

# UC Berkeley

## UC Berkeley Electronic Theses and Dissertations

### Title

Constraining Ill-Posed Inverse Problems in Neural Electrophysiology via Biophysically Detailed Forward Simulation

### Permalink

<https://escholarship.org/uc/item/3w84c5vz>

### Author

Baratham, Vyassa

### Publication Date

2021

Peer reviewed|Thesis/dissertation

Constraining Ill-Posed Inverse Problems in Neural Electrophysiology via Biophysically  
Detailed Forward Simulation

by

Vyassa Baratham

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Adjunct Professor Kristofer Bouchard, Co-chair

Associate Professor Michael DeWeese, Co-chair

Professor Daniel Feldman

Associate Professor Na Ji

Summer 2021

Constraining Ill-Posed Inverse Problems in Neural Electrophysiology via Biophysically  
Detailed Forward Simulation

Copyright 2021  
by  
Vyassa Baratham

# Abstract

## Constraining Ill-Posed Inverse Problems in Neural Electrophysiology via Biophysically Detailed Forward Simulation

by

Vyassa Baratham

Doctor of Philosophy in Physics

University of California, Berkeley

Adjunct Professor Kristofer Bouchard, Co-chair

Associate Professor Michael DeWeese, Co-chair

Biophysically detailed simulation is an invaluable tool for understanding experimental data when those data do not uniquely determine the underlying state of the system, a situation we refer to as an *ill-posed inverse problem*. Such problems arise frequently in the study of biological systems with many degrees of freedom. This dissertation presents simulation-based approaches to two ill-posed inverse problems in neural electrophysiology. First, using a large volume of simulated data, we demonstrate that a Convolutional Neural Network can be trained to determine the conductances of various ion channels in a neuron from its somatic membrane potential in response to a current injection. Next, we use a simulation to study the cellular origin of electrical signals recorded at the surface of the brain, and find that they are produced primarily in layers V and VI of the cortex, contrary to the intuition that neurons closer to the electrode should contribute more of the signal. In both cases, simulation is a natural way to incorporate biological constraints to rule out certain *a priori* plausible solutions. Our results show how the massive throughput, fine-grained control over model parameters, and access to underlying ground-truth details within a simulation can be utilized to overcome the ill-posedness that many biological problems exhibit when stated in physical terms.



Dedicated to Sanjaya B.

# Contents

<b>Contents</b>	<b>ii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Ill-posed Inverse Problems . . . . .	1
1.2 Inverse Problems in Neural Electrophysiology . . . . .	1
1.2.1 Toy Model of Biological Constraints on an Inverse Problem . . . . .	2
1.3 Biophysically Detailed Forward Models Incorporate Biological Constraints .	4
1.4 Computational Simulation of Neural Activity . . . . .	5
1.4.1 Cable Equation; NEURON Simulation Environment . . . . .	5
1.4.2 Blue Brain Project . . . . .	5
<b>2 Inference of Ion Channel Conductance from Somatic Membrane Potential</b>	<b>6</b>
2.1 Background and Motivation . . . . .	6
2.2 Forward Model . . . . .	6
2.2.1 Neuronal models . . . . .	7
2.2.1.1 Izhikevich - phenomenological spiking neuron model . . . . .	7
2.2.1.2 Hodgkin-Huxley - ball and stick compartmental model . . . . .	8
2.2.1.3 Mainen and Sejnowski - biophysically detailed compartmental model . . . . .	8
2.2.1.4 Blue Brain Project - biophysically detailed compartmental models . . . . .	8
2.2.2 Stimulation . . . . .	9
2.3 Inverse Model . . . . .	9
2.4 Results . . . . .	11
2.4.1 CNNs can accurately predict ionic conductances in different neuronal models . . . . .	11
2.4.1.1 Izhikevich : CNNs can regress parameters from a phenomenological point cell model . . . . .	11
2.4.1.2 Hodgkin-Huxley : CNNs can regress conductances in remote dendritic compartments . . . . .	12
2.4.1.3 Mainen and Sejnowski: CNNs can regress conductances in a biophysical model . . . . .	14

2.4.1.4	BBP: CNNs can regress most conductances in an advanced biophysical model . . . . .	15
2.4.1.5	BBP: CNNs show promise to regress conductances across cells . . . . .	15
2.4.2	Accurate CNN predictions qualitatively reproduce input waveform . . . . .	18
2.5	Discussion . . . . .	19
2.5.1	CNNs are able to regress ionic conductances from somatic potentials . . . . .	19
2.5.2	Inferring ionic conductances in biophysically detailed models requires significant computational effort . . . . .	19
2.6	Acknowledgements . . . . .	20
2.7	Tables of parameter values and CNN prediction accuracy . . . . .	20
<b>3</b>	<b>Reconstructing Neuronal Sources of Cortical Surface Electrical Potentials</b>	<b>23</b>
3.1	Background and Motivation . . . . .	23
3.2	Forward Model . . . . .	23
3.2.1	Cortical Column Simulation . . . . .	23
3.2.2	Computation of CSEP from simulated neuronal activity . . . . .	25
3.3	Results . . . . .	25
3.3.1	Stimulus-evoked cortical surface electrical potentials exhibit large peaks in the high-gamma range . . . . .	25
3.3.2	Biophysical <i>in silico</i> cortical column reproduces <i>in vivo</i> observed ECoG response . . . . .	27
3.3.3	<i>In silico</i> cortical column predicts experimentally observed relationship between response magnitude and frequency . . . . .	29
3.3.4	Evoked $\mu$ ECoG responses originate in infragranular layers . . . . .	30
3.3.5	Evoked $\mu$ ECoG responses originate in sources 800-1400 $\mu$ m below the surface . . . . .	32
3.4	Discussion . . . . .	34
<b>4</b>	<b>Conclusion</b>	<b>35</b>
4.1	Biophysical simulation is a valuable tool for studying ill-posed inverse problems in neural electrophysiology . . . . .	35
4.2	Future Work . . . . .	36
4.2.1	<i>In vitro</i> inference of ion channel conductance . . . . .	36
4.2.2	Probing deeper into the origins of ECoG signals . . . . .	37
4.2.3	Explicit inverse model of CSEPs . . . . .	37
4.2.4	Finite Element Model of neural activity and extracellular potential . . . . .	38
	<b>Bibliography</b>	<b>40</b>
	<b>A Derivation of Line Source Approximation</b>	<b>43</b>
	<b>B Computation of CSEP contributions from simulation</b>	<b>45</b>

B.1	Laminar contributions . . . . .	45
B.2	200 $\mu\text{m}$ slice contributions . . . . .	45

# Chapter 1

## Introduction

### 1.1 Ill-posed Inverse Problems

When studying any system of nontrivial complexity, experimental observations may not uniquely determine the underlying state of the system - there may instead be a family of underlying states consistent with any experimental observation. The task of determining which state from within that family actually gave rise to the observation is deemed an *ill-posed inverse problem* [14]. Many ill-posed inverse problems arise when biological systems are described in terms of the physics of their constituents. That is, in biological systems there are often many physically possible realizations of a given high-level observation. Intuitively, simulations should be a useful tool for studying such inverse problems, provided that they model the system's biophysics with sufficient detail to capture the variety of ways it could produce a given high-level observation. Specifically, we propose that biology imposes constraints on this variety, and that a simulation of sufficient biophysical realism to capture those constraints can be used to gain insight into which of the possibilities is/are employed by Nature.

### 1.2 Inverse Problems in Neural Electrophysiology

Neural electrophysiology - the observation of electrical signals produced by the brain - has long been a key method in the field of neuroscience, and has provided countless insights into brain function and organization. However, as neuroscientists probe finer structures of the brain at higher resolution, use of population-level electrophysiological methods becomes challenging due to the fact that electrical recordings typically reflect a superposition of the activity of all neurons in the vicinity of the recording electrode, limiting the spatial and neuronal resolution at which we can investigate the brain using such methods. One approach to overcoming this limitation is to use the superposed observed signal to reconstruct the underlying contributions from distinct sub-populations of neurons. This is an example of an ill-posed inverse problem because there exist many different distributions of activity across

individual neurons or sub-populations that would give rise to the same observed population-level signal. In the absence of external information, such as a detailed model of the signal generation, it is generally not clear *a priori* which of these solutions reflects the true neural activity distribution at the time of recording.

Oftentimes, similarly ill-posed problems in neuroscience attain unique solutions with the help of biological knowledge or measurements, such as the dynamics of ion channels, the connectivity between various types of neurons, or the density of neurons in a particular region of the brain. Section 1.2.1 provides a toy example of how "biological" constraints of this sort may help obtain a unique solution to an ill-posed inverse problem. The following chapters 2 and 3 present two ill-posed inverse problems in neural electrophysiology which were studied using a biophysically detailed computer simulation.

### 1.2.1 Toy Model of Biological Constraints on an Inverse Problem

To begin to understand how biological constraints can help mitigate the ill-posedness of a given inverse problem, we now consider a toy example which roughly resembles the problem treated in chapter 3, that of inferring the neuronal activity which gave rise to a given Cortical Surface Electrical Potential (CSEP). The problem is to use the CSEP  $y_{\text{tot}}$  recorded at the cortical surface at one point in time to recover the activity  $x_i$  of  $n$  neuronal sources located at distances  $r_i$  below the surface at that same point in time. For this example, we consider a simplified forward model in which the contribution of each source to the signal recorded at the surface is given by  $y_i = x_i/r_i$ , and the contributions superpose linearly so that  $y_{\text{tot}} = \sum_i y_i$ . We seek the activity levels  $x_i$  of the  $n$  sources such that  $y_{\text{tot}} = 2$ .

To summarize, this ill-posed inverse problem asks us to solve for each of the  $x_i$ 's in the following equation:

$$\sum_{i=1}^n \frac{x_i}{r_i} = 2 \quad (1.1)$$

Clearly, finding the  $x_i$ 's that gave rise to  $y_{\text{tot}} = 2$  will require some information beyond just that observation of  $y$ . To begin with, we do not have the values of  $n$  and the  $r_i$ 's. Thus, the first piece of biological knowledge we incorporate into our solution of this inverse problem is:

**Constraint 1:** The number of neurons contributing to the CSEP is  $n = 3$ .

Here we have chosen a tractable but unrealistically small number of neurons contributing to the surface signal. This allows us to begin to specify the possible solution set, as we now know that the  $x_i$ 's will form an ordered triple. After Constraint 1, the inverse problem is to find  $(x_1, x_2, x_3)$  such that

$$\frac{x_1}{r_1} + \frac{x_2}{r_2} + \frac{x_3}{r_3} = 2 \quad (1.2)$$

Currently, there is no ordered triple that would not be able to solve equation 1.2, given that there are as yet no constraints on the  $r_i$ 's. Put another way, the solution set is still far too

big to be of any practical use. Specifically, Constraint 1 (trivially) limits our solution to:

**After Constraint 1:**  $(x_1, x_2, x_3) \in \mathbb{R}^3$

Next, we will specify the  $r_i$ 's:

**Constraint 2:** The 3 neurons are located at distances  $r_1 = 1$ ,  $r_2 = 2$ ,  $r_3 = 3$  below the surface.

After Constraint 2, the inverse problem looks like:

$$x_1 + \frac{x_2}{2} + \frac{x_3}{3} = 2 \quad (1.3)$$

And the solution set can be expressed parametrically as:

**After Constraint 2:**  $(x_1, x_2, x_3) \in \{(s, t, 6 - 3s - \frac{3}{2}t) \mid s, t \in \mathbb{R}^2\}$

While the two previous pieces of biological knowledge were explicitly present in the mathematical representation of our inverse problem, other constraints affect the system more indirectly. Our next constraint describes the neural activity itself, and is loosely inspired by the first-approximation observation that at any given point in time, neurons are either firing, or in a resting state:

**Constraint 3:** The activity level of a neuron must be a non-negative integer.

Constraint 3 does not change the mathematical form of the inverse problem from equation 1.3, but it permits only 5 solutions, so the entire solution set can now be specified explicitly:

**After Constraint 3:**  $(x_1, x_2, x_3) \in \{(0, 0, 6), (0, 2, 3), (0, 4, 0), (1, 2, 0), (2, 0, 0)\}$

In summary, by incorporating constraints about the biology of the system into our forward model, we have seen the solution set to this toy problem change from all ordered triples, to a 2 dimensional space of triples, to just 5 specific possibilities. The drastic reduction in the solution set's dimensionality and cardinality previews the effects of our use of biophysically detailed forward modeling to constrain similarly ill-posed inverse problems in chapters 2 and 3.

## 1.3 Biophysically Detailed Forward Models Incorporate Biological Constraints

While the toy model of section 1.2.1 is simple enough that we may explicitly write down the effects of each constraint on the permitted solution set (i.e. it is simple to show mathematically which solutions are inconsistent with each constraint), biological knowledge is not always as mathematically convenient as this. For example, much is known about the electrical properties of neuronal membranes: their capacitance per unit area, their physical structure and dimensions, the voltage dependence of their various ion channels, and how all these properties vary with cell type. These properties put constraints on the electrical phenomena that neurons can exhibit, but it is highly nontrivial to understand how those constraints affect downstream observable properties of the system, such as the electrical potentials recorded at the surface.

One way to incorporate biological knowledge when solving an ill-posed inverse problem is to implement a biophysically detailed forward model of the phenomenon you would like to invert. In a forward model, any biological constraint can be implemented as long as it is understood how that constraint interacts with the biophysics of the model. For example, knowledge of a cell's location fixes the coordinates of that cell wherever they appear in the model's computations of biophysics properties (for example, in the forward model presented in chapter 3, these coordinates appear in the denominator of equation 3.2). More detailed biological knowledge such as the voltage dependence of an ion channel has more complicated interactions with the biophysics; accordingly, more complex logic is required to compute the effects of these interactions. In our simulations of neural activity, we use ion channel dynamics programmed using the NMODL model description language [17]. Compilation of an NMODL source file results in a dynamically linked library which can be used by NEURON at runtime to simulate the channel.

Once constrained appropriately, the forward model can be left to run, and the behavior of the system will tell us something about the effects of the constraints. The final step is to extract this information in a form that provides insight into the inverse problems that inspired our biophysically detailed modeling to begin with. Chapters 2 and 3 describe how biophysically detailed simulations were used to approach two ill-posed inverse problems related to neural electrophysiology. Chapter 2 shows that simulations enable fine-grained parameter control and extremely high throughput which permit the development of an explicit inverse model that is able to reconstruct the densities of several different ion channels throughout the morphology of a neuron just by reading the somatic membrane potential during an amplitude-modulated current injection. Chapter 3 shows how the experimentally inaccessible ground truth information available from a simulation can teach us about the neuronal origins of the Cortical Surface Electrical Potential (CSEP) recorded by micro-electrocorticography ( $\mu$ ECoG).

Our approaches to the problems presented in chapters 2 and 3 employ biophysically detailed simulations of the activity of neurons in the brain. The following section provides the



necessary background on the biophysics of neurons and the software tools used to implement our simulations, with more details presented in subsequent chapters as necessary.

## 1.4 Computational Simulation of Neural Activity

In the following two chapters, real inverse problems are treated with the help of biophysically detailed simulations of individual neurons or populations of neurons. In these problems, we are primarily interested in the electrical signals produced by these neurons, which arise from their interactions with the charged ions of the extracellular medium. These interactions are mediated by the cell membrane, a phospholipid bilayer studded with ion channels which selectively permit or block ions from entering or exiting the cell. The bulk of the membrane itself also serves as a dielectric between the intracellular and extracellular spaces, allowing the membrane to separate charges and thus act as a capacitor.

### 1.4.1 Cable Equation; NEURON Simulation Environment

The forward models used in the following two chapters are implemented in the NEURON simulation environment [10]. NEURON simulates compartmental models, in which neuron(s) are modeled as a series of small cylindrical segments over which the membrane potential (the drop in electrical potential experienced when crossing the neuronal membrane) and currents through the membrane can be taken as approximately constant. The evolution of the membrane potential is given by the Cable Equation:

$$c_m \frac{\partial V}{\partial t} + i_m(x, t) = \frac{1}{r} \frac{\partial^2 V}{\partial x^2} \quad (1.4)$$

Where  $c_m$  is the membrane capacitance per unit length,  $r$  is the axial resistance of the compartment per unit length,  $V$  is the membrane potential, and  $i_m$  is the current per unit length entering or leaving the compartment through the membrane.

### 1.4.2 Blue Brain Project

The neuronal models used here were originally implemented by the Blue Brain Project (BBP) in a model of a small piece of rat sensory cortex. Consisting of about 30,000 neurons whose biophysical properties were tuned in accordance with a large body of experimental measurements, the BBP model represents the current state of the art in biophysically detailed simulation of large neural populations. Parts of the model have been released to the public, including 5 experimentally tuned reconstructed neurons (or algorithmically generated clones thereof) representing each of the 210 cell types in the column. These cell models play key roles in chapters 2 and 3. In chapter 3 we also use the rest of the publicly available BBP data to reproduce the entire model in our own software stack.

## Chapter 2

# Inference of Ion Channel Conductance from Somatic Membrane Potential

## 2.1 Background and Motivation

Understanding the nature of computations performed by single neurons (i.e., how individual neurons integrate their inputs) will require a detailed picture of electrical properties throughout the neuron’s morphology, including ionic conductances (directly related to the density of ion channels in the membrane), membrane resistance, etc. However, these microscopic variables are experimentally difficult or impossible to measure directly. Thus, we seek to infer their values from measurements of more accessible properties, such as the membrane potential recorded at the soma. This inverse problem is ill-posed: in principle, there may be many sets of generative parameters (ionic conductances, membrane resistances, etc.) which give rise to the same somatic membrane potential. Despite this, we show here that Convolutional Neural Networks (CNNs) are able to directly regress these generative parameters from time-varying membrane potentials in response to an input stimulus (current injection), thus solving this inverse problem on simulated neurons. Our method is able to perform high-accuracy regressions of microscopic electrical parameters on neuronal models of varying complexity, from a phenomenological model of point neurons described by 4 parameters, up to biophysically detailed neurons including the Blue Brain Project cells described by 30+ parameters including spatial variation. For the latter models, high throughput (beyond what would be practical experimentally) is necessary to produce enough training data to obtain accurate results.

## 2.2 Forward Model

To validate our method of predicting ionic conductances from somatic membrane potentials, we used cable equation simulations via the NEURON simulator as described in section 1.4.1. We applied our method to data obtained from three types of simulated neurons: (i) a

point neuron modeled phenomenologically using the technique of Izhikevich [19], (ii) a toy "ball and stick" morphology obeying Hodgkin-Huxley dynamics, and (iii) a biophysically detailed compartmental reconstruction of a layer V pyramidal neuron released by Mainen and Sejnowski. To evaluate our method's performance on a biophysically diverse set of cells of realistic complexity, we then applied it to the neurons in the Blue Brain Project model.

### 2.2.1 Neuronal models

The neuronal models we used for testing our method reflect a progressive increase in complexity, which allowed us to analyze precisely how the method works in various conditions. For example, the simplest model we used (Izhikevich, section 2.2.1.1) was a point neuron, which allowed us to assess whether our method could regress local membrane parameters. The next simplest (Hodgkin-Huxley, section 2.2.1.2) was a neuron with two compartments, which allowed us to separately assess whether our method could regress membrane parameters describing compartments other than the soma, where the membrane potential (the CNN's input) was recorded. Later models in the progression help us to assess our method on biophysically realistic cells.

Below, we describe each of the neuronal models used. The conductances (denoted  $\bar{g}$  in the tables at the end of this chapter) are *maximum* conductances (over the course of an action potential, the conductance of each channel varies between 0 and  $\bar{g}$ ).

#### 2.2.1.1 Izhikevich - phenomenological spiking neuron model

The Izhikevich model [19] is a remarkably simple phenomenological approximation of a diverse range of neuronal spiking behaviors based on just 4 parameters  $a$ ,  $b$ ,  $c$ , and  $d$ . It does not account for morphological effects; i.e. the Izhikevich model is a point cell model. The Izhikevich model can be used to represent a wide range of firing patterns by varying the parameter values. Spiking neuron models such as Izhikevich are mostly used to represent neurons in a large scale neuronal simulation with many point neurons. The Izhikevich model is not readily interpretable in terms of biophysical properties of the simulated neuron.

To emphasize the simplicity of the Izhikevich model, we reproduce its entire specification here:

$$v' = 0.04v^2 + 5v + 140 - u + I \quad (2.1)$$

$$u' = a(bv - u) \quad (2.2)$$

$$\text{if } v > 30\text{mV, then } v \leftarrow c, u \leftarrow u + d \quad (2.3)$$

The simulated membrane potential is given by  $v$ ,  $u$  is an internal variable accounting for the recovery of the membrane, and  $I$  represents the electrical current input into the cell. In our study, the Izhikevich model is implemented as a `POINT PROCESS` in NMODL, then inserted into a zero-length cellular compartment in the NEURON simulator.

### 2.2.1.2 Hodgkin-Huxley - ball and stick compartmental model

The Hodgkin-Huxley model is one of the first conductance-based biophysical models of neuronal activity to be developed which was capable of explaining the generation and propagation of action potentials using sodium and potassium ions [18]. It does this by keeping track of the activation/deactivation levels of sodium and potassium channels and providing differential equations to specify how these activation/deactivation levels vary as a function of time and the membrane potential.

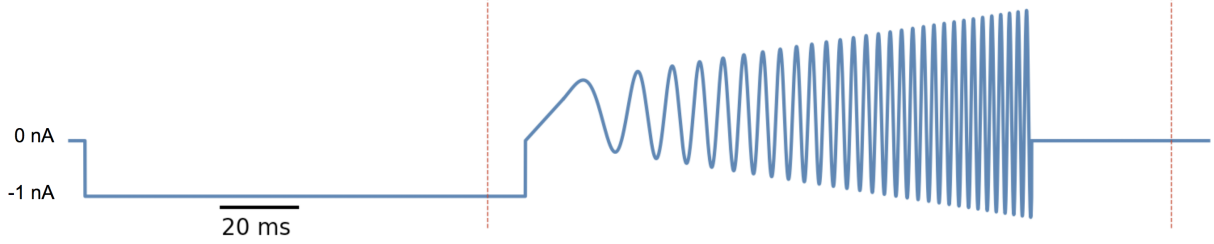
Hodgkin-Huxley dynamics are implemented in NEURON via an NMODL file called `hh.mod` which is included in all standard distributions of the software. Using this channel model, we implemented HH-type dynamics in a sphere connected to a cylinder, representing a simplified soma and dendrite, respectively. The adjustable parameters in our Hodgkin-Huxley 7-parameter model were the maximal ionic conductances of various ion channels in each of the two compartments (for most channels, the conductances are set independently in the two compartments).

### 2.2.1.3 Mainen and Sejnowski - biophysically detailed compartmental model

In one of the first studies to demonstrate the importance of dendritic structure to the spiking behavior of neurons [23], Mainen and Sejnowski used a biophysically detailed compartmental model of a neuron which has since been adopted by the computational neuroscience community in many studies requiring biophysically detailed neuronal models. Thus, it was a natural choice as the first biophysically detailed neuron on which to test our method. In order to examine the model at different levels of complexity, and to facilitate comparison with the Izhikevich and Hodgkin-Huxley neurons, we start by varying only 4 of the 10 free parameters in the model, then increase to 7, before finally varying all 10 parameters. Frozen (non-varied) parameters are held at the original values used by Mainen and Sejnowski.

### 2.2.1.4 Blue Brain Project - biophysically detailed compartmental models

The cells in the Blue Brain Project model represent a significant advance in biophysically detailed compartment models over the cell published by Mainen and Sejnowski. This owes not only to the vastly superior computational resources available today, but also to a large body of experimental observations about the ion channels present in neuronal membranes, and how those channels vary with cell type. The incorporation of these observations resulted in cell models with over 30 parameters describing many types of ion channels gating several ionic species, with different densities (conductances) at different sections of the cell. Importantly, the variety of neuronal models available from BBP captures much of the variability between different neocortical cells in rats.



**Figure 1:** Time course of injected current used to stimulate neuronal activity. Red lines demarcate the time window from which somatic membrane potentials were analyzed.

### 2.2.2 Stimulation

To drive activity in our simulated neurons, we injected current in the form of a chirp shown in figure 1. The chirp stimulus is a sinusoidal wave with continuously varying frequency and amplitude:

$$I(t) = C \cdot (I_0 + I_c \cdot t \cdot \sin(\omega t)) \quad (2.4)$$

where  $I_c = 11$  nA/s,  $\omega = 0.64$  rad/ms,  $I_0 = 0.5$  nA. We added a hyperpolarizing current lasting 110 ms to prevent spontaneous spikes prior to the onset of the stimulating chirp, as well as an 80 ms period of no current injection post-chirp to allow for some spontaneous (i.e., non stimulus-driven) activity.

For the Izhikevich neuron, current was injected via the variable  $I$  in equation 2.1. For all other models, current injections were simulated using the `IClamp` functionality provided in NEURON. `IClamp` simulates the effect of a current clamp, a technique which allows experimenters to control the electrical current passing through a neuronal membrane at a point. The `IClamp` processes were added to the center of the somatic compartment in each model.

## 2.3 Inverse Model

Explicitly, an inverse model for the current problem is a function  $F(X(t)) = \mathbf{Y}$  which takes as input the membrane potential  $X(t)$  and outputs a multi-valued floating point vector  $\mathbf{Y}$  corresponding to the predicted parameter values (e.g., ionic conductances, membrane resistances, etc.). We employ a Convolutional Neural Network (CNN) to compute this function.

CNNs have been successfully used in a variety of machine learning tasks, e.g., 2D image processing. Since a biological neuron’s activity (membrane potential) is typically recorded as a single valued time series, the CNN was constructed from several 1-dimensional (1D) CNN layers interleaved with 1D-Pooling layers to extract information about repetitive features (neuron spikes), followed by a block of Fully Connected (FC) layers intended to capture

global features of the trace. The last FC layer returned regressed parameters modelling a neuron. Those physical parameters were linearly scaled to the unitless range  $[-1,1]$ . The predictor  $F(\cdot)$  had the capacity to produce values with amplitudes up to 20% larger than necessary by design, to allow out-of-range predictions.

The natural amplitude of the membrane potential varied between  $[-80, 50]$  mV. To emphasize variability near the neuron’s resting potential, we shifted each trace so that its minimum value was +1mV and took the logarithm, as follows:

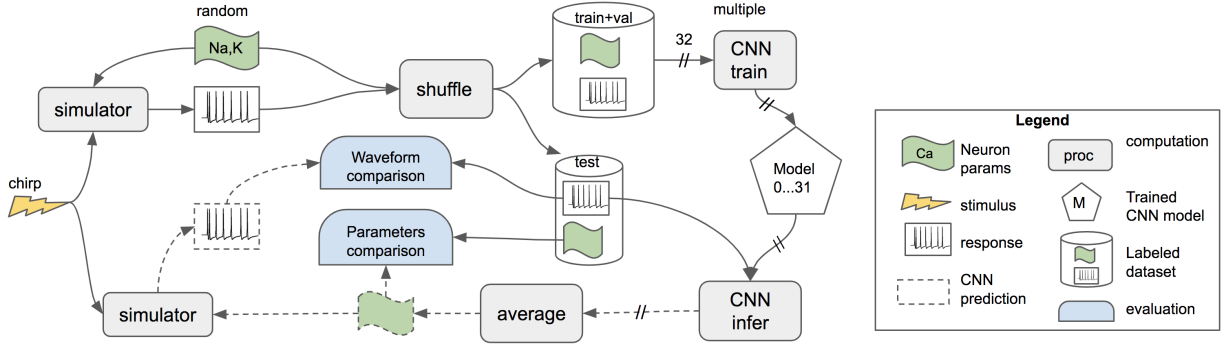
$$X'(t) = \log_{10} \left( 1 + (X(t) - \min_t X(t)) / \text{mV} \right) \quad (2.5)$$

This guarantees that the range of input values  $X'$  is limited to  $[0,3]$  and only a small fraction of this range is designated to spikes, which are of large amplitude. The dimension of  $X(t)$  was 9000 time bins.

The  $\{X(t), \mathbf{Y}\}$  dataset was randomly divided into train/validate/test sets in a 7/1/1 ratio. During CNN training, we reduce the learning rate at each step if there was no improvement in the validation loss, and terminate learning when no further improvement of loss is observed. We executed 32 independent trainings on the same train and validation datasets, randomly shuffling the examples each time. The final CNN-predictions were averaged over the 32 sets of learned weights  $\theta$ .

The optimal configuration of the layers of the CNN and the training method were determined by means of a random search of the hyper-parameter space [30]. We executed 780 CNN training experiments with 14 randomly varied hyper-parameters (hyper-parameters include number of CNN and FC layers, number of features per layer, optimizer type, loss function, etc.). Each training lasted between 4 and 6 GPU-hours, depending on how quickly the loss stabilized. The best model was not necessarily the one with the smallest one-time training loss value. By repeatedly training the same model using random permutations of the same data, we confirmed that training was stable, and that the final validation loss achieved by each training instance was minimally variable across permutations.

The best-performing CNN model we tested has 744,000 weights ( $\theta$ ) and consists of 6 CNN, 3 Pool, and 7 FC layers. The CNN layers had 18 or 25 features, the flatten layer had 45,000 units, the feature count for the FC layers varied between 148 and 46 (with deeper layers having fewer features), the last FC layer output had the dimension equal to the number of regressed parameters ( $\mathbf{Y}$ ). The CNN kernel size was 5, the pool size was 3, and the dropout rate between FC layers was 0.01. These hyper-parameters were determined via Adam Optimization [22] with a learning rate reduction factor of 0.14 and using mean absolute error (MAE) loss.



**Figure 2:** Workflow diagram describing generation of voltage traces, CNN training, prediction of ion channel densities from voltage traces input, and validation of CNN predictions.

## 2.4 Results

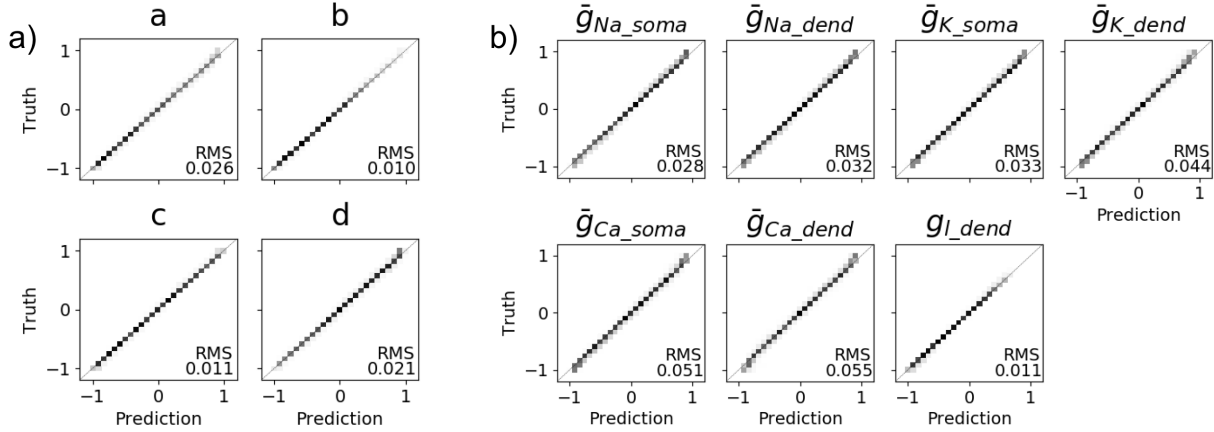
### 2.4.1 CNNs can accurately predict ionic conductances in different neuronal models

To train and validate our CNN for predicting ionic conductances from somatic membrane potentials, we recorded large datasets using the neuronal models described in section 2.2. Each observation in the dataset consisted of a randomly chosen set of microscopic parameters, as well as the membrane potential reported by NEURON during the time range demarcated by the red dotted lines in figure 1. Up to two million examples per neuronal model were generated, then split into training/validation/test sets. The training and validation examples were then used to train 32 independent CNN models, each of which had the same architecture but started with a different initialization of the weights and processed the training data in a different order. The 32 CNN predictions made for the 'test' dataset were averaged, resulting in a single prediction of the parameter values, drawn as a green banner with dashed outline in Fig. 2. The CNN-predicted parameters were compared using the ground truth values and the residual errors averaged over the sampled parameter space.

After optimizing the CNN hyperparameters as described in section 2.3, we are able to recover a majority of parameters for all of the models we trained on, with varying accuracy.

#### 2.4.1.1 Izhikevich : CNNs can regress parameters from a phenomenological point cell model

We started with the simplest model, the Izhikevich neuron, described in section 2.2.1.1. The predictions for the Izhikevich dataset are shown in Fig. 3a). The regression residues, defined as the difference between the truth and predicted value, have a RMS between 1-2%, averaged over the whole sampled dynamic range of respective parameters  $[-1,1]$ . Despite training on only  $3 \times 10^4$  examples, the CNN achieved high prediction accuracy. This can be attributed



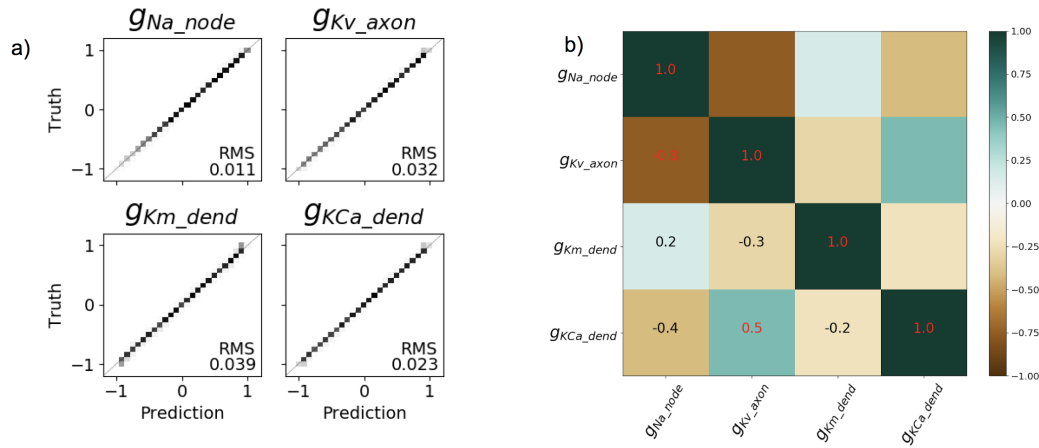
**Figure 3:** Correlation between neuron parameters used by the simulator (ground truth) and the values predicted by the CNN-model trained on the respective dataset. The average RMSE of the difference between the truth and predictions is shown in the lower right corner of each parameter plot. a) Izhikevich simulator with 4 varied parameters, b) Hodgkin-Huxley simulator with 7 varied parameters.

to the relative simplicity of the Izhikevich model, where the parameters contribute directly to the voltage response.

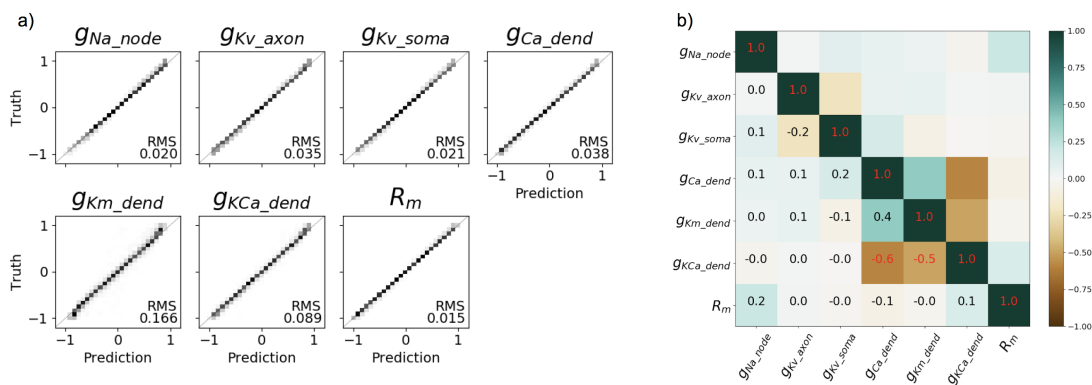
#### 2.4.1.2 Hodgkin-Huxley : CNNs can regress conductances in remote dendritic compartments

Next, we simulated the Hodgkin-Huxley ball and stick model, which is not a point neuron but has a simplified neuronal morphology and 7 parameters. This model was used to study whether a CNN could be trained to separate contributions from the same ion channels in two different compartments. We trained the CNN with  $4.5 \times 10^5$  examples. The CNN predictions of the parameter values are shown in Fig. 3b. The CNN trained for the Hodgkin-Huxley 7-parameter model made accurate predictions with small residues below 6%. The highest prediction accuracy (residue RMS of 1.1%) was achieved for the leak conductance at the dendrite  $g_{I\_dend}$ . This is the only conductance in the Hodgkin-Huxley 7 parameter model that was not allowed to take different values at the soma and the dendrite. The most difficult parameters to predict were the correlated somatic and dendritic calcium conductances (residue RMS of 5.1% and 5.5%, respectively). We attribute this to the challenge of distinguishing the contribution of the calcium conductance in dendritic compartments to the voltage response recorded at the soma.

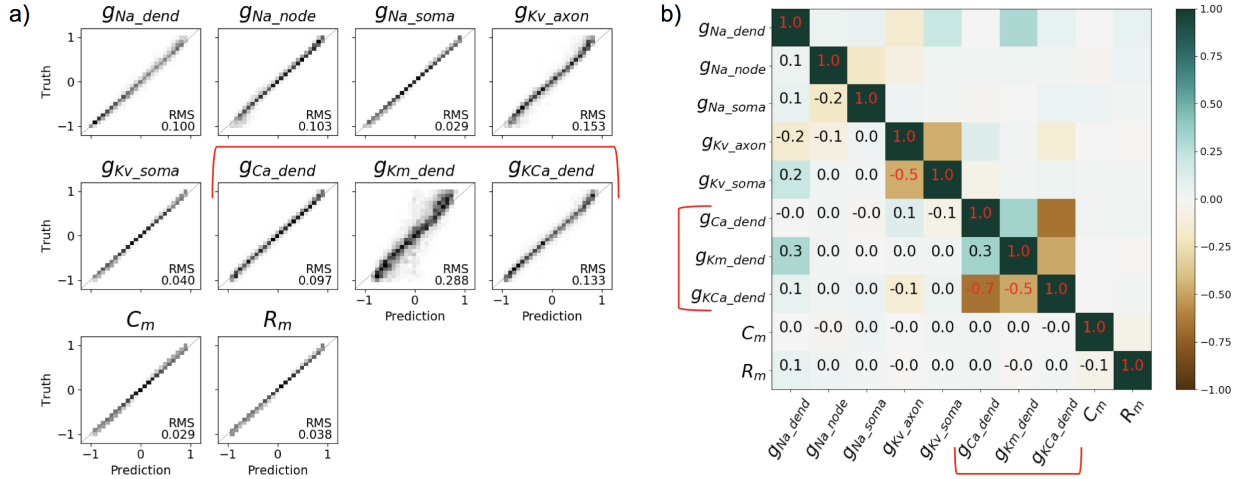




**Figure 4:** CNN prediction of Mainen 4-parameter neuron model. (a) Correlation between the true parameters and predictions. (b) correlation between residuals of predictions depicted by color scale, where 0 means no correlation and (-)1 means strong (anti)correlation.



**Figure 5:** CNN prediction of Mainen 7-parameter neuron model.



**Figure 6:** CNN predictions for 10 parameter Mainen dataset. The red brackets mark the most difficult parameters to predict, a highly correlated triplet of dendrite conductances describing Kv, Km, and KCa channels.

#### 2.4.1.3 Mainen and Sejnowski: CNNs can regress conductances in a biophysical model

Next, we studied how the trained CNN would predict the ionic conductances of a biophysically detailed model with realistic morphology such as Mainen and Sejnowski [23]. We tested this model with gradually increasing complexity by varying 4, 7, and 10 parameters while holding the rest constant (Table 2.4, Figs. 4, 5, and 6, respectively). When 4 parameters were varied, prediction errors were less than 4%. Increases in the average prediction error were seen as more parameters were varied (Table 2.4). The 10 parameter model introduced the new challenge of multiple variables controlling the same ionic conductance *in the same compartment*. For example, the total potassium conductance can be controlled by any of the following channels (parenthetical symbols correspond to parameter names in table 2.4): voltage-gated potassium channels (Kv), potassium M-channels (Km), or calcium-activated potassium channels (KCa). This redundancy results in a fundamental lack of parameter identifiability. Indeed, these three parameters proved to be the most strenuous to resolve: for the 10 parameter Mainen model, the CNN was able to determine them with residues of 15%, 29%, and 13%, respectively. However, since the Mainen model is sufficiently complex to represent many of the underlying mechanisms of neuronal firing, this is an example of how a CNN can be a useful tool to predict ion channel densities of neurons from *in vitro* voltage responses.

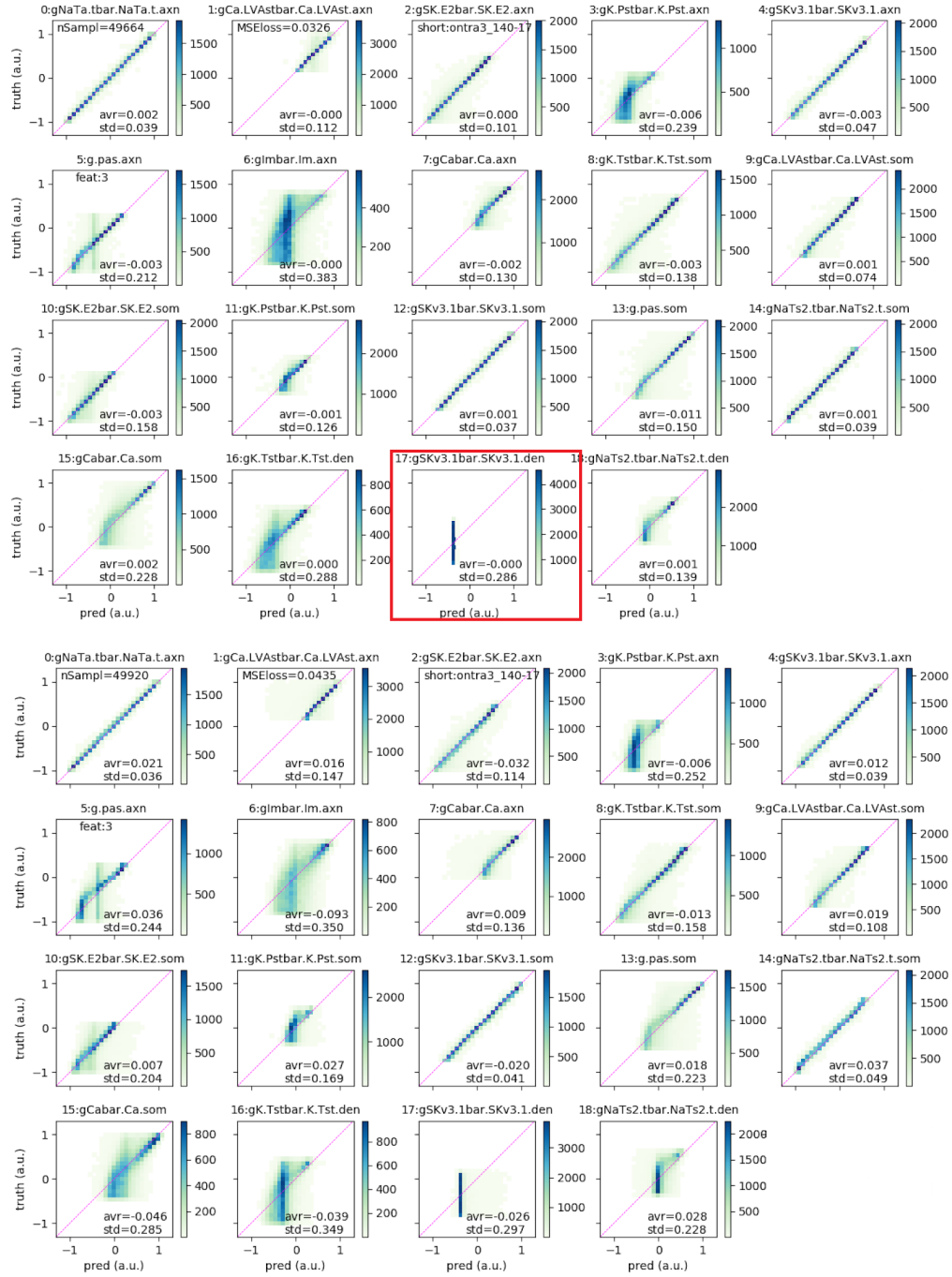
#### 2.4.1.4 BBP: CNNs can regress most conductances in an advanced biophysical model

The final targets for our method were the biophysically detailed neurons from the Blue Brain Project model. Here we sought to evaluate our CNN on a variety of cells of different types. We started by training the CNN on only one BBP cell at a time, as we did for the previous models. We chose two double bouquet cells (DBC), one from cortical layer IV and one from layer VI. The results are shown in the top halves of figures 7 and 8 respectively. Because the BBP cells simulate many more ion channels than the Mainen and Sejnowski model, we encountered the same difficulty predicting correlated parameters, such as conductances that affect the same ionic species (indeed, many of the most difficult conductances to predict are Potassium currents, just like we saw with the Mainen model), but here the effect is stronger, resulting in the CNN failing to predict some parameters entirely, such as the ones bracketed in red, where the CNN made nearly the same predictions for almost all examples. Subsequent analysis revealed that those parameters the CNN was unable to resolve have minimal effect on the membrane potential waveform under the conditions we studied.

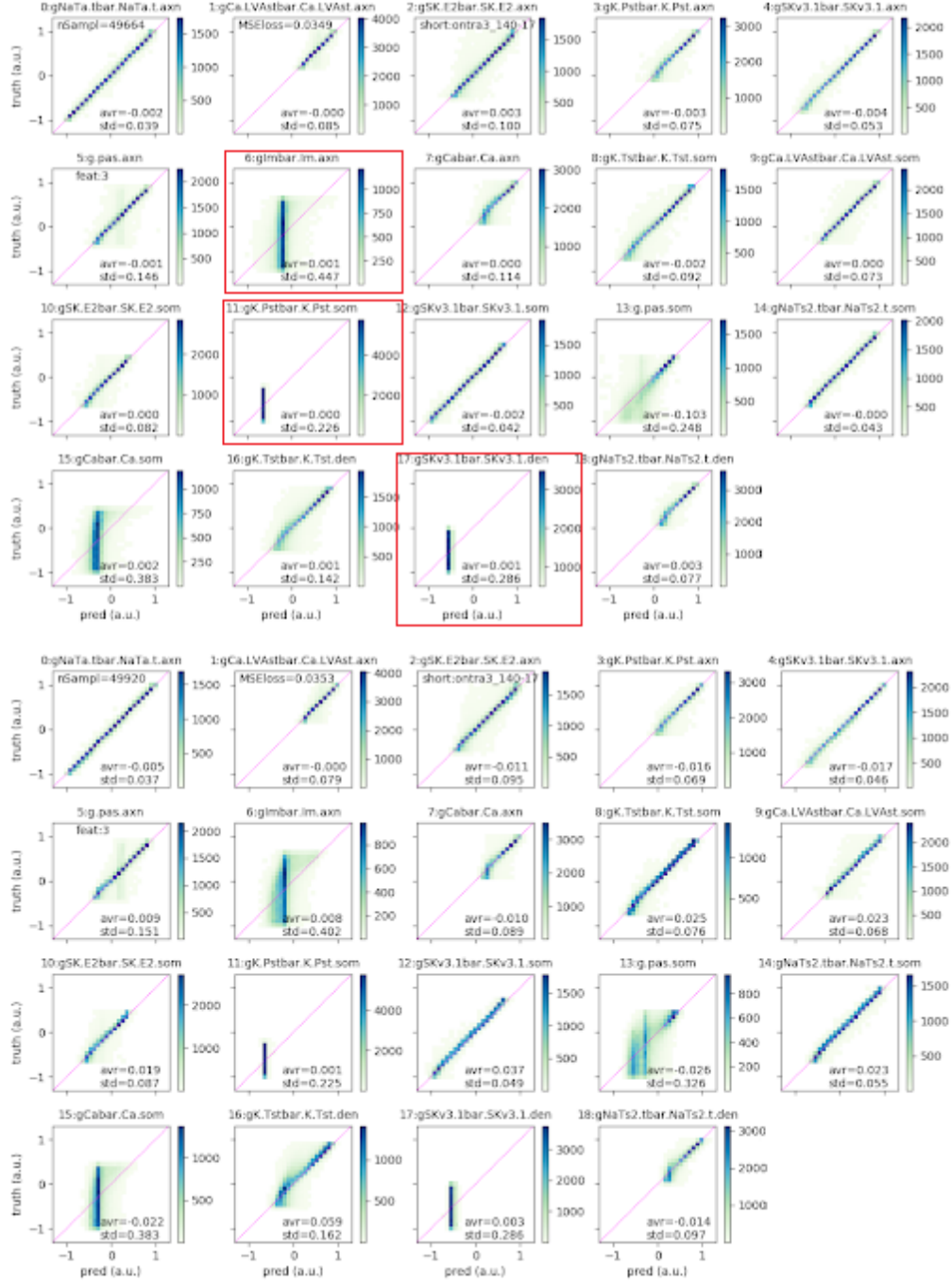
#### 2.4.1.5 BBP: CNNs show promise to regress conductances across cells

Finally, to evaluate whether CNNs could regress conductances across cells, we trained one CNN using data from 8 different inhibitory cells of different celltypes. The two cells shown in figures 7 and 8 were not used in the training of this CNN, but their data were used for testing. This simulates one way of applying our technique in an experimental setting where we would likely need to train our CNN on a variety of cells, then use it to make predictions on novel cells (see section 4.2.1). When we applied this strategy to our *in silico* data, we found that the CNN trained on 8 inhibitory cells was able to make many good predictions of ionic conductances in cells that it had never seen before. The results are shown in the bottom halves of figures 7 and 8.

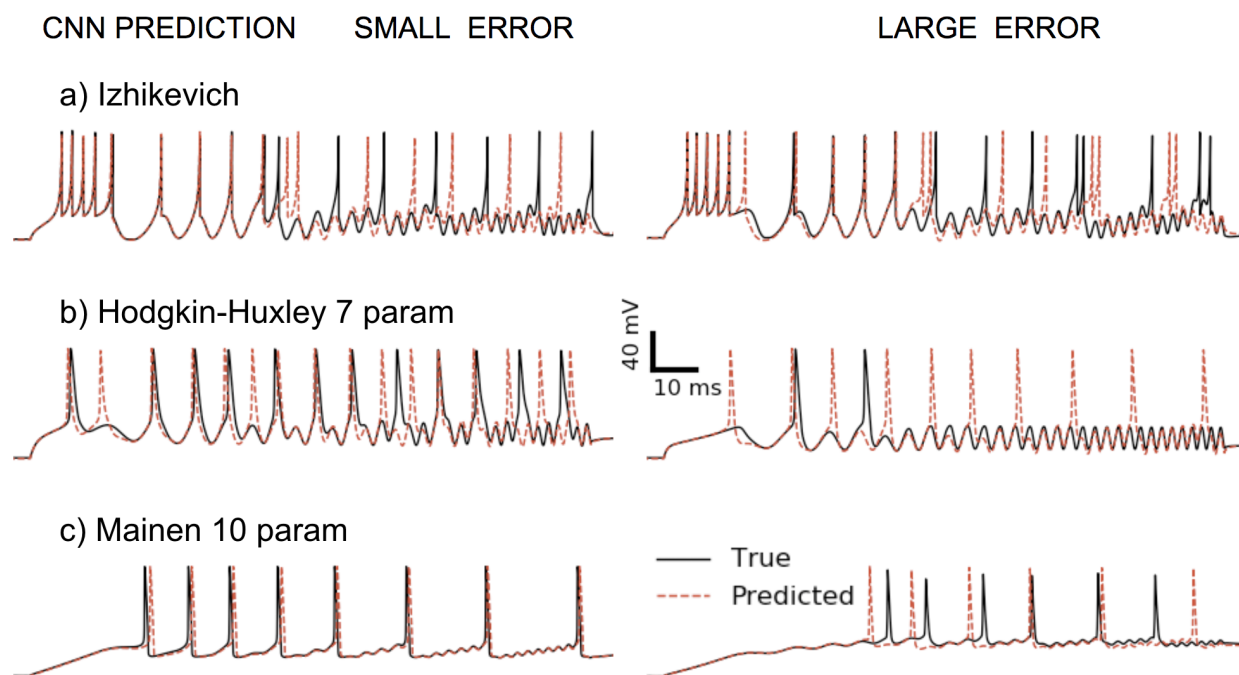
Not surprisingly, the accuracy of CNN predictions generally decreased when going from a CNN trained on data from the same cell as the test data to a CNN trained on 8 cells distinct from the testing cell. For example, the layer IV cell shown in figure 7 experienced in a change in MSE of loss across testing samples of  $0.0326 \rightarrow 0.0424$ , an increase of 30%. Interestingly, however, some cells saw very little reduction in accuracy of CNN predictions even when switching from a CNN trained on that cell's data to one trained on 8 other cells' data. For example, the layer VI cell shown in figure 8 experienced a change in MSE loss of  $0.0344 \rightarrow 0.0348$ , an increase of only 1.2%. This suggests that 1.) there indeed exists an inverse model for the current problem which is general enough to solve it across a variety of cell types, and 2.) our method is likely to be able to learn this inverse model using data from a reasonably small number of cells, provided they are of sufficient diversity.



**Figure 7:** CNN predictions for L4\_DBC cell from Blue Brain. **Top:** CNN trained only on L4\_DBC data. Red brackets indicate poorly-resolved parameters. MSE of loss across testing examples is 0.0326. **Bottom:** CNN trained on data from 8 inhibitory cells. MSE of loss across testing examples is 0.0424



**Figure 8:** CNN predictions for L6\_DBC cell from Blue Brain. **Top:** CNN trained only on L6\_DBC data. Red brackets indicate poorly-resolved parameters. MSE of loss across testing examples is 0.0344. **Bottom:** CNN trained on data from 8 inhibitory cells. MSE of loss across testing examples is 0.0348.



**Figure 9:** Selected examples of membrane potential traces generated from CNN predictions (red, dotted), overlaid with the true membrane potential (black, solid). "Small error" means predictions were within 0.5 RMSE of the true parameters on average, "large error" means the average error across parameters was  $>0.5$  RMSE. See section 2.4.2.

## 2.4.2 Accurate CNN predictions qualitatively reproduce input waveform

Although our primary goal was to accurately regress the numerical values of microscopic parameters describing a neuron, we were interested to see how well the regressed parameters reproduced the spiking activity of the simulated neuron, in order to gain some qualitative insight into the accuracy of our CNN's predictions in terms of actual neuronal behavior. To do this, we divided the testing examples into two classes: those that resulted in well-predicted parameters ("small error", where the predicted parameters were less than an average of 0.5 RMSE from the true value) and those that resulted in poorly-predicted parameters ("large error", where the predicted parameters were more than an average of 0.5 RMSE from the true value). We then re-ran the simulations using the CNN-predicted parameters and compared the resulting membrane potentials with the test examples themselves (ground truth). This comparison is shown for a few exemplars in figure 9. On the left are the "small error" examples, while on the right are the "large error" examples. In general, we observe more similarity between the waveforms from predicted and ground truth parameters for the "small

error” cases than for the ”large error” cases. This shows that, in many cases, our CNN is able to make predictions of microscopic parameters that capture most of the important variability required to reproduce the neuron’s response to a stimulus.

## 2.5 Discussion

### 2.5.1 CNNs are able to regress ionic conductances from somatic potentials

In this study we explored the applicability of CNNs to predict ionic channels’ conductances in neuronal models from their somatic membrane potential in response to a standardized stimulus. We trained the same CNN on datasets from 3 different types of neuronal models: Izhikevich spiking models, Hodgkin-Huxley reduced morphology models, and biophysically detailed models (Mainen and Sejnowski, as well as the Blue Brain Project cells). The CNN with optimized hyperparameters was able to recover a majority of generative parameters for all of the 8 models we trained it on with varying accuracy. The most challenging parameters to predict were those that were highly correlated, either due to having the same ion channels in several compartments or multiple types of ion channels gating the same ionic species. This issue was more prevalent as the complexity of the models increased.

In previous studies of individual cells [11, 12] and small circuits [27], neuronal behavior was found to be robust to certain variations in low-level parameters such as ionic conductances. However, these studies either examined only certain properties of the neuronal response (e.g. firing rate in response to a current injection [11] or qualitative activity pattern - silent vs. tonically firing vs. bursting [12]), or output behavior of networks [27]. Here, our CNN operates on the full, raw voltage response of the cell. Our results suggest that when disparate sets of parameters produce similar neuronal firing, there are actually minute differences in the membrane potential which can be used to constrain the parameter values more strongly than the overall behavior (firing rate or qualitative activity pattern).

### 2.5.2 Inferring ionic conductances in biophysically detailed models requires significant computational effort

In order to obtain accurate results on the BBP cells, we found that we needed on the order of 500,000 samples from each cell. Due to the computational complexity of simulating activity in such detailed neuronal models, this was a highly nontrivial task. To simulate 500,000 samples of 500ms of neural activity for each of 200 cells consisting of hundreds or thousands of segments per cell, we used Cori, a Cray XC70 supercomputer installed at the National Energy Research Scientific Computing Center at Lawrence Berkeley National Lab (Berkeley, CA). Our jobs used up to 4096 nodes of the system, with 128 threads running on each node. Each thread ran a separate instance of NEURON (scripted via Python) in a Docker-like container, which achieved total isolation of the user spaces of each process, while still

allowing them to write to disk in parallel via the Lustre filesystem [28] installed as Cori’s scratch filesystem. The complete dataset required hundreds of thousands of node-hours to produce, and occupies more than 100 TB on disk.

Of course, obtaining such samples experimentally would be completely infeasible - not only would the volume of data take prohibitively long to obtain, a neuron *in vitro* is not likely to withstand so many repeated injections of large currents without breaking down. Additionally, our method requires fine-grained control over the conductances of ion channels in order to generate training data. This can only be done in a simulation, as ion channels are not readily adjustable in a physical neuron.

## 2.6 Acknowledgements

The work presented in this chapter was done in collaboration with Roy Ben-Shalom (University of California, San Francisco) and Jan Balewski (National Energy Research Scientific Computing Center at Lawrence Berkeley National Lab, Berkeley, CA) [7].

## 2.7 Tables of parameter values and CNN prediction accuracy

**Table 2.1:** Ranges of parameters used by Izhikevich model and accuracy of CNN predictions, averaged over 32 trained models.

Name	Base	Range		Units	CNN accuracy
a	0.01	0.01	0.10	<i>ms</i>	2.6%
b	0.2	0.1	0.4	(1)	1.0%
c	-65	-80	-50	<i>mV</i>	1.1%
d	2	0.5	5	<i>mV</i>	2.1%
CNN-training		end-loss MSE			5.4e-4
		num. examples			29k
		GPU time (min)			7.2
		epochs			38



**Table 2.2:** Ranges of 4 or 7 parameter sets used by Hodgkins-Huxley ball+stick model and respective CNN prediction accuracies. The value of  $g_{l\_soma}$  was locked to  $g_{l\_dend}$ .

Paramater	Base	Range		Units	CNN accuracy		
					7par	4parH	4parE
$\bar{g}_{Na\_soma}$	500	250	1000	$pS/\mu m^2$	2.8%	1.3%	1.7%
$\bar{g}_{Na\_dend}$	500	250	1000	$pS/\mu m^2$	3.2%	1.7%	-
$\bar{g}_{K\_soma}$	10	5	20	$pS/\mu m^2$	3.3%	1.8%	-
$\bar{g}_{K\_dend}$	10	5	20	$pS/\mu m^2$	4.4%	2.0%	4.6%
$\bar{g}_{Ca\_soma}$	1.5	0.75	3	$pS/\mu m^2$	5.1%	-	3.6%
$\bar{g}_{Ca\_dend}$	1.5	0.75	3	$pS/\mu m^2$	5.5%	-	-
$g_{l\_dend}$	0.6	0.3	0.9	$mS/cm^2$	1.1%	-	0.7%
CNN-training		end-loss MSE			2.8e-3	5.2e-4	1.4e-3
		num. examples			449k	116k	88k
		GPU time (min)			150	33	26
		epochs			48	42	46

**Table 2.3:** Ranges of 10 parameters used by Hodgkin-Huxley two dendrites model and respective CNN prediction accuracies. The value of  $g_{l\_apic}$  and  $g_{l\_basal}$  were locked to  $g_{l\_soma}$ .

Parameter	Range		Units	CNN accuracy
$\bar{g}_{Na\_soma}$	250	1000	$pS/\mu m^2$	7.9%
$\bar{g}_{Na\_apic}$	250	1000	$pS/\mu m^2$	3.1%
$\bar{g}_{Na\_basal}$	250	1000	$pS/\mu m^2$	17%
$\bar{g}_{K\_soma}$	5	20	$pS/\mu m^2$	17%
$\bar{g}_{K\_apic}$	5	20	$pS/\mu m^2$	6.9%
$\bar{g}_{K\_basal}$	5	20	$pS/\mu m^2$	39%
$\bar{g}_{Ca\_soma}$	0.75	3	$pS/\mu m^2$	23%
$\bar{g}_{Ca\_apic}$	0.75	3	$pS/\mu m^2$	7.1%
$\bar{g}_{Ca\_basal}$	0.75	3	$pS/\mu m^2$	50%
$g_{l\_soma}$	0.25	1	$mS/cm^2$	1.9%
CNN-training		end-loss MSE		0.059
		num. examples		1.4M
		GPU time (min)		620
		epochs		72

**Table 2.4:** Ranges of 4,7 or 10 parameter sets used by Mainen neuron simulator and respective CNN prediction accuracies.

Paramater	Base	Range	Units	CNN accuracy		
				10par	7par	4 par
$gNa_{dend}$	20	10	250 $S/cm^2$	10%	-	-
$gNa_{node}$	30	15	80 $kS/cm^2$	10%	2.0%	1.1%
$gNa_{soma}$	0.02	0.01	60 $kS/cm^2$	2.9%	-	-
$gKv_{axon}$	2	1	5 $kS/cm^2$	15%	3.5%	3.2%
$gKv_{soma}$	0.2	0.1	2 $kS/cm^2$	4.0%	2.1%	-
$gCa_{dend}$	300	150	600 $mS/cm^2$	9.7%	3.8%	
$gKm_{dend}$	100	50	200 $mS/cm^2$	29%	17%	3.9%
$gKCa_{dend}$	3	1.5	6 $S/cm^2$	13%	8.9%	2.3%
$C_m$	0.6	0.3	1.5 $\mu F/cm^2$	2.9%	-	-
$R_m$	30	15	60 $k\Omega cm^2$	3.8%	1.5%	-
CNN-training			end-loss MSE	0.023	6.6e-3	9e-4
			num. examples	780k	400k	77k
			GPU time (min)	310	110	20
			epochs	51	48	38

## Chapter 3

# Reconstructing Neuronal Sources of Cortical Surface Electrical Potentials

### 3.1 Background and Motivation

Electrocorticography (ECoG; Figure 10) is a neural recording technique where signals are obtained from electrodes placed directly on the cortical surface. ECoG provides broad spatial coverage with an intermediate ("mesoscale") level of detail between scalp-based electroencephalography (EEG) and intracortical recordings. Figure 10a shows a photomicrograph of a  $\mu$ ECoG grid implanted in a rat.

Of particular interest to neuroscientists studying mesoscale phenomena is the distribution of neural activity through the depth of cortex (the axis perpendicular to the plane containing the electrodes), which is anatomically divided into several layers consisting of distinct cell types whose different purposes are as yet largely unknown. Understanding the laminar distribution of neural activity contributing to cortical surface electrical potentials (CSEPs) recorded via ECoG would enable ECoG studies to probe the different functions of each cortical layer. In this chapter, we show how biophysically detailed modeling helps us understand the makeup of auditory-evoked CSEPs in terms of contributions from the various cortical layers.

### 3.2 Forward Model

#### 3.2.1 Cortical Column Simulation

The signals we wish to study, cortical surface electrical potentials recorded via ECoG, are produced by large networks of interconnected neurons. Thus, we require a detailed model of a large population of neurons representing at least one column of sensory cortex in the rat. The Blue Brain Project (BBP; section 1.4.2) provides such a model, implemented in NEURON. To gain insight into the laminar contributions to evoked ECoG signals, we reproduced this

model using publicly released data, including the spatial location and connectivity matrix of all cells in the column, as well as tuned and experimentally validated models of individual neurons from all cortical laminae including their electrical characteristics (ion channel models and associated parameters such as spatially varying ionic conductances, membrane resistance/capacitance, etc.) and their detailed morphologies based on full reconstructions of neurons observed experimentally (or algorithmically generated clones thereof), reflecting the full diversity of neurons currently known to be found in rat somatosensory cortex. The publicly available BBP dataset provides 5 reconstructed (or cloned and slightly tweaked) morphologies for each of 207 distinct cell types. Each neuron in our model is represented by one of the 5 morphologies for that neuron's cell type, chosen at random and rotated by a random angle about a line passing through the cell's soma parallel to the longitudinal axis of the column. With this model, we were able to qualitatively recreate the evoked spectrum observed in the experimental data.

Neurons in our simulated column are innervated by synapses from three populations: 1.)  $N_{thal} = 5000$  excitatory thalamic neurons conveying feed-forward sensory input (thalamo-cortical connections) modeled as rate-modulated Poisson spike trains, 2.)  $N_{bkg,e} = 25000$  excitatory and  $N_{bkg,i} = 25000$  inhibitory background cortical neurons from other columns (external cortico-cortical connections) also modeled as Poisson spike trains, and 3.) other neurons in the simulated column (internal cortico-cortical connections). The rate constant of the Poisson processes generating thalamocortical spike trains increases from a baseline rate of  $\nu_{thal}^{base} = 1$  hz to a stimulus-induced rate of  $\nu_{thal}^{stim} = 35$  hz for 50 ms out of every 1000ms (onset and offset are cosine ramps 5 ms in duration), reflecting the temporal structure of tone pips in our experimental preparation, while the rate constant of the external cortico-cortical spike trains remains constant at  $\nu_{bkg} = 7$  hz for the duration of the simulation. Thalamic synapses are distributed within the column in a depth-dependent manner, with peaks at 670  $\mu\text{m}$  and 1300  $\mu\text{m}$  below the cortical surface (Fig. 11b). Synapses from background neurons are formed on neuronal segments in the simulated column with probability proportional to each segment's surface area.

Synapses from all populations produce membrane currents  $i_{syn}(t)$  according to the equation:

$$i_{syn}(t) = G(V_m - E_{rev})(e^{\tau/\tau_2} - e^{\tau/\tau_1}) \quad (3.1)$$

Where  $t$  is the time since the synapse was activated,  $V_m$  is the membrane potential,  $E_{rev}$  is the reversal potential of the synapse,  $G$  is the weight (max conductance) of the synapse, which is randomly drawn from a lognormal distribution with different center and spread for each input source, and  $\tau_1$  and  $\tau_2$  are the time constants of the exponential activation/deactivation of the synapse. The values of these parameters for different types of synapses is given in the table below:

Synapse type	$\tau_1$ (ms)	$\tau_2$ (ms)	$E_{rev}$ (mV)
AMPA (e $\rightarrow$ e)	1.0	3.0	0
AMPA (e $\rightarrow$ i)	0.1	0.5	0
GABA (i $\rightarrow$ e)	2.7	15.0	-70
GABA (i $\rightarrow$ i)	0.2	8.0	-70

We applied a modest amount of hand-tuning of these parameters to achieve reasonable baseline firing rate (3-10 Hz) during time periods when the thalamocortical spike trains fire at  $\nu_{thal}^{base}$ , and to reproduce the experimentally observed sharp transient stimulus-evoked response (after the transition to  $\nu_{thal}^{stim}$ ) within the simulated column.

### 3.2.2 Computation of CSEP from simulated neuronal activity

In order to compute the electrical potential produced by the currents generated by the simulated column, we consider the extracellular space to be electrically homogenous and isotropic, permitting a simple expression for the electrical potential produced by each neuronal segment in the simulation in terms of the current passing through the membrane of that segment. These can then be summed over all segments to obtain the total extracellular potential due to all cells in the simulation. The resulting formula is known as the Line Source Approximation (LSA):

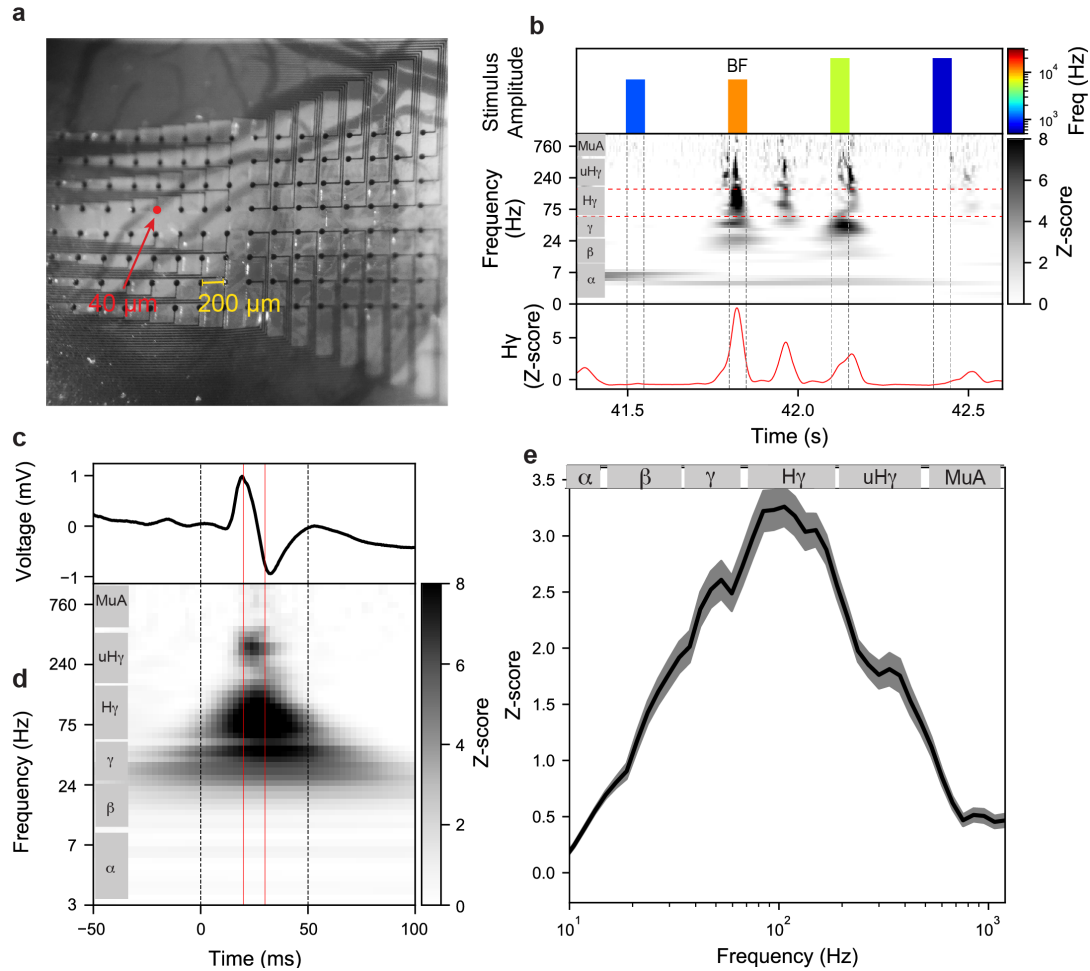
$$V(\mathbf{r}, t) = \sum_i \frac{i_i(t)}{4\pi\sigma} \int \frac{d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} \quad (3.2)$$

Where  $V(\mathbf{r}, t)$  is the extracellular potential,  $i_i(t)$  the current going through neuronal segment  $i$  at time  $t$ ,  $\sigma = 0.3$  S/m is the conductivity of the extracellular medium, and the variable of integration  $\mathbf{r}'$  runs from one end of segment  $i$  to the other. The sum in equation 3.2 runs over all segments  $i$  in the simulated column. A derivation of equation 3.2 can be found in appendix A. To account for the nonzero spatial extent of the  $\mu$ ECoG electrode, we compute  $V(\mathbf{r}, t)$  at 100 randomly and uniformly sampled points within the 20  $\mu$ m radius of the  $\mu$ ECoG electrode and average them.

## 3.3 Results

### 3.3.1 Stimulus-evoked cortical surface electrical potentials exhibit large peaks in the high-gamma range

We began our study of ECoG signals by obtaining experimental recordings of stimulus-evoked surface potentials from auditory cortex in a rat. Figure 10a shows a photomicrograph of a  $\mu$ ECoG grid implanted in the animal. To collect *in vivo* CSEP recordings, we played auditory tone pips with varying frequency and intensity (amplitude) at 5 recording locations in 4 rats, obtaining signals such as the one depicted in Figure 10b, which depicts the average signal recorded over several repetitions of the frequency that elicited the largest response



**Figure 10: Stimulus-evoked cortical surface electrical potentials exhibit large peaks in the high-gamma range.**

**a.** Photomicrograph of an 8x16  $\mu$ ECoG grid (pitch: 200  $\mu$ m, contact diameter: 40  $\mu$ m) on the surface of rat primary auditory cortex (A1).

**b.** Top: tone stimulus played during experimental recordings. Middle: z-scored spectral decomposition of single-trial evoked cortical surface electrical potentials from a single electrode. Bottom: High-gamma component of single-trial evoked cortical surface electrical potentials indicated by horizontal dashed lines in the middle panel.

**c.** Trial-averaged evoked cortical surface electrical potential on one  $\mu$ ECoG electrode in response to presentations of that electrode's best tuned frequency.

**d.** Trial-averaged neural spectrogram for the electrode shown in **c** in response to presentations of its best tuned frequency. Dashed vertical lines in **c** & **d** represent stimulus onset and offset. Red vertical lines in **c** & **d** correspond to the time window of extracted evoked response used for subsequent analysis.

**e.** Grand-average (mean  $\pm$  s.e.) z-scored amplitude as a function of frequency across all tuned electrodes ( $N = 333$ ).

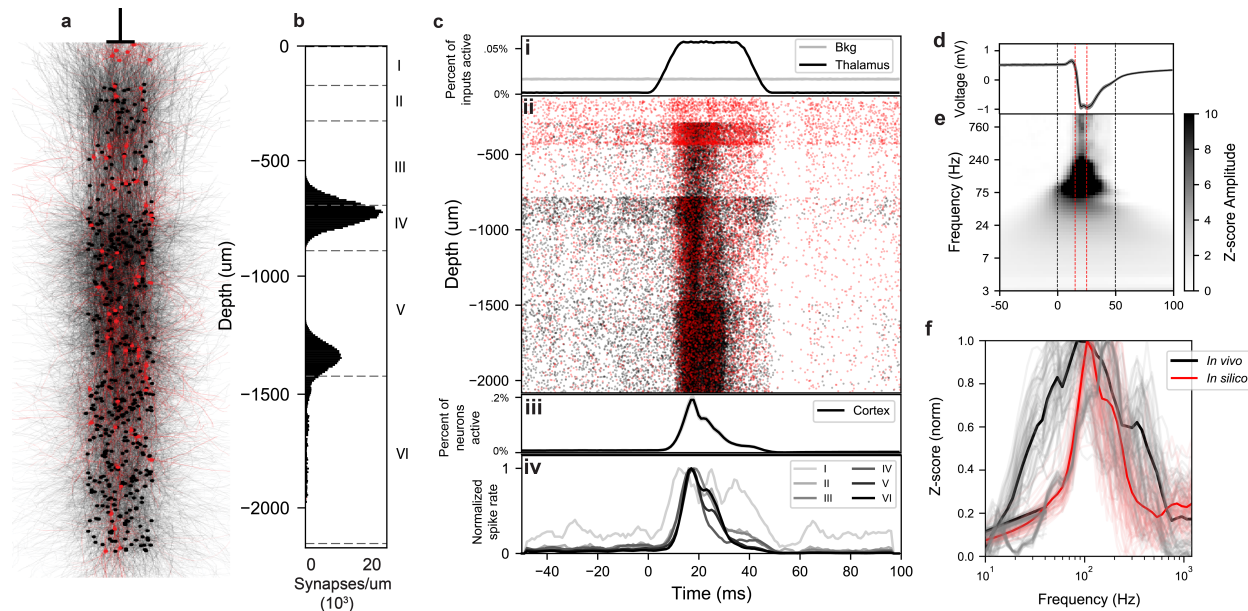
on a selected electrode. To understand the spectral structure of the neural response, we computed spectrograms of the CSEP, such as the one shown in Figure 10c. To summarize the frequency content of evoked CSEPs, we extracted mean z-scored responses across all neural frequency components in a  $\pm 5$ ms window around the time of the peak high-gamma response (we define high-gamma as the 70-170 Hz component of the signal, indicated by red vertical lines in Fig 10b-c). We included all electrodes with a tuned response in the high-gamma band ( $N = 333$  electrodes from 5  $\mu$ ECoG placements on auditory cortex in 4 rats). Figure 10d plots the averaged ( $N = 333$  electrodes, mean  $\pm$  s.e.) z-scored response as a function of frequency. On average, we found that evoked responses were unimodally peaked around the H $\gamma$ -band, with notable responses in the multi-unit activity range (MuA,  $>500$ Hz).

### 3.3.2 Biophysical *in silico* cortical column reproduces *in vivo* observed ECoG response

The column model and  $\mu$ ECoG electrode are depicted in Figure 11a, where circles indicate the locations of (a subset of) neuronal soma (black: excitatory neurons; red: inhibitory neurons). Stimulus-evoked input to the column is provided by activating (with Poisson spike trains) thalamocortical synapses located throughout the column according to the distribution shown in Figure 11b, which also displays the cortical layers.

An example of the column's activity is displayed in Figure 11c. The biophysical neurons in the column received thalamic input in the form of Poisson spike trains that were modulated in time to emulate our tone stimulus (Fig.11c.i, black) and background Poisson spike trains (Fig.11c.i, grey) that were not modulated by the stimulus. In Figure 11c.ii we show the spike times (black: excitatory neurons; red: inhibitory neurons) in response to one presentation of the input stimulus (Fig. 11c.i). Neurons are arranged by depth below the surface, which allows us to visualize the laminar boundaries as sharp changes in the density of firing reflecting different cell densities across layers. The fraction of neurons in the column firing action potentials is displayed in Figure 11c.iii as a function of time. The time-to-peak of about 15-20 ms in most layers (Fig. 11c.iv), as well as the following period of slightly elevated activity until stimulus offset, are both consistent with the *in vivo* recordings.

The biophysical model produces CSEPs (Fig. 11d-f) consistent with the high-frequency transient onset response observed *in vivo*. We computed the electrical potential at the cortical surface of the simulated column using the Line Source Approximation, and processed the simulated data identically to the experimental data. Average (mean  $\pm$  s.e.,  $N = 60$  stimulus presentations) raw evoked cortical surface electrical potential from the model is plotted in Figure 11d, while the spectral content of the simulated CSEP is shown in Figure 11e (dashed black lines: stimulus; dashed red line: 10ms window around the peak high-gamma response). The extracted z-scored (max normalized) response as a function of frequency for the simulated CSEP is shown Figure 11f, as well as the experimental data (black). The high-frequency content of CSEPs collected experimentally and CSEPs generated by the



**Figure 11: Biophysical *in silico* cortical column reproduces *in vivo* observed  $\mu$ ECoG response**

**a.** Rendering of a random sub-selection of 626 neurons in the simulated column (about 2% of the total). Black: excitatory neurons; red: inhibitory. Circles represent somas, lines represent dendritic structures. The position of the simulated  $\mu$ ECoG electrode relative to the column is shown above.

**b.** Distribution of synapses from thalamus along the depth of the simulated cortical column.

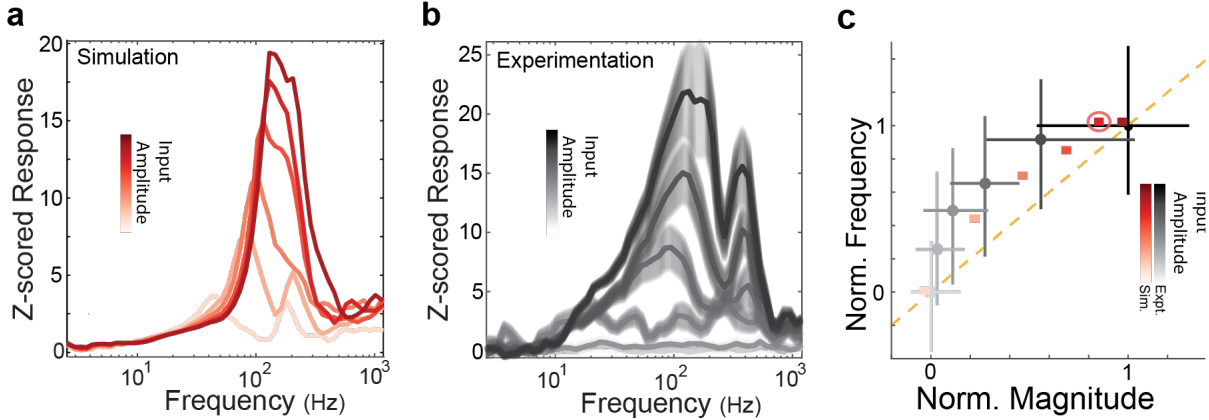
**c.** Data from one simulated stimulation and pre/post-stimulus silence **i.** Population spike rate of thalamic and background cortical spike trains activating synapses in the column. **ii.** Spike raster of all neurons in the column vs. soma depth (y-axis). Note that differences in raster density in part reflect differences in neuron density across cortical layers. **iii.** Population spiking (fraction of neurons spiking in 1 ms) of biophysically detailed cortical neurons. **iv.** Cell-averaged spike rate of biophysically detailed neurons in each layer. Darker shades indicate deeper layers.

**d.** CSEP computed by the Line Source Approximation from all neurons in the column during a 150 ms window centered around the 50 ms “tone pip” stimulation.

**e.** Spectrogram of the CSEP in panel **d**, z-scored to baseline

**f.** Frequency content of CSEP during 10 ms centered at the response peak (indicated with dotted red lines in panels **d** and **e**), z-scored to baseline. Individual electrode averages from experimental results are in grey, black is grand average. Individual stimulus presentations from simulations are in pink, red is grand average. All traces are normalized to their respective maxima.





**Figure 12:** *In silico* cortical column predicts experimentally observed relationship between response magnitude and frequency

- a. Average z-score as a function of frequency in six simulations with variable input amplitude.
- b. Average z-score as a function of frequency in the experimental data for six different stimulus amplitudes.
- c. Normalized response magnitude vs. normalized response frequency for experimental data (black, mean  $\pm$  s.d.) and for simulations (red). Each data point corresponds to the response frequency and magnitude associated with a distinct input magnitude (response magnitude increases monotonically with input magnitude). Circled point indicates the input magnitude used in Figures 11,13,14. Orange dashed line is unity.

simulations are qualitatively similar. Both experimental and simulated CSEPs exhibit a peak frequency of  $\sim 100$  Hz, and the spread of the signal around the peaks are overlapping. Thus, biophysical simulations accurately recreate key aspects of experimentally acquired data, indicating that they are a good forward model of CSEP generation.

### 3.3.3 *In silico* cortical column predicts experimentally observed relationship between response magnitude and frequency

The model makes testable predictions regarding the relationship between the magnitude and frequency content of CSEP responses. In the simulations, we varied the magnitude of the excitatory input to the network by increasing the mean firing rate of the thalamic spike trains during the stimulus. Analogously, in the experiments, we monitored the evoked CSEP in response to varying sound amplitudes at each electrode's best frequency.

For the simulations, Figure 12a displays the average normalized evoked response to stimulation of different amplitudes as functions of frequency (input magnitude given by color saturation, indicated by inset color bar). We observed that the magnitude of CSEP response depended monotonically on input magnitude. More interestingly, we found that as the magnitude of the response increased, so did the frequency content of that response. This

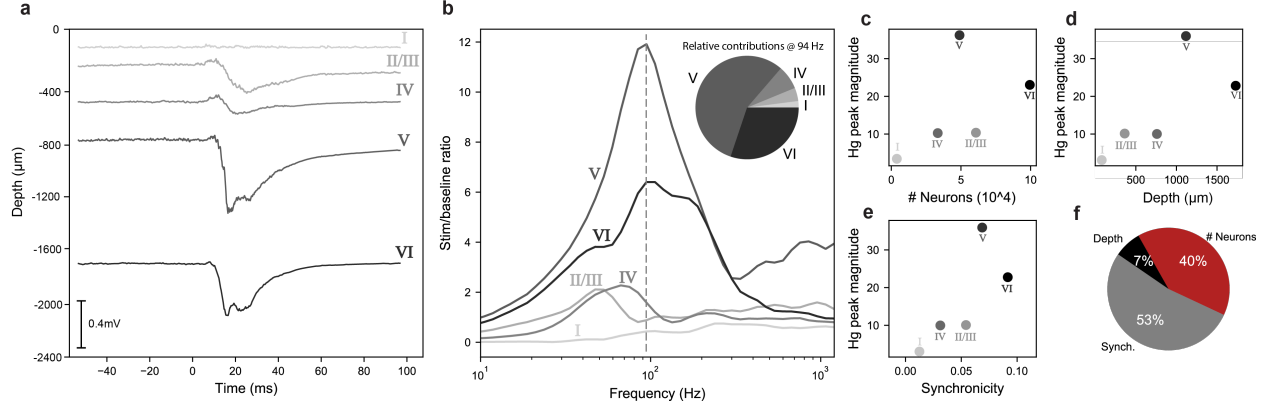
can be seen as a sweep towards the upper right of the individual traces as input magnitude increases. We quantified the relationship between response magnitude (maximum z-scored response across frequencies) and peak frequency (frequency at maximum response). The pink-to-red squares in Figure 12c display the normalized maximum response magnitude vs. normalized peak frequency for varying input amplitudes (color saturation demarcates magnitude of input, circled square demarcates input used in Figs. 11, 13, and 14). Intuitively, these effects were mediated by an increase in the population mean firing rate and spike synchrony resulting from increased input spike rate.

Next, we sought to determine if this relationship between magnitude and frequency existed in the experimental data. Figure 12b displays the z-scored CSEP at an example electrode as a function of frequency in response to the BF stimulus presented at different amplitudes (see inset color bar). Similar to Figure 12a, we observed a sweep towards the upper right of the individual traces with increasing input amplitude (Fig. 12b). For frequency tuned electrodes, we calculated the same quantities (maximum response magnitude and frequency at maximum response) as a function of the amplitude of auditory input at the electrode’s best frequency in the tone stimuli (Fig. 12c; grey-to-black circles, mean  $\pm$  s.d.,  $N \in [206, 299]$ ). As in the simulations, we observed that increasing the input magnitude resulted in an increase in both the magnitude of the peak response and frequency at the peak response. Further, there is a striking correspondence in the curvature of response frequency vs. response magnitude plots derived from experimental and simulation data (Fig. 12c). These results demonstrate a prediction made by the model that was confirmed by a novel experimental finding.

### 3.3.4 Evoked $\mu$ ECoG responses originate in infragranular layers

We next utilized the model to understand the spatial distribution of the generating sources of the CSEP. A key feature of the biophysical model is that CSEP calculation is separate from the numerical simulation of the neurons in the column, enabling us to calculate CSEPs from arbitrary samples of neuronal segments in the column without perturbing the activity at all. We first examined the contributions to the CSEP from cortical layers by computing each layer’s contribution to the CSEP individually. This was done by taking the sum in equation 3.2 over only neuronal segments belonging to neurons in an individual cortical layer (see appendix B.1 for more details on the computation).

Figure 13a plots the raw evoked CSEP (scale bar in inset) as a function of time for each layer, and indicates the average depth of neuronal somas for the layers. Surprisingly, we found that layers V and VI produce the largest evoked potentials, despite being the furthest away, while neurons in superficial layers contribute very little of the total CSEP. Figure 13b shows the frequency content of each contribution during a 10ms window surrounding the response peak. The inset shows the relative magnitudes of the layers contributions in the band centered at 94 Hz, the apex of the high-gamma peak, shown as a dotted vertical line in the main panel. As with the raw evoked potential, we found that infragranular layers also



**Figure 13: Evoked  $\mu$ ECoG responses originate in infragranular layers**

- a.** Contributions to the simulated CSEP from anatomical layers. Top-to-bottom: cortical layers I through VI. The sum of these contributions is the total CSEP.
- b.** Frequency content of the laminar contributions during stimulus peak. Layer V and VI contributions dominate the high-gamma peak.
- c.** Magnitude at peak frequency of each cortical layer's CSEP contribution vs. number of neurons in the layer.
- d.** Magnitude at peak frequency of each cortical layer's CSEP contribution vs. average distance of cell bodies in the layer from the recording electrode.
- e.** Magnitude at peak frequency of each cortical layer's CSEP contribution vs. synchronicity of somatic membrane potentials averaged over all pairs of neurons in the layer.
- f.** Pie chart showing the relative importance of these three factors in a linear model of the high-gamma peak contribution magnitudes of anatomical layers.

contribute most to the high-gamma component of ECoG responses: 51% from layer V, 35% from layer VI, and the remaining 14% coming from layers I-IV.

The results above appear counter-intuitive when the contribution of sources is viewed only as a function of distance. However, in addition to distance, the number of sources and their correlations are additional biophysical factors that dictate the contribution of neuronal populations to a distally recorded signal. *A priori*, the relative importance of these factors to determining laminar contributions to evoked CSEPs in a full-scale cortical column model is not clear. Thus, we plotted each layer's peak high-gamma responses as a function of the number of simulated neurons in (Fig. 13c), the average distance of soma in each layer from the recording electrode (Fig. 13d), and the synchronicity between somatic membrane potentials in each layer during the stimulus (Fig. 13e). We note that in Figure 13d the peak high-gamma response vs. depth shows a positive slope, contrary to the physical principle that individual neurons further from the electrode will contribute less to the signal. However, as is evident from these plots, there are correlations between depth and the other variables. For example, deeper layers tend to contain more neurons. Thus, we fit a regularized linear model to predict peak high-gamma magnitude across layers as a function of depth, number

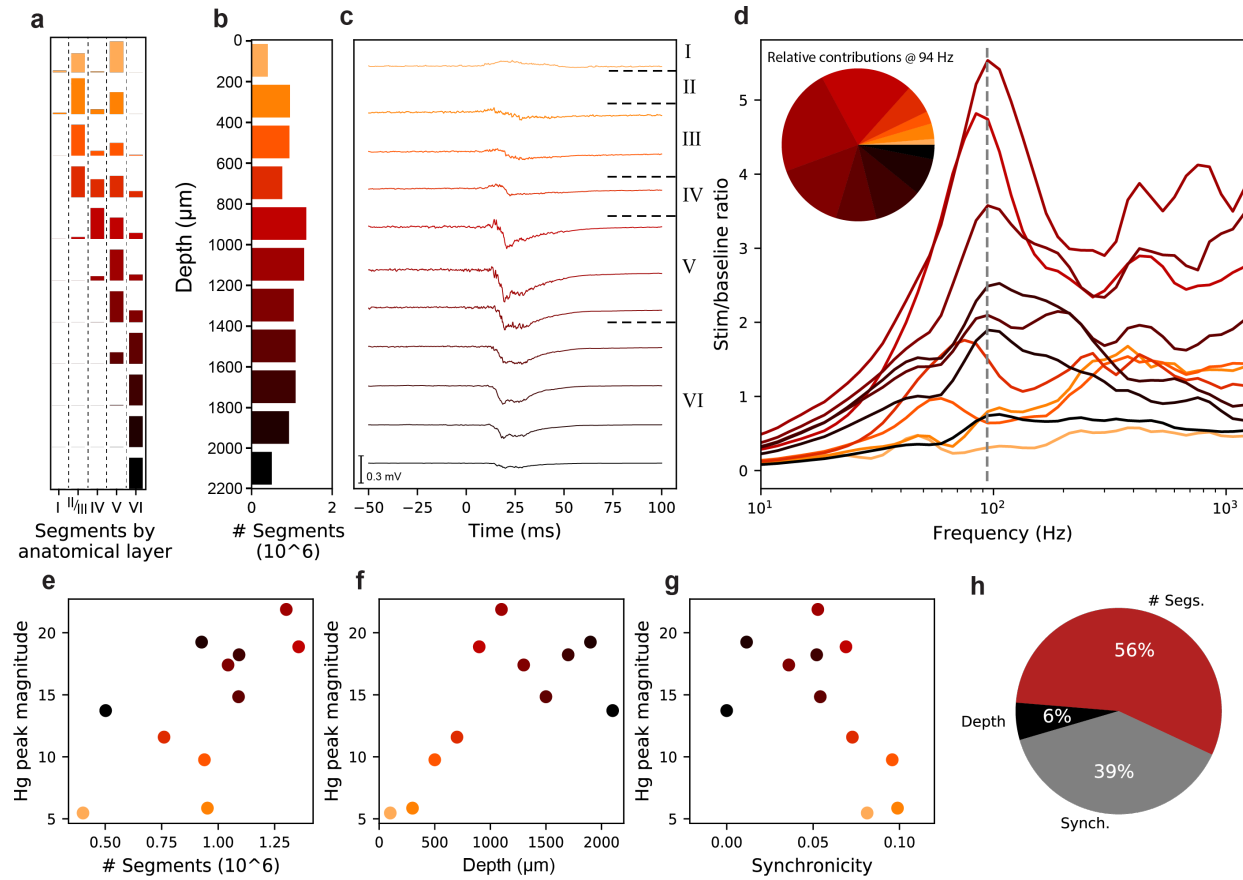
of segments, and synchronicity of neuronal soma simultaneously, which fit the data well ( $R^2 = 0.98$ ). The relative magnitudes of the fit coefficients are plotted in Figure 13f, which shows that the number of segments and synchronicity between cells are much larger determinants of source contributions than is depth. Thus, infragranular layers contribute 86% of evoked CSEP responses due to their increased number of neurons and increased synchronicity.

### 3.3.5 Evoked $\mu$ ECoG responses originate in sources 800-1400 $\mu$ m below the surface

The previous results indicate that layers V and VI are the dominant sources of evoked CSEPs. However, due to the large, extended morphology of some neurons relative to the column depth, knowledge of the largest contributing anatomical layers does not necessarily imply precise knowledge of the spatial distribution of segments generating CSEPs. For example, the apical tufts of many layer V pyramidal neurons reach into layer I. Thus, we next isolate contributions to the CSEP from 200  $\mu$ m slices of the column. Again, this is done by limiting the sum in equation 3.2, this time to only the neuronal segments located within a particular range of depths below the surface (see appendix B.2 for details).

Most slices contain segments from neurons in more than one layer, and a given neuron can contribute to more than one slice. The breakdown of segments in each slice by anatomical layer is shown in Figure 14a, where each color represents one slice. For each slice, five bars are shown displaying the number of segments in that slice belonging to neurons in the five cortical layers. For example, the top slice is dominated by segments from layer V neurons (Fig. 14a, 4th column). The total number of neuronal segments in each slice is shown in Figure 14b, which makes clear that the slices between 800-1200  $\mu$ m have the most segments. Figure 14c shows CSEPs calculated only from segments in the slices as a function of depth (CSEP scale bar is inset). The largest contributors to the evoked responses are the slices located from 800-1400  $\mu$ m below the surface, i.e., in layer V. Interestingly, the top slice was the only one to experience a positive voltage deflection from baseline for the entire duration of the stimulus. Next, we extracted a 10 ms window around the peak of the CSEP response and analyzed the frequency content of each slice's contribution within that window. The results are shown in Figure 14d. The inset shows the relative magnitudes of the slices' contributions at 94 Hz, the apex of the high-gamma peak, shown as a dotted vertical line. Here we see that the slices spanning 800-1400  $\mu$ m are also the ones contributing most to the high-gamma peak (56% total), which is where layer V somas are located. Thus, this analysis strongly suggests that layer V somas are the major generating source of evoked ECoG signals.

As with the layer contributions, we sought to ascertain the relative importance of the number of segments in the slice (Fig. 14e), the depth of the slice below the surface (Fig. 14f), and the synchronicity of membrane potentials of segments within the slice (Fig. 14g) in determining the high-gamma peak contribution magnitude. The results of a regularized linear regression predicting high-gamma peak from those parameters ( $R^2 = 0.91$ ) are shown



**Figure 14: Evoked  $\mu$ ECoG responses originate in sources 800-1400  $\mu$ m below the surface**

**a.** Proportional breakdown of segments by anatomical layer. Most slices contain segments from neurons in multiple cortical layers. Bars represent proportion of total segments in the slice, different slices not to scale.

**b.** Total number of simulated neuronal segments in each 200  $\mu$ m axial slice of the column.

**c.** Contributions to the CSEP from 200  $\mu$ m slices, organized by depth (top: cortical surface). The sum of these contributions is the total CSEP shown in Figure 4b.

**d.** Frequency content of the slice contributions during stimulus peak, colored by slice depth. Slices containing somas of layer V neurons dominate the high-gamma peak.

**e.** Magnitude at peak frequency of each slice's CSEP contribution vs. number of neuronal segments in the slice.

**f.** Magnitude at peak frequency of each slice's CSEP contribution vs. average distance of segments in the slice from the recording electrode

**g.** Magnitude at peak frequency of each slice's CSEP contribution vs. average synchronicity in the slice.

**h.** Pie chart showing the relative importance of the three factors in our linear model of the slices' high-gamma peak contribution magnitudes.

in Figure 14h. Similar to the layer contributions, we find that the number of segments and the synchronicity are the most important factors determining the magnitude of a slice’s contribution to the CSEP.

### 3.4 Discussion

Our cortical column model accurately reproduced the spectral content and the relationship between frequency and magnitude of experimentally observed ECoG responses. By studying the detailed ground truth of simulations that reproduced these properties of *in vivo* CSEPs, we find that the intuition that ECoG signals are generated primarily in superficial layers is incorrect. Instead, our analysis implicates neurons in cortical layer V as the primary source of the signal recorded at the surface, with layer VI also contributing substantially. Similarly, analysis of contributions to the CSEP by depth showed that slices of the column containing layer V soma produce most of the signal observable at the surface. Subsequent analysis found that the density (number) and synchrony of neurons to be more important than depth in determining a population’s contribution to the surface signal. While layers II/III/IV are closer to the recording electrode than layers V/VI, they have fewer neurons [29, 24] and reduced synchrony [3, 1]. Layers V/VI are composed of predominately excitatory pyramidal cells [24, 1], and pyramidal neuron action potentials contribute most to the high-frequencies examined here [9]. Thus, we conclude that evoked high-gamma at the surface is a biomarker of layer V/VI pyramidal neuron firing rates.

Additionally, the simulation results suggest a word of caution for the interpretation of multi-unit activity (and LFPs more broadly) recorded both at the surface and intracortically [21, 26, 31]. In particular, we found that activity in the multi-unit activity range (500-1000Hz) at the surface was predominantly generated by neurons in layer V, sources which are very distant (1-1.5mm) from the recording electrode. Hence, previous reports of ‘single-unit’ recording from the cortical surface, as well as intracortically recorded multi-unit activity (e.g., laminar polytrodes, Utah arrays, etc.) may contain contributions from distal, but numerous and synchronous, neurons.

Although we are unable to construct an explicit inverse model for the problem of CSEP generation (in contrast with the problem of chapter 2, where a neural network was used to explicitly invert the model), our biophysically detailed modeling provides insights that experiments alone cannot. For example, the finding that evoked ECoG high-gamma is primarily generated by neurons in layer V provides a potential explanation of the robust tuning to exogenous variables found here and elsewhere (e.g., auditory stimuli, vocal tract articulators [8], etc.). In particular, neurons in layer V have previously been found to have sharper tuning curves than neurons in layer II/III [5, 4, 16]. Thus, while cortical surface electrical stimulation may activate broadly connected neurons in layer II/III [16, 13], recordings from the surface can reflect the finely tuned responses and precise projections of neurons in layer V [4, 16].

# Chapter 4

## Conclusion

### 4.1 Biophysical simulation is a valuable tool for studying ill-posed inverse problems in neural electrophysiology

This dissertation presents two applications of biophysically detailed simulations of neural activity towards the goal of helping to solve ill-posed inverse problems related to the electrophysiology of neurons and neural networks. In chapter 2, we simulate a large volume of data reflecting the changes in spiking activity caused by very fine modulations of the ionic conductances through a neuronal membrane, and show how a CNN can directly regress the conductances which gave rise to a given somatic potential. Using this method, we are able to infer microscopic parameters in a variety of neuronal models of varying complexity, including the ion channel conductances in the biophysically detailed cells of the BBP model of rat somatosensory cortex. In chapter 3, we develop a simulation of  $\mu$ ECoG signals based on the same BBP model, and show that it provides access to the information we need to determine the spatial localization of the sources generating those signals. We find that  $\mu$ ECoG signals recorded at the surface of the brain are primarily due to activity in infragranular layers, hundreds of microns from the electrode below the cortical surface. We attribute this to the large numbers of neurons and greater synchrony between neurons in those layers. Together, the techniques presented in chapters 2 and 3 demonstrate a few of the key properties of biophysical simulation which make them valuable for studying ill-posed inverse problems in electrophysiology:

1. **Simulations are faster than experiments.** Some techniques that require large volumes of training data may be infeasible due to the time required to collect data. Neural networks for many scientific problems of interest, such as the CNN used in chapter 2, may be subject to this limitation and therefore may benefit from simulated data.

2. **Simulations offer fine-grained control over input parameters.** Understanding the behavior of a system throughout the space of its input parameters can often help solve ill-posed inverse problems. In many systems, such as the neurons in chapter 2, it is difficult to probe the parameter space experimentally. This is another limitation which may be overcome via simulation in situations where the biophysics are well-understood.
3. **Simulations provide unparalleled access to biophysical details.** Most biophysical simulations keep track of many more degrees of freedom than the user is ultimately interested in. For example, in chapter 2, the only simulation result we analyze is the somatic potential, but in order to compute this quantity, the simulation must keep track of the membrane potential on *every* segment of the neuron simultaneously. In chapter 3, we are interested in studying the CSEP, but in order to compute the CSEP, the simulation must keep track of the membrane current on each neuronal segment (among countless other internal variables, often instantiated on each segment in the simulation). These individual segment membrane currents were the key data whose analysis yielded our conclusions about the spatial localization of CSEP generators.

## 4.2 Future Work

### 4.2.1 *In vitro* inference of ion channel conductance

In chapter 2, we presented a technique for inferring the conductances of various ion channels using only the somatic membrane potential during a current injection, which is much easier to measure directly than the conductances themselves. It was important to use an input that was experimentally accessible because our goal is ultimately to bring our method to the lab bench, where scientifically useful predictions of ion channel density can be made on real neurons.

While our results are promising, and we are able to demonstrate some degree of generalizability across neuronal morphologies and celltypes, there are a few limitations which must be overcome before our method becomes experimentally viable. Firstly, the volume of training data required to make accurate predictions on biophysically detailed cells remains prohibitively large, even when data from multiple cells are combined. Additionally, it is unclear how experimental sources of variability in the spiking behavior of neurons will affect the CNN’s learning rate or ability to make predictions. Preliminary experiments using neurons *in vitro* show some amount of variability between multiple membrane potential waveforms obtained from a single neuron in one recording session using the chirp stimulus shown in Fig. 1, though it is not yet known to what degree the experimental variability we observed will affect the prediction accuracy.



### 4.2.2 Probing deeper into the origins of ECoG signals

While the simulations in chapter 3 have already revealed much about the biophysical origin of stimulus-evoked ECoG signals beyond what we learned from experimentation alone, there is still a wealth of information available in the simulation which could provide further insights into the origin of these signals. For example, as discussed in section 3.4, the results of figures 13 and 14, taken in conjunction, strongly suggest that layer V soma are responsible for the majority of the ECoG signal. By examining the detailed spatial distribution of the membrane currents throughout each neuron in the simulation, it should be possible to determine the effect of contributions to the surface from soma versus from dendrites. For example, one could simply separate these two types of contributions and compare their magnitudes. We expect somatic contributions to exceed dendritic contributions for most cell types in most layers.

Another possibility enabled by the accessibility of the detailed spatial distribution of currents in the simulation is to understand the importance of dipoles in generating surface potentials. When charges (or currents) of opposing polarity exist in close proximity, the charge (or current) configuration possesses a property known as its dipole moment, and the potential produced by such a configuration is strongly directional. In the brain, neurons can exhibit strong dipole moments when the current through their apical dendrites opposes that through their basal dendrites, or through their soma, which happens often in order to maintain electroneutrality [9]. The link between dipoles and electroencephalography (EEG) is well-understood [9, 25], but it is not clear whether the same analysis applies to ECoG electrodes which are much closer to the sources (dipole analysis of electric fields is typically most useful in the limit of large distances from the source, as such analyses usually neglect higher moments of the charge distribution’s multipole expansion which are responsible for its near-field effects [20]).

### 4.2.3 Explicit inverse model of CSEPs

The problem of determining ionic conductances from somatic membrane potentials  $X(t)$  was solved in chapter 2 by a CNN which learned an explicit inverse model allowing it to predict a vector  $\mathbf{Y}$  of numerical values for the target conductances  $\bar{g}_x$  (or target parameters  $a$ ,  $b$ ,  $c$ , and  $d$  for the Izhikevich model). However, to determine the spatial and neuronal origins of stimulus-evoked CSEPs, we employed a different strategy in which the contributions  $V_{j(k)}(\mathbf{r}, t)$  to the CSEP from layers (slices) of the column were directly computed from the simulation ground-truth, the membrane currents  $i_i$ . We propose that the CNN approach may be applicable to the ECoG origin problem as well. Such a CNN would take as input the full CSEP  $V(\mathbf{r}, t)$  and return the vector of layer (slice) contributions  $V_{j(k)}(\mathbf{r}, t)$ . An alternative approach would be to have the CNN return the distribution  $i_i(z)$  of membrane currents as a function of depth  $z$  below the surface (binned to a suitable spatial resolution).

A major impediment to the application of these CNN methods to the CSEP origin problem is the dataset likely required for sufficient training. While the parameters of the

single cell models are easily and directly modifiable *in silico* in order to generate a large volume of training data uniformly sampling the space of ionic conductance values (we can easily assign new conductances to any compartment of a cell in the simulator), it is less clear how to vary the spatial distribution of activity in our cortical column model for training data generation. One option would be to change the spatial distribution of thalamic input to the column (Fig. 11b), but there is not much room for variation before the input distribution no longer reflects the actual *in vivo* distribution. Another option would be to simulate electrical or optogenetic stimulation or suppression of cells at various depths. This has the benefit of being experimentally realizable. Indeed, future studies (in preparation) will use optogenetic perturbations to assess the relative contributions of different layers and celltypes to surface signals. None of these potential solutions, however, offers quite the same level of control as modifying the conductances in a single cell model, which will make *uniform* sampling a significant challenge.

#### 4.2.4 Finite Element Model of neural activity and extracellular potential

The insights that a simulation can provide into a given inverse problem are only as good as the models which the simulation implements. While the Cable Equation (employed by the NEURON simulator used here) remains the top choice for most biophysically detailed *in silico* studies in neuroscience today, it does not provide a model of extracellular space, thus precluding the possibility of simulating ephaptic interactions between cells (communication that occurs via electrical fields rather than synapses) using the Cable Equation alone. That is, Cable Equation models treat all neuronal segments as totally independent, unless they are in *direct* contact with one another (i.e., they are neighboring segments on the same neuron, or they are connected by a synapse). This is computationally convenient, but may fail to capture the behavior of neurons accurately, especially their synchrony, as ephaptic interactions tend to synchronize populations of neurons [15]. Recently, simulations employing Finite Element techniques to solve Maxwell's equations directly on a mesh representing the neuronal membrane as well as both intracellular and extracellular spaces have been successful at capturing the effects of extracellular potentials acting back onto the populations of neurons that produced them [2]. Such studies also permit modeling of non-cylindrical neuronal segments and the smoothing of hard edges where cylinders connect [2].

The lack of an explicit model of extracellular space has another crucial effect when extracellular potentials such as the CSEP are considered. As discussed in section 3.2.2, in order to develop an analytic expression for the extracellular potential due to a neuronal segment in a Cable Equation simulation (the Line Source Approximation), we treat each point on the segment in isolation, and we assume extracellular space to be electrically homogenous. These assumptions are incompatible with the presence of cell membranes - both neuronal and glial - where the conductivity is near zero, while the ionic solution making up most of the rest of the extracellular space has very high conductivity. Furthermore, we assume extracellular

space to be isotropic, which is incompatible with the known laminar and columnar structure of cortex. Inhomogeneities in the extracellular space induce variations in the amplitude of simulated potentials owing to the secondary fields they produce, which superpose with the primary fields produced by the neurons themselves. It has been shown that the amplitude of these secondary fields varies with frequency [6] - an important effect to capture for analyses in the frequency domain such as the one in chapter 3 (see Figs. 10e, 11f, 12ab, 13b, and 14d). While our results show that today's neural simulations are accurate enough to help solve inverse problems in a range of important areas, these improvements to the simulation model will likely increase the power of neural simulations to provide insight into some of the more challenging ill-posed inverse problems in the field.

# Bibliography

- [1] Hillel Adesnik and Alexander Naka. “Cracking the Function of Layers in the Sensory Cortex”. In: *Neuron* 100 (Dec. 2018), pp. 1028–1043. DOI: 10.1016/j.neuron.2018.10.032.
- [2] Andres Agudelo-Toro and Andreas Neef. “Computationally efficient simulation of electrical activity at cell membranes interacting with self-generated and externally imposed electric fields”. In: *Journal of Neural Engineering* 10.2 (2013). URL: <https://iopscience.iop.org/article/10.1088/1741-2560/10/2/026019>.
- [3] Craig A. Atencio and Christoph E Schreiner. “Auditory cortical local subnetworks are characterized by sharply synchronous activity”. In: *J Neurosci.* 33.47 (Nov. 2013). DOI: 10.1523/JNEUROSCI.2014-13.2013.
- [4] Craig A. Atencio and Christoph E. Schreiner. “Laminar Diversity of Dynamic Sound Processing in Cat Primary Auditory Cortex”. In: *Journal of Neurophysiology* 103.1 (2010). PMID: 19864440, pp. 192–205. DOI: 10.1152/jn.00624.2009. eprint: <https://doi.org/10.1152/jn.00624.2009>. URL: <https://doi.org/10.1152/jn.00624.2009>.
- [5] Craig A. Atencio and Christoph E. Schreiner. “Spectrotemporal Processing in Spectral Tuning Modules of Cat Primary Auditory Cortex”. In: *PLOS ONE* 7.2 (Feb. 2012), pp. 1–15. DOI: 10.1371/journal.pone.0031537. URL: <https://doi.org/10.1371/journal.pone.0031537>.
- [6] Claude Bédard, Helmut Kröger, and Alain Destexhe. “Model of low-pass filtering of local field potentials in brain tissue”. In: *Phys. Rev. E* 73 (5 May 2006), p. 051911. DOI: 10.1103/PhysRevE.73.051911. URL: <https://link.aps.org/doi/10.1103/PhysRevE.73.051911>.
- [7] Roy Ben-Shalom et al. “Inferring neuronal ionic conductances from membrane potentials using CNNs”. In: *bioRxiv* (2019). DOI: 10.1101/727974. eprint: <https://www.biorxiv.org/content/early/2019/08/06/727974.1.full.pdf>. URL: <https://www.biorxiv.org/content/early/2019/08/06/727974.1>.
- [8] Kristofer E Bouchard et al. “Functional organization of human sensorimotor cortex for speech articulation”. In: *Nature* 495.7441 (Mar. 2013), pp. 327–332. DOI: 10.1038/nature11911.

- [9] György Buzsáki, Costas A. Anastassiou, and Christof Koch. “The origin of extracellular fields and currents — EEG, ECoG, LFP and spikes”. In: *Nature Reviews Neuroscience* 13 (May 2012), pp. 407–420. DOI: 10.1038/nrn3241. URL: <https://www.nature.com/articles/nrn3241>.
- [10] Nicholas T. Carnevale and Michael L. Hines. *The NEURON Book*. Cambridge, UK: Cambridge University Press, 2006.
- [11] W. R. Foster, L. H. Ungar, and J. S. Schwaber. “Significance of conductances in Hodgkin-Huxley models”. In: *Journal of Neurophysiology* 70.6 (1993). PMID: 7509859, pp. 2502–2518. DOI: 10.1152/jn.1993.70.6.2502. eprint: <https://doi.org/10.1152/jn.1993.70.6.2502>. URL: <https://doi.org/10.1152/jn.1993.70.6.2502>.
- [12] Mark S. Goldman et al. “Global Structure, Robustness, and Modulation of Neuronal Models”. In: *Journal of Neuroscience* 21.14 (2001), pp. 5229–5238. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.21-14-05229.2001. eprint: <https://www.jneurosci.org/content/21/14/5229.full.pdf>. URL: <https://www.jneurosci.org/content/21/14/5229>.
- [13] Michael S.A Graziano, Charlotte S.R Taylor, and Tirin Moore. “Complex Movements Evoked by Microstimulation of Precentral Cortex”. In: *Neuron* 34.5 (2002), pp. 841–851. ISSN: 0896-6273. DOI: [https://doi.org/10.1016/S0896-6273\(02\)00698-0](https://doi.org/10.1016/S0896-6273(02)00698-0). URL: <https://www.sciencedirect.com/science/article/pii/S0896627302006980>.
- [14] Jacques Hadamard. “Sur les problèmes aux dérivées partielles et leur signification physique”. In: *Princeton University Bulletin* (1902), pp. 49–52.
- [15] Kyung-Seok Han et al. “Ephaptic Coupling Promotes Synchronous Firing of Cerebellar Purkinje Cells”. In: *Neuron* 100 (2018). DOI: 10.1016/j.neuron.2018.09.018. URL: [https://www.cell.com/neuron/pdfExtended/S0896-6273\(18\)30820-1](https://www.cell.com/neuron/pdfExtended/S0896-6273(18)30820-1).
- [16] Thomas C Harrison, Oliver G S Ayling, and Timothy H Murphy. “Distinct cortical circuit mechanisms for complex forelimb movement and motor map topography”. In: *Neuron* 74.2 (Apr. 2012), pp. 397–409. DOI: 10.1016/j.neuron.2012.02.028.
- [17] Michael L. Hines and Nicholas T. Carnevale. “Expanding NEURON’s Repertoire of Mechanisms with NMODL”. In: *Neural Computation* 12 (2000), pp. 839–851.
- [18] Alan L. Hodgkin and Andrew F. Huxley. “A quantitative description of membrane current and its application to conduction and excitation in nerve”. In: *The Journal of physiology* 117.4 (Aug. 1952). 12991237[pmid], pp. 500–544. ISSN: 0022-3751. DOI: 10.1113/jphysiol.1952.sp004764. URL: <https://www.ncbi.nlm.nih.gov/pubmed/12991237>.
- [19] Eugene M. Izhikevich. “Simple model of spiking neurons”. In: *IEEE Transactions on Neural Networks* 14.6 (Nov. 2003), pp. 1569–1572. ISSN: 1045-9227. DOI: 10.1109/TNN.2003.820440.
- [20] John David Jackson. *Classical electrodynamics*. 3rd ed. New York, NY: Wiley, 1999. ISBN: 9780471309321. URL: <http://cdsweb.cern.ch/record/490457>.

- [21] Dion Khodagholy et al. “NeuroGrid: recording action potentials from the surface of the brain”. In: *Nature Neuroscience* 18 (2015), pp. 310–315.
- [22] Diederik P. et al. Kingma. “Adam: A Method for Stochastic Optimization”. In: *CoRR* abs/1412.6980 (2014). arXiv: 1412.6980.
- [23] Zachary F. Mainen. “Influence of dendritic structure on firing pattern in model neocortical neurons.” In: *Nature* (1996).
- [24] Henry Markram et al. “Reconstruction and Simulation of Neocortical Microcircuitry”. eng. In: *Cell (Cambridge)* 163.2 (2015), pp. 456–492. ISSN: 0092-8674.
- [25] Paul L. Nunez and Ramesh Srinivasan. *Electric Fields of the Brain*. New York: Oxford University Press, 2006.
- [26] Angelique C Paulk et al. “Microscale Physiological Events on the Human Cortical Surface”. In: *Cerebral Cortex* (Mar. 2021). bhab040. ISSN: 1047-3211. DOI: 10.1093/cercor/bhab040. eprint: <https://academic.oup.com/cercor/advance-article-pdf/doi/10.1093/cercor/bhab040/36650575/bhab040.pdf>. URL: <https://doi.org/10.1093/cercor/bhab040>.
- [27] Astrid A. Prinz, Dirk Bucher, and Eve Marder. “Similar network activity from disparate circuit parameters”. In: *Nature Neuroscience* 7.12 (2004), pp. 1345–1352. ISSN: 1546-1726. DOI: 10.1038/nn1352. URL: <https://doi.org/10.1038/nn1352>.
- [28] Philip Schwan. “Lustre: Building a File System for 1,000-node Clusters”. In: *Proceedings of the Linux Symposium*. Ottawa, Ontario, 2003, pp. 380–386.
- [29] G. M. Shepherd, ed. *The Synaptic Organization of the Brain*. Oxford: Oxford University Press, 1990.
- [30] Leslie N. Smith. “A disciplined approach to neural network hyper-parameters: Part 1 - learning rate, batch size, momentum, and weight decay”. In: *CoRR* abs/1803.09820 (2018). arXiv: 1803.09820.
- [31] Eric M Trautmann et al. “Accurate Estimation of Neural Population Dynamics without Spike Sorting”. In: *Neuron* 103 (2010), pp. 292–308.

# Appendix A

## Derivation of Line Source Approximation

Equation 3.2, the Line Source Approximation for the extracellular current due to a neuron's activity, can be derived for a single point in time by first considering an isolated point current source (neuronal segment) embedded in a perfectly homogeneous, isotropic, and purely Ohmic medium, emitting a current  $i_i$ . By symmetry, the current spreads out to an area of  $4\pi r^2$  at a distance of  $r$  from the source and thus has a density of

$$J = \frac{i_i}{4\pi r^2} \quad (\text{A.1})$$

We now invoke the differential form of Ohm's law, dropping the polar and azimuthal terms, which will vanish after integration in the radial direction:

$$J = \sigma E = -\sigma \frac{dV}{dr} \quad \left( + \dots \frac{dV}{d\phi} + \dots \frac{dV}{d\theta} \right)$$

We replace the left hand side with the current density from (A.1):

$$\frac{i_i}{4\pi\sigma r^2} = -\frac{dV}{dr}$$

Then integrate radially from a reference point at  $R_{ref}$  to the measurement point at  $R$ :

$$\frac{i_i}{4\pi\sigma} \left( \frac{1}{R} - \frac{1}{R_{ref}} \right) = V(R) - V(R_{ref})$$

Then we choose our reference point  $R_{ref} = \text{infinity}$  and define  $V(R_{ref}) = 0$ :

$$V(R) = \frac{i_i}{4\pi\sigma R} \quad (\text{A.2})$$

This is the *Point Source Approximation* for the extracellular potential at  $R$  due to a point source of current at the origin. If the source is a distance  $R'$  from the origin, the potential is

$$V(R) = \frac{i_i}{4\pi\sigma} \frac{1}{R - R'}$$

Finally, equation 3.2 can be obtained for a given time  $t$  by integrating this expression along the length of each neuronal segment and summing over all segments.



# Appendix B

## Computation of CSEP contributions from simulation

### B.1 Laminar contributions

By summing only neuronal segments belonging to neurons in an individual cortical layer, we obtain the contributions  $V_j(\mathbf{r}, t)$  to the CSEP from that layer:

$$V_j(\mathbf{r}, t) = \sum_{i \in L_j} \frac{i_i(t)}{4\pi\sigma} \int \frac{d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|}$$

Where  $j \in [1, 6]$  denotes the layer whose contribution to  $V(\mathbf{r}, t)$  is represented by  $V_j(\mathbf{r}, t)$ , and  $L_j$  is the set consisting of all neuronal segments comprising neurons in layer  $j$ . The full signal is then the sum of these contributions:

$$V(\mathbf{r}, t) = \sum_j V_j(\mathbf{r}, t)$$

### B.2 200 $\mu\text{m}$ slice contributions

By summing only the neuronal segments located within a particular range of depths below the surface, we obtain contributions to the CSEP from 200  $\mu\text{m}$  slices of the column. The CSEP contribution  $V_k(\mathbf{r}, t)$  from the  $k^{\text{th}}$  slice below the surface is given by:

$$V_k(\mathbf{r}, t) = \sum_{S_k} \frac{i_i(t)}{4\pi\sigma} \int \frac{d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|}$$

Where  $S_k = \{i : z_i \in [k, k + 1] \times 200\mu\text{m}\}$  is the set of neuronal segments whose midpoints are between  $200k$  and  $200(k + 1)$   $\mu\text{m}$  below the surface. The full signal is then the sum of these contributions:

$$V(\mathbf{r}, t) = \sum_k V_k(\mathbf{r}, t)$$

Note that most neurons' dendritic arbors extend beyond the slice boundaries, therefore each slice contains segments from neurons in a multitude of layers, and a given neuron may contribute to multiple slices.