchronous-irregular states. *Nature Communications* **12**, 6057 (2021).
163. Moldakarimov, S., Bazhenov, M., Feldman, D. E. & Sejnowski, T. J. Structured networks support sparse traveling waves in rodent somatosensory cortex. *Proceedings of the National Academy of Sciences* **115**, 5277–5282 (2018).
164. Meijer, H. G. & Coombes, S. Travelling waves in a neural field model with refractoriness. *Journal of mathematical biology* **68**, 1249–1268 (2014).
165. Nevian, T., Larkum, M. E., Polsky, A. & Schiller, J. Properties of basal dendrites of layer 5 pyramidal neurons: a direct patch-clamp recording study. *Nat Neurosci* **10**, 206–214. ISSN: 1097-6256 (Print) (2 2007).
166. Stuart, G., Schiller, J & Sakmann, B. Action potential initiation and propagation in rat neocortical pyramidal neurons. *J Physiol* **505**, 617–632. ISSN: 0022-3751 (Print) (1997).
167. McIntyre, C. C. & Grill, W. M. Excitation of central nervous system neurons by nonuniform electric fields. *Biophysical journal* **76**, 878–888 (1999).
168. Matsumoto, G & Tasaki, I. A study of conduction velocity in nonmyelinated nerve fibers. *Biophysical journal* **20**, 1–13 (1977).
169. Schmidt, H. *et al.* Axonal synapse sorting in medial entorhinal cortex. *Nature* **549**, 469–475 (2017).
170. Micheva, K. D. *et al.* A large fraction of neocortical myelin ensheathes axons of local inhibitory neurons. *elife* **5**, e15784 (2016).
171. Micheva, K. D. *et al.* Distinctive structural and molecular features of myelinated inhibitory axons in human neocortex. *Eneuro* **5** (2018).
172. Avermann, M., Tömm, C., Mateo, C., Gerstner, W. & Petersen, C. C. Microcircuits of excitatory and inhibitory neurons in layer 2/3 of mouse barrel cortex. *Journal of neurophysiology* **107**, 3116–3134 (2012).

173. Stedehouder, J. *et al.* Local axonal morphology guides the topography of interneuron myelination in mouse and human neocortex. *Elife* **8**, e48615 (2019).
174. Campagnola, L. *et al.* Local connectivity and synaptic dynamics in mouse and human neocortex. *Science* **375**, eabj5861 (2022).
175. Shusterman, V. & Troy, W. C. From baseline to epileptiform activity: a path to synchronized rhythmicity in large-scale neural networks. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics* **77**, 061911 (2008).
176. Schwalger, T., Deger, M. & Gerstner, W. Towards a theory of cortical columns: From spiking neurons to interacting neural populations of finite size. *PLoS computational biology* **13**, e1005507 (2017).
177. Rosenbaum, R. & Doiron, B. Balanced networks of spiking neurons with spatially dependent re-current connections. *Physical Review X* **4**, 021039 (2014).
178. Connors, B. W. & Gutnick, M. J. Intrinsic firing patterns of diverse neocortical neurons. *Trends in neurosciences* **13**, 99–104 (1990).
179. Davis, Z. W., Busch, A., Steward, C., Muller, L. & Reynolds, J. Horizontal cortical connections shape intrinsic traveling waves into feature-selective motifs that regulate perceptual sensitivity. *Cell reports* **43** (2024).
180. Beaulieu-Laroche, L. *et al.* Allometric rules for mammalian cortical layer 5 neuron biophysics. *Nature* **600**, 274–278 (2021).
181. Rattay, F. Analysis of models for external stimulation of axons. *IEEE transactions on biomedical engineering*, 974–977 (1986).
182. Gerstner, W., Kistler, W., Naud, R. & Paninski, L. *Neuronal Dynamics* (Cambridge University Press, Cambridge, 2014).
183. Swadlow, H. A. Monitoring the excitability of neocortical efferent neurons to direct activation by extracellular current pulses. *Journal of neurophysiology* **68**, 605–619 (1992).
184. Naud, R. & Longtin, A. Linking demyelination to compound action potential dispersion with a spike-diffuse-spike approach. *The Journal of Mathematical Neuroscience* **9**, 3 (2019).
185. Dale, H. *Pharmacology and nerve-endings* 1935.
186. Eccles, J. C., Fatt, P & Koketsu, K. Cholinergic and inhibitory synapses in a pathway from motor-axon collaterals to motoneurons. *The Journal of physiology* **126**, 524 (1954).
187. Coombs, J. S., Eccles, J. C. & Fatt, P. Excitatory synaptic action in motoneurons. *The Journal of physiology* **130**, 374 (1955).
188. Paszke, A. *et al.* Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems* **32** (2019).
189. French, A. P. *Vibrations and waves* (CRC press, 2017).
190. Muller, L., Chavane, F., Reynolds, J. & Sejnowski, T. J. Cortical travelling waves: mechanisms and computational principles. *Nature Reviews Neuroscience* **19**, 255–268 (2018).

191. Ortuño, T., Grieve, K. L., Cao, R., Cudeiro, J. & Rivadulla, C. Bursting thalamic responses in awake monkey contribute to visual detection and are modulated by corticofugal feedback. *Frontiers in behavioral neuroscience* **8**, 198 (2014).
192. Krook-Magnuson, E., Armstrong, C., Oijala, M. & Soltesz, I. On-demand optogenetic control of spontaneous seizures in temporal lobe epilepsy. *Nature communications* **4**, 1376 (2013).
193. Hadjiabadi, D. *et al.* Maximally selective single-cell target for circuit control in epilepsy models. *Neuron* **109**, 2556–2572 (2021).
194. Murphy, K. R. *et al.* A tool for monitoring cell type-specific focused ultrasound neuromodulation and control of chronic epilepsy. *Proceedings of the National Academy of Sciences* **119**, e2206828119 (2022).
195. Stuck, M. & Naud, R. *Burstprop for learning in spiking neuromorphic hardware* in *Proceedings of the 2023 International Conference on Neuromorphic Systems* (2023), 1–5.
196. Schuman, C. D. *et al.* Opportunities for neuromorphic computing algorithms and applications. *Nature Computational Science* **2**, 10–19 (2022).
197. Bender, F. *et al.* Synaptic dynamics as a tunable substrate shaping neuronal activity sequences. *bioRxiv*, 2026–03 (2026).
198. Brown, P. Bad oscillations in Parkinson’s disease. *Journal of neural transmission Supplementum* **70**, 27 (2006).
199. Wang, D. D. *et al.* Pallidal deep-brain stimulation disrupts pallidal beta oscillations and coherence with primary motor cortex in Parkinson’s disease. *Journal of Neuroscience* **38**, 4556–4568 (2018).
200. Drouot, X. *et al.* Functional recovery in a primate model of Parkinson’s disease following motor cortex stimulation. *Neuron* **44**, 769–778 (2004).
201. Beuter, A., Lefaucheur, J.-P. & Modolo, J. Closed-loop cortical neuromodulation in Parkinson’s disease: An alternative to deep brain stimulation? *Clinical Neurophysiology* **125**, 874–885 (2014).
202. Underwood, C. F. & Parr-Brownlie, L. C. Primary motor cortex in Parkinson’s disease: Functional changes and opportunities for neurostimulation. *Neurobiology of Disease* **147**, 105159 (2021).
203. Rizzone, M. *et al.* Deep brain stimulation of the subthalamic nucleus in Parkinson’s disease: effects of variation in stimulation parameters. *Journal of Neurology, Neurosurgery & Psychiatry* **71**, 215–219 (2001).
204. Herrington, T. M., Cheng, J. J. & Eskandar, E. N. Mechanisms of deep brain stimulation. *Journal of neurophysiology* **115**, 19–38 (2016).