Modelling structure-function relationships from neurons to networks in the cerebral cortex

Doctoral thesis

to obtain a doctorate (PhD)

from the Faculty of Medicine

of the University of Bonn

Rieke Fruengel

from Darmstadt, Germany

2025

Written with authorization of

the Faculty of Medicine of the University of Bonn

First reviewer: Marcel Oberlaender

Second reviewer: Tatjana Tchumatchenko

Day of oral examination: 13.06.2025

From the Max Planck Institute for Neurobiology of Behavior – caesar

1 Table of Contents

1 Table of Contents 3
2 List of abbreviations 5
3 Introduction 7
3.1 Cortex-inspired artificial neural networks 10
3.2 Neuron-network interactions which shape L5PT receptive fields 12
3.3 Subcellular scale thalamocortical wiring specificity 13
4 Methods 16
4.1 Cortex-inspired ANNs 16
4.2 Reduced models 22
4.2.1 Biophysically detailed models and multi-scale model simulations 23
4.2.2 Reduced model inference and simulations 24
4.2.3 Analysis 27
4.3 Thalamocortical wiring specificity 28
4.3.1 In vivo experiments 28
4.3.2 Multi-scale model simulations 29
4.3.3 Multi-scale model manipulations 31
4.3.4 Categorising sensory- vs non sensory-evoked CaAPs 32
5 Results 33
5.1 Sparse connectivity enables efficient information processing in cortex-like neural networks 33
5.1.1 Training and manipulating cortical column ANNs 33
5.1.2 Sparse connectivity enables time- and data- efficient training of large and recurrent networks 36
5.1.3 Sparsely connected networks form distributed, robust representations 40

5.1.4 Sparse connectivity facilitates efficient training in networks with fixed excitatory and inhibitory nodes 43
5.1.5 Beyond random connectivity: degree sequences and network motifs 46
5.2 Network-neuron interactions underlying sensory responses of layer 5 pyramidal tract neurons in barrel cortex 51
5.2.1 Reduction of the multi-scale models into analytically tractable models 51
5.2.2 Input-output computation is robust to morphological and biophysical diversity 56
5.2.3 Linear reduced models capture effects of nonlinear mechanisms 58
5.2.4 Contribution of different input pathways to sensory responses 61
5.2.5 Network vs neuron contributions to receptive field variability 64
5.3 The origin of bursts in sensory-evoked cortical output 66
5.3.1 Characterisation of calcium potentials in the distal dendrites 66
5.3.2 The network origin of CaAPs and triplets 67
5.3.3 Early triplets and CaAPs have different origins 74
5.3.4 Triplets distinguish awake from anaesthetised animals 76
6 Discussion 78
6.1 Sparse connectivity enables efficient information processing in cortex-like neural networks 79
6.2 Network-neuron interactions underlying sensory responses of layer 5 pyramidal tract neurons in barrel cortex 83
6.3 The origin of bursts in sensory-evoked cortical output 85
7 Abstract 88
8 List of Figures 89
9 Supplementary Material 91
10 References 106

2 List of abbreviations

AAV	Adeno-associated virus
AMPA	α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
ANN	Artificial neural network
AP	Action potential
AUROC	Area under the receiver operating characteristic curve
BP	Primary bifurcation point of the apical dendrite
CaAP	Calcium action potential
ChR2	Channelrhodopsin
EPSP	Excitatory post-synaptic potential
GLM	Generalised linear model
IC	Intracortical
IPSP	Inhibitory post-synaptic potential
L2/3PN	Layer 2/3 pyramidal neuron
L4SP	Spiny neurons in L4
L5IT	Layer 5 intratelencephalic neurons
L5PT	Layer 5 pyramidal tract neuron
L6CC	Layer 6 cortico-cortical neuron
L6CT	Layer 6 corticothalamic neurons
LFP	Local field potential
NMDA	N-methyl-D-aspartate
POm	Posteromedial nucleus of the thalamus

PSTH	Peri-stimulus time histogram
PW	Principal whisker
ReLU	Rectified linear unit
RNN	Recurrent neural network
SW	Surround whisker
TC	Thalamocortical
uPSP	Unitary post-synaptic potential
VPM	Ventral posterior medial nucleus of the thalamus
WNI	Weighted net input

3 Introduction

Understanding how a network's structure relates to its function as an information processing system is challenging. Since Ramón y Cajal discovered over a century ago that neurons from different parts of the nervous system have very different morphologies, investigating how the structure of neurons and networks in the brain contributes to their function has been an active field of study. Unravelling structure-function relationships in brain networks remains challenging, not least due to the overwhelming number and diversity of neurons, and the intricacy of the connected networks they form. Complexity arises from interactions between dynamic systems at different scales, from nonlinear biophysical properties of individual neurons to long-range anatomical projections between distant brain regions. Furthermore, measurements at the required level of detail and temporal resolution often remain elusive experimentally, and even with the necessary measurements, inferring causal relationships from observational data is often intractable.

To address these challenges, I here largely work in a system in which I have access to any data one could want: a highly detailed, multi-scale model of the rat primary vibrissal somatosensory cortex (barrel cortex), the brain region which processes sensory information from the animal's whiskers (Feldmeyer et al., 2013). The multi-scale model is built from decades of empirical anatomical and functional measurements of the barrel cortex at various scales. These data are used to constrain the model for simulations which reproduce experimental measurements of cortical function. Structurally, it consists of a dense neuropil built from *in vivo* labelled reconstructed dendrite and axon morphologies of all major cortical cell types, as well as thalamocortical axons (Egger et al., 2014). Using empirical measurements of synapse density and axo-dendritic projection patterns, the model provides realistic estimates for connectivity at cellular and subcellular scales. These connectivity estimates were shown to not only be accurate in terms of pairwise connectivity, but also to reproduce empirical measurements of synapse clustering and network motif occurrences (Udvary et al., 2022). Each cell in the model can be activated according to *in vivo* measured cell type-specific activity during a whisker stimulus, which in combination with synaptic

connectivity from the anatomical model provides a realistic estimate of the synaptic activity impinging onto any cell in the model during a sensory stimulus (Egger et al., 2020).

The multi-scale model has been used to perform well-constrained simulations of layer 5 pyramidal tract neurons (L5PTs), the major output cell type of the neocortex (Harris and Shepherd, 2015), in response to sensory stimuli of the whiskers (Egger et al., 2020). This required the inclusion of a biophysically detailed single-cell model of a L5PT, which can transform synaptic input into spiking output. The multi-scale model has since been extended to include a diverse set of biophysically detailed multi-compartmental models of L5PTs, which span the full range of morphological and biophysical diversity of the cell type, while all reproducing the characteristic electrophysiology and sensory responses of L5PTs (Bast and Oberlaender, 2021). However, although the multi-scale model accurately reproduces a wealth of experimental measurements, it does not *per se* provide much understanding about the cellular and circuit mechanisms underlying sensory processing.

Modelling provides an additional benefit beyond perfect access to data: the ability to exactly repeat a trial and make targeted manipulations. In order to disentangle the contributions of various neuron and network properties, one may want to perform manipulations to remove or modify specific features and assess the effect of this manipulation on the network's function. However, these kinds of specific manipulations, for example altering connectivity or biophysical properties, or modifying the activity of a specific cell type without changing any others, are often inaccessible experimentally. Furthermore, even if keeping all conditions identical in an experiment, it is highly likely that e.g. the exact spatiotemporal pattern of synapse activations will differ from a previous trial, or that the starting membrane voltage before a stimulus will not be identical. These many sources of trial-to-trial variability make it challenging to determine the precise causal effect of a manipulation, leaving aside substantial cell-to-cell and animal-to-animal variability (Marder and Taylor, 2011). Stochasticity is a fundamental feature of brain function, starting with noisy ion channel conductance (Rusakov et al., 2020). The ability to keep everything exactly the same, and only make targeted manipulations, is therefore unique to modelling, and allows us to investigate fine-grained causal relationships.

When investigating structure-function relationships in brain networks, we could work at a range of different scales: from protein interactions on a molecular level (Agnati et al., 2008), through cellular-scale synaptic connectivity, to brain regions connected by white matter tracts, the brain fundamentally consists of links and interactions between its components. Formally, a network is defined as a set of objects (nodes) of which at least some pairs are somehow related - these relations are denoted by edges between the nodes. Here, I focus on cellular scale networks, where each neuron represents a node in the network, and synaptic connections between neurons form the network's edges. Hereby, we describe the network's structure, comprising the neurons and their properties, and the anatomical connections between them. Within the cellular scale networks which we model here, we may describe components at different levels of detail. For instance, we may note simply whether a neuron is excitatory or inhibitory, or we may consider its specific cell type and location. We may represent each neuron as a very simplified computational unit, like an integrate-and-fire model, or account for intricate morphological detail and nonlinear biophysical properties. Regarding edges, we may deem it sufficient to simply use a binary connectivity matrix (a square matrix where each entry indicates whether or not a pair of neurons are connected), or we may care about the specific number and dendritic location of synapses which make up each connection. The aim of this thesis is to use modelling approaches at three levels of detail to investigate how structure affects function in cortical networks:

1. Macro-scale architecture and connectivity, abstracted through artificial neural networks (ANNs) with cortex-inspired architecture.
2. Circuit-level, accounting for biophysical detail and realistic synaptic input to uncover how interactions between neuronal properties and network input shape the receptive fields of L5PTs in barrel cortex.
3. Subcellular scale, following the observation that direct primary thalamocortical (TC) innervation to L5PTs has a density peak around the region which initiates calcium action potentials (CaAPs) in the apical dendrite. What is the relevance of this wiring specificity for calcium-mediated burst firing in response to sensory stimuli?

3.1 Cortex-inspired artificial neural networks

Cortical networks are very sparsely connected. In fact, it has recently been shown that in a given subvolume of sensory cortex, <1 % of neurons with overlapping axons and dendrites will form a synaptic connection (Udvary et al., 2022). What is the relevance of such sparse connectivity for a network's function? Surprisingly, sparse connectivity was shown to impair information processing in ANNs, making training more difficult and leading to worse performance (Evci et al., 2019). Do these findings in ANNs imply that sparse connectivity also impairs information processing in biological neural networks, or are there conditions under which sparsity may be beneficial?

Although the ANNs underlying much of modern deep learning were originally inspired by the brain, there are major differences between conventional ANNs and biological neural networks. While neurons are arranged in layers consisting of different cell types, there are abundant recurrent connections within and between layers (reviewed in Singer, 2021). Meanwhile, conventional ANNs are typically initialised with dense, feedforward connectivity, meaning that each node in a given layer is connected unidirectionally to all nodes in the subsequent layer. In the brain, a neuron is either excitatory or inhibitory (Dale's law; Eccles, 1976), whereas conventional ANNs place no such constraints on weights, allowing individual nodes to have both positive and negative outgoing connections. Determining the relevance of these differences in structure for network function is challenging. It may seem an obvious approach to take the empirically measured connectivity from e.g. a dense electron microscopic reconstruction or the multi-scale model, and use it to construct an ANN replica with biologically realistic connectivity. Unfortunately, such a detailed replica of cortical connectivity is challenging to compare to other architectures in order to understand which structural properties are actually relevant for function. How can we isolate the effect of individual features of cortical networks on information processing in a network?

To this end, I systematically generate and train artificial neural networks (ANNs) constrained by selected, interpretable features of cortical network architecture. I constructed ANNs with different degrees of sparsity in the hidden layers – the sparser the network, the fewer nodes were connected to each other by a trainable weight. Connectivity in cortex is highly recurrent,

so I compared the effect of sparsity in feedforward ANNs and recurrent neural networks (RNNs). It has been suggested that the degree of sparsity in cortex is affected by the size of the brain, with larger brains having sparser connectivity (Herculano-Houzel et al., 2010). This relationship between network size and sparsity has also recently been shown experimentally in CA3 hippocampus (Watson et al., 2025), with humans having sparser connectivity than rats, and rats, in turn, having sparser connectivity than mice. Therefore, I here investigated sparsity and recurrence in networks of different sizes. I find that in networks which are large and recurrent, like cortical networks, sparse connectivity enables networks to achieve better performance when training time or data are limited than conventional, dense connectivity.

An important function of sensory areas in the cortex is to encode inputs in a way that enables them to be distinguished by downstream areas. Similarly, ANNs must encode their inputs in the activations of nodes in the hidden layers in order to allow classification at the output layer. I therefore next investigated how our ANNs represent their inputs in the activations of the hidden layer nodes, and to what extent connectivity in the hidden layer affects this input representation. I find that although in the case of large, recurrent networks, both sparsely and densely connected networks form sparse representations of their inputs, sparse networks use more distributed representations which are more robust to neuronal noise.

Finally, I constructed and trained ANNs which obey Dale's principle, which states that a neuron releases the same set of neurotransmitters at all of its synapses (Eccles, 1976); broadly, each neuron is either excitatory or inhibitory and must remain that way. For ANNs, this corresponds to each node having exclusively positive or negative outgoing weights, and this sign remaining unchanged throughout the training process. It has been found that applying such constraints to conventional ANNs often impairs their training (Cornford et al., 2020). Here, I constructed Dale-compliant RNNs with a proportion of inhibitory nodes corresponding to the proportion of inhibitory neurons reported in sensory cortex (Meyer et al., 2011). I find that the training of densely connected networks is indeed severely slowed by Dale's principle, but that sparse connectivity enables Dale's networks to train almost as efficiently as their unconstrained counterparts.

3.2 Neuron-network interactions which shape L5PT receptive fields

L5PTs are the main output cell type of the cortex, integrating synaptic inputs along their extensive, biophysically complex dendrites which span all cortical layers, and sending action potential (AP) outputs to subcortical targets (Harris and Shepherd, 2015; Rojas-Piloni et al., 2017). Following a sensory (whisker) stimulus, L5PTs in the barrel cortex show fast and reliable AP responses (de Kock et al., 2007). How L5PTs transform synaptic input into AP output – their effective input-output computation – remains unclear, in no small part due to many sources of trial-to-trial and cell-to-cell variability, which make it challenging to elucidate general computational principles. L5PTs show substantial variability in their dendritic morphology (Hay et al., 2013). Biophysical properties (i.e., the density and distribution of ion channels on the dendrite) may also vary greatly from cell to cell: while this is challenging to measure experimentally, it has been shown that the characteristic electrophysiology of L5PTs can be reproduced by biophysically detailed models with vastly different biologically plausible ion channel expression and utilisation (Bast and Oberlaender, 2021). Additional variability arises from the synaptic input received by the cell: it has been found experimentally that the spatiotemporal input pattern (i.e., when and where on the dendrite synapses are active) is highly variable from cell-to-cell and even trial-to-trial within the same experimental condition (Chen et al., 2011; Varga et al., 2011; Rochefort and Konnerth, 2012; Jia et al., 2014; Scholl et al., 2021). We would therefore like to disentangle the relevance of morphological and biophysical properties for the input-output computation performed by L5PTs, and determine how AP responses arise from the interplay between network input and neuronal properties.

Addressing these questions would ideally require simultaneous measurements of sensory-evoked synaptic input at high spatiotemporal resolution and AP output, as well as a morphological reconstruction of the dendrites and their ion channel distributions. While this remains challenging experimentally, such data is accessible in the multi-scale model, which reproduces the fast sensory-evoked responses of L5PTs seen *in vivo*. The multi-scale model provides realistic estimates of the timing and location of synapse activations which a L5PT receives from the network after a whisker stimulus, and uses biophysically detailed models to simulate how L5PTs transform these synaptic inputs into AP outputs (Egger et al., 2020).

Here, we developed an approach which reduces the multi-scale model into analytically tractable models which capture the input-output computation of L5PTs upon single whisker deflections, while preserving cell-to-cell and trial-to-trial variability. These reduced models revealed that three features are sufficient to predict an AP response: the count of active synapses in a time window, their distance from the soma, and the time since the previous AP of the L5PT. They show that this input-output computation is qualitatively preserved across L5PTs with different morphologies and biophysical properties. Thanks to the interpretability of the reduced models, we use them to investigate the circuit mechanisms underlying another property of L5PT sensory-evoked responses: broad and heterogeneous receptive fields.

The receptive fields of L5PTs are broader than those of any other cell type in the same column. The multi-scale model had previously predicted that the circuit mechanism underlying these broad receptive fields is a horizontal dissemination of thalamocortical input by layer 6 cortico-cortical cells (L6CCs). This prediction was successfully tested *in vivo* by pharmacological manipulation (Egger et al., 2020). The reduced models confirmed that L6CC is the primary source of excitation to L5PTs in columns which are not somatotopically aligned to the stimulated whisker, and thereby acts to horizontally spread thalamocortical input, which results in broad receptive fields in L5PTs. However, the origin of the inter-cell variability in receptive field shapes remained unexplored. Using the reduced models, we show that a L5PT's location, and therefore the presynaptic cells it is connected to which determine the synaptic input it receives from the network, is the primary determinant of receptive field shape, and that morphology and biophysics play only a minor role.

3.3 Subcellular scale thalamocortical wiring specificity

L5PTs have been observed to respond to some stimuli with high frequency bursts of two or more APs both *in vitro* (Larkum et al., 1999) and *in vivo* (Larkum and Zhu, 2002). In response to a sensory stimulus *in vivo* in awake rats, an increase in high frequency bursts distinguishes L5PTs from all other excitatory cell types in the column (De Kock et al., 2021). What is the functional relevance of these bursts? Being the main source of output from the cortex, L5PT axons target, among other downstream areas, higher order thalamic nuclei (the posteromedial nucleus (POm) for L5PTs in the barrel cortex, (Rojas-Piloni et al., 2017)).

These corticothalamic synapses are strong, but also strongly depressing (Groh et al., 2008) – this means that individual APs or bursts are filtered out by POm, but a synchronous volley of bursts from L5PTs can overcome this depression and in turn lead to burst firing in POm neurons (Mease et al., 2017). This increase in POm activity can enhance sensory processing in the barrel cortex via feedback thalamocortical excitation (Mease et al., 2016). Despite their apparent importance for sensory processing, the origin of these bursts in cortical output from L5PTs remains unknown.

One theory, proposed following *in vitro* observations in L5PTs, is coincidence detection (Larkum et al., 1999). In slice experiments, L5PTs elicit single AP responses when subjected to a brief current injection at the soma. If this somatic current injection is paired with a slightly delayed current injection at the distal dendrite, which on its own produces no somatic response, the cell instead produces a high frequency burst of APs. This burst is driven by a plateau potential at the apical dendrite, which is mediated by calcium currents. The plateau potential originates from the coincidence of the distal dendritic current injection with the backpropagation of depolarisation from the somatic AP caused by the somatic current injection, which then supralinearly combine and result in a CaAP. Unfortunately, backpropagating APs were shown to have little influence on dendritic calcium potentials *in vivo* (Helmchen et al., 1999; Murayama et al., 2007), and coincidence detection is therefore unlikely to explain bursts in cortical output.

Rather, L5PT burst firing *in vivo* is thought to arise from a local activation of calcium channels (Takahashi et al., 2016). What could be the source of this localised excitatory input to the apical dendrites? The primary thalamus (ventral posterior medial nucleus (VPM) for barrel cortex) is the main source of sensory information to the cortex, and VPM axons form synapses with most cell types across all cortical layers in barrel cortex. The seminal anatomical finding for the present study was that in addition to a high number of synapses on the basal dendrites of L5PTs, VPM axons form a high density of synapses to the region of the apical dendrite which initiates CaAPs (Bast et al., 2021). The CaAP initiation zone lies below the primary bifurcation point (BP) of the apical dendrite – the laminar location of this zone can therefore vary greatly depending on the L5PT's dendritic morphology. Nonetheless,

this high density VPM innervation to the CaAP initiation zone is preserved across L5PTs with very different morphologies. What is the functional relevance of this wiring specificity in the TC-L5PT pathway?

A major challenge with disentangling the relevance of this direct TC input is that L5PTs receive synaptic input from virtually all intracortical cell types (Lefort et al., 2009), all of which respond with similar latencies to sensory stimuli (de Kock et al., 2007). Therefore, it is theoretically possible for all of these intracortical pathways to provide input to L5PTs and contribute to bursting. As these different pathways remain highly challenging to isolate experimentally, the multi-scale model provides sufficient access to disentangle the effects of individual pathways on CaAPs and burst firing in L5PTs.

We find that direct input from VPM is necessary for bursts in sensory-evoked L5PT responses, but not sufficient. Removing all direct TC input to L5PTs abolished bursts, generally leaving single AP responses, whereas removing all intracortical sensory-evoked input and only leaving direct TC input almost fully abolished somatic responses. In general, burst responses arise from a combination of direct sensory input via TC innervation, indirect sensory input from intracortical cell types and non sensory-evoked input prior to the stimulus. We find that bursts with 2 APs (doublets) and bursts with 3 or more APs (triplets) have different origins: while doublets frequently occur without a CaAP and only require TC input to the basal dendrites, triplets are almost always associated with a CaAP and additionally require direct TC input to the distal dendrite. As an exception to this finding, I identify a distinct class of early triplet response, which only rely on direct TC input to the basal dendrites as the associated CaAPs are driven by pre-stimulus input.

4 Methods

4.1 Cortex-inspired ANNs

Author contributions: Rieke Fruengel & Marcel Oberlaender. MO conceived and designed the study. RF developed the approach, performed training and analysed results. RF wrote the paper with feedback from MO. Additional investigations by Stephan Henk and Enru Qiu as thesis projects for a BSc Biology programme. SH investigated the effect of network motif occurrences, and EQ investigated the effect of degree sequences. Both projects were supervised by RF.

Corresponding publication: Fruengel, R. & Oberlaender, M. (under review) 'Sparse connectivity enables efficient information processing in cortex-like artificial neural networks'.

The Introduction, Methods, Results and Discussion sections are in part adapted from the submitted manuscript. Figures 3-8 and Supplementary Materials are taken directly from the submitted manuscript.

All code needed to reproduce the results in this study can be found at:
https://github.com/mpinb/sparseANNs

First, I built an ANN with connectivity directly taken from the C2 column of the multi-scale model (Fig. 1). I extracted a binary connectivity matrix between all intracortical cells in the C2 column (n = 23087). These formed the hidden layer, with one node representing each neuron. There is no convenient way to split these intracortical neurons into layers with exclusively feedforward connectivity, as would be the case for a conventional ANN, due to the high degree of recurrence, so I put all 23087 nodes into a single, recurrent hidden layer. This means that any node could be connected to any other node, allowing the creation of a replica of the connectivity in the cortical column, where only nodes representing neurons which are connected in the cortical column were connected by a trainable weight (other edge weights were set to zero). I initialised nodes representing inhibitory neurons (2869 nodes) with all negative outgoing weights (-0.001 to 0), and nodes representing excitatory neurons (20398

nodes) with all positive outgoing weights (0 to 0.001). Weights were drawn from a uniform distribution. I chose to use the MNIST dataset as a first benchmark training task (Lecun et al., 1998). Each image in the MNIST dataset is 28 pixels wide, and 28 pixels tall. To adapt this image-based dataset into a time series for classification by a recurrent network, I split each image into rows and provided one row to the ANN per time step, meaning each input had a length of 28, and there were 28 time steps (Fig. 1).

These 23087 cortical neurons received sensory input from 350 different neurons in VPM according to the barrel cortex model. To account for connectivity from VPM to the C2 column, I randomly assigned each of the 350 VPM neurons to one of the 28 columns in the image – at each time step, each node in the hidden recurrent layer would receive input only from the columns in the image which corresponded to the VPM neurons it was connected to. In the final time step, output values were collected from all nodes representing L5PTs (1086 nodes), and passed to a linear decoder layer, representing a downstream target of the barrel cortex. There were 10 nodes in the output layer, one for each class in the dataset, and outputs were one-hot encoded (meaning the output layer node representing the target class should have an output of 1, and all other outputs should be zero). I used the ADAM optimiser for gradient descent, with a learning rate of 0.001.

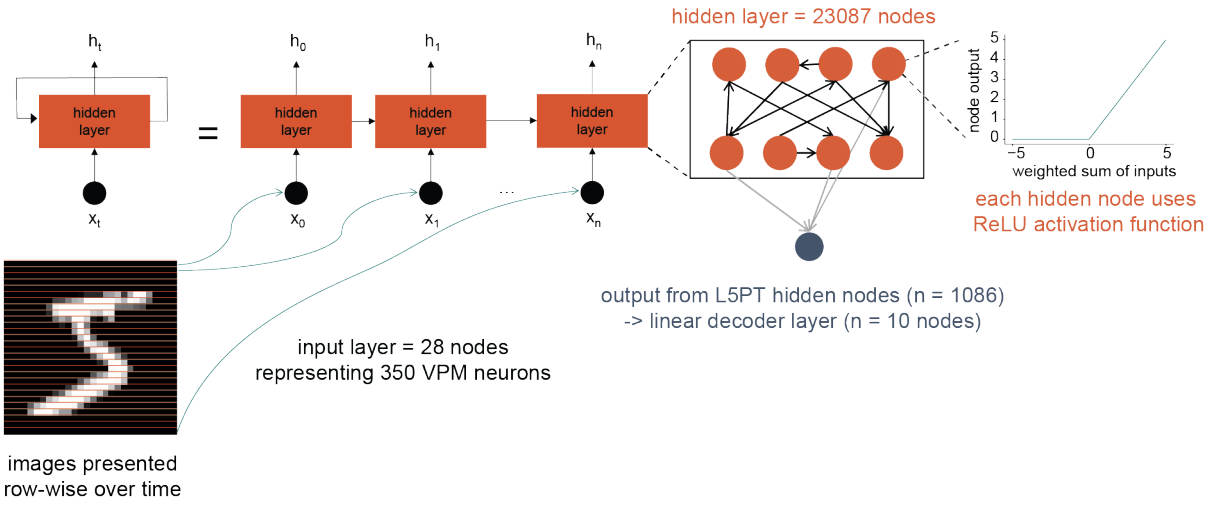

Figure 1. Cortical column ANN performing MNIST image classification.

To determine the relevance of specific connectivity, I shuffled the connectivity matrix between the hidden layer nodes. Everything else, including the identity of excitatory and inhibitory nodes, input and output layer connectivity and training parameters remained the same. Next, I removed all nodes representing neurons from cortical layers 1-4, and as a control for this manipulation, removed an equal number of nodes at random (both manipulations leave 8788 nodes/neurons in the hidden layer). The final manipulation was to decode from other cell types. Instead of connecting L5PTs to the output layer, I connected L5ITs (1575 nodes) or L6CTs (3359 nodes).

I systematically investigated the effect of size, sparse connectivity and recurrence on the training and performance of ANNs (Fig. S1A). I began in recurrent neural networks (RNNs). To this end, I generated a range of ANNs with different numbers of nodes in the hidden layer, and with different connection probabilities between the hidden layer nodes. All nodes were arranged in a single recurrent hidden layer, and any node in the hidden layer could be connected to any other node, with the number of such connections determined randomly by the connection probability. The lower the connection probability the sparser the network, i.e., the fewer nodes were connected to each other by a trainable weight. The weights of edges between hidden layer nodes were randomly initialised from a uniform distribution between -0.001 and 0.001, with weights between unconnected nodes fixed at zero throughout. For each set of hyperparameters (network size & sparsity) I trained 10 networks with different random initialisations, which differ in their connectivity matrix (specifying which nodes are connected by a trainable weight) and in the values of edge weights at initialisation. Results are given as mean ± standard deviation unless otherwise specified. All nodes in these ANNs use the rectified linear unit (ReLU) activation function unless otherwise specified.

I evaluated our networks on three benchmark machine learning tasks, MNIST (Lecun et al., 1998) and CIFAR10 (Krizhevsky, 2012) image classification, and Sleep-EDF (Kemp et al., 2000) sleep-stage classification from EEG recordings. The image-based tasks were modified for recurrent networks to encode each image as a time series by slicing each image row-wise and presenting one row of pixel values (one value per pixel for grayscale images in MNIST,

three values per pixel for RGB images in CIFAR10) in each time step (Fig. S1B). The size of the input layer was equal to the width of the image (MNIST, 28), the width of the image multiplied by 3 to account for RGB colour channels (CIFAR10, 32x3), or one for the single-channel time series in the Sleep-EDF dataset. Each input layer node was connected to all hidden layer nodes by a trainable weight. The output from all the recurrently connected hidden layer nodes in the last time step was then passed to a linear output layer with 5 (Sleep-EDF) or 10 (MNIST & CIFAR10) nodes for classification (one for each class in the dataset, target classes were one-hot encoded). Networks were trained by backpropagation using the ADAM optimiser for gradient descent, with a learning rate of 0.001. I used the cross entropy loss function. All networks were implemented and trained in PyTorch v1.8.1.

To assess the single-epoch learning performance of our networks, I trained all networks for one epoch, meaning that each data point from the training dataset was presented once (60000 data points for MNIST, 50000 data points for CIFAR10, 398370 data points for Sleep-EDF). Then, the accuracy of the networks was tested on the corresponding testing dataset, which consists of 10000 (MNIST & CIFAR10) or 170730 (Sleep-EDF) unseen examples.

To assess whether recurrent network architecture was necessary for sparsity to be beneficial, I repeated the same training of networks with different hyperparameters with feedforward networks. Here, hidden layer nodes were split into two hidden layers. Nodes in the first layer could be connected to nodes in the second hidden layer by a feedforward connection, with the number of such connections determined randomly by the connection probability. To feedforward networks, the whole image was provided at once as input (flattened to a 1D vector) for MNIST and CIFAR10 datasets, and for the Sleep-EDF dataset, the whole EEG sequence was passed at once. The length of the resulting input vector determined the size of the input layer.

To evaluate the effect of a different activation function, I replaced the ReLU activation function in hidden layer nodes with the hyperbolic tangent (tanh) function. All other hyperparameters remained unchanged, and I trained three different random initialisations for each set of hyperparameters with this new activation function.

To assess the performance of our networks with limited training data, I modified the MNIST and CIFAR10 tasks. The full training datasets contain 6000 or 5000 examples of each class for MNIST and CIFAR10 respectively. I selected a smaller random subset of examples (50, 100, 500 or 1000 per class) and used only these for training. The test dataset remained unchanged. ANNs were trained for 50 epochs with early stopping if their performance stopped improving, and we then compared their best test accuracy.

In order to investigate how our networks encode their inputs, I determined how hidden layer nodes contribute to a classification. To this end, I took a sparsely connected network (size 10000 hidden nodes, connection probability 0.1) and a densely connected network (size 10000 hidden nodes, connection probability 1.0). Then, I gave each network the full testing dataset of MNIST (10000 images) as input. After the network had processed each image, I recorded the activation values (outputs) of all hidden layer nodes at the last time step, which is the activation value which is passed to the output layer for classification. I performed these measurements before the networks had received any training, and after one epoch of training. I first consider what proportion of hidden layer nodes send a non-zero activation to the output layer in response to each image. Then, I also consider the magnitude of these non-zero activations. I used sklearn.feature_selection.mutual_info_regression from scikitlearn (Pedregosa et al., 2011) to calculate the mutual information between the responses to the whole testing dataset for 10000 randomly sampled node pairs. To investigate the effect of neuronal noise, I randomly set the output of a proportion of hidden layer node outputs to zero at test time.

I next trained networks which obey Dale's principle (Eccles, 1976), meaning that each node had either exclusively positive or negative outgoing weights, and this sign remained unchanged throughout the training process. I initialised networks with 11.5 % inhibitory nodes, corresponding to the proportion of inhibitory neurons reported in sensory cortex (Meyer et al., 2011). Inhibitory nodes were initialised with random, all negative outgoing weights, and excitatory nodes with random, positive outgoing weights (all weights still had a magnitude randomly chosen according to a uniform distribution between 0 and 0.001).

Whether a node had a trainable connection to another node was purely determined by the connection probability, and was not affected by their excitatory or inhibitory nature. After each weight update, if any weight's sign would be reversed, its value was set to zero instead. All other training parameters were the same as for previous networks (with the exception of sparse to dense networks in Fig. S4, where weights initialised at zero were allowed to become nonzero during training). For Fig. 6F, I initialised networks with 50 % inhibitory nodes to evaluate the effect of structural E/I balance.

Finally, we investigated the effect of higher-order features of cortical connectivity on the training and performance of ANNs. First, we explore the effect of varying the occurrence of recurrent network motifs (Fig. 7A). We began with random connectivity matrices with connection probability 0.14 to match the barrel column, with either 100, 1000 or 10000 nodes. In these matrices, we replaced empty network motifs from randomly sampled triplets with fully recurrent network motifs, increasing the number of recurrent motifs by 300 % and 600 % for networks with 100 and 1000 nodes, and by 50 % and 100 % in networks with 10000 nodes (the difference in the extent of modification between small and large networks is due to the large computational cost of sampling and modifying network motifs in larger networks). To maintain the overall number of edges and therefore connection probability in the network, we compensated for this addition of edges by removing edges from sparse feedforward motifs. To evaluate the effect of a network with only recurrent motifs, we began with an empty connectivity matrix and added fully recurrent motifs until the same overall connection probability was reached as for the other networks. We trained all networks on MNIST for 10 epochs.

We investigated the effect of maintaining biological degree sequences. The degree sequence is the distribution of in- and out-degrees of nodes in a network, the in-degree being the number of incoming connections each node receives, and the out-degree being the number of outgoing connections from each node. First, we calculated the in- and out-degree distributions for one column (C2) of the barrel cortex model. We then sample from this distribution in two different ways: for 'uncorrelated' networks, each node receives an in- and

out-degree randomly sampled from the corresponding distributions. For ‘correlated’ networks, each node receives both the in- and out-degree from a randomly sampled neuron in the cortical column – this way, the correlations between in- and out-degree which are seen in the cortical network (Landau et al., 2016; Udvary et al., 2022) are preserved. To test the effect of including Dale’s principle, we calculate the cortical in- and out-degree distributions separately for excitatory and inhibitory neurons, and then construct a network with excitatory and inhibitory nodes as before, sampling each node’s in- and out-degree from a neuron of the corresponding type. We built ANNs with uncorrelated, correlated, correlated-Dale’s and random (as a control) connectivity matrices with 500, 5000 and 10000 hidden layer nodes, and trained all networks on MNIST and CIFAR10 for 100 epochs.

4.2 Reduced models

Author contributions: Arco Bast*, Rieke Fruengel*, Christiaan P. J. de Kock & Marcel Oberlaender. *These authors contributed equally. AB & MO conceptualised the study and provided supervision. AB developed the model reduction approach (estimation of spatiotemporal filters, nonlinearity and post-spike penalty from simulation data of biophysically detailed models), developed a pipeline to estimate spatiotemporal filters via optimisation of the AUROC score accelerated by GPU computing, generated reduced models to predict APs from synaptic input and validated the predictions with respect to *in vivo* data. RF extended the approach to robustly work with different biophysically detailed models and to robustly estimate the post-spike penalty, improved compatibility with larger datasets, performed generalisability analysis, performed comprehensive evaluation of model performance and applied the pipeline to different biophysical models. AB & RF performed simulations and analysed data. CDK contributed experimental data. AB & RF wrote the paper with MO.

Corresponding publication: Bast*, A., Fruengel*, R., de Kock, C.P.J. & Oberlaender, M. (2024) ‘Network-neuron interactions underlying sensory responses of layer 5 pyramidal tract neurons in barrel cortex’, PLOS Computational Biology. Edited by H. Cuntz, 20(4), p. e1011468. Available at: https://doi.org/10.1371/journal.pcbi.1011468.

The Methods, Results and Discussion sections are in part adapted from the publication. Figures and Supplementary Materials are taken directly from the publication.

Code for reduced model inference is provided here:
https://doi.org/10.5281/zenodo.10730382

Training data for the reduced models (synapse activation times and spike times) and parameters for all reduced models: https://doi.org/10.7910/DVN/QV7JIF

4.2.1 Biophysically detailed models and multi-scale model simulations

The work in this section was performed by AB and is included for completeness.

We selected 5 L5PT dendrite reconstructions that are representative of the morphological variability of this cell type. Multi-compartmental models were generated for these morphologies as described previously (Hay et al., 2011; Egger et al., 2020). Briefly, a multi-objective evolutionary algorithm was used to find parameters for the passive leak conductance and the density of Hodgkin-Huxley type ion channels on soma, basal dendrite, apical dendrite and axon initial segment, such that the neuron model is able to reproduce characteristic electrophysiological responses to somatic and dendritic current injections of L5PTs within the experimentally observed variability, including back-propagating APs, calcium APs, and AP responses to prolonged somatic current injections, using a method adapted from Hay *et al.* (2011) with additions described in Bast and Oberlaender (2021). We incorporated the IBEA algorithm (van Geit et al., 2016) for optimisation. The optimisation was terminated if there was no progress or when acceptable models had been found. We repeated the optimisation process several times. From each independent run, we selected one model for which the maximal deviation from the experimental mean in units of standard deviation across all objectives was minimal (0.9–1.9 mean STDs across objectives).

We embedded the dendrite morphologies selected for biophysically detailed models in the network model of the barrel cortex (Egger et al., 2014) at 81 locations within the cortical barrel column representing the C2 whisker, which is located approximately in the centre of the barrel

cortex. For all *in silico* data presented in this study, the principal whisker (PW) is hence the C2 whisker. The locations were the column centre, and equally spaced in a grid with a distance of 50 µm between adjacent somata. For each of the 81 locations, we extracted the location of presynaptic neurons in VPM and barrel cortex that provide input to the respective L5PT (Fig. 7D) and where along the L5PT dendrite they synapse (Fig. 7E) from the multi-scale model. To generate synaptic activity, we activated all presynaptic cells in the barrel cortex model according to experimental recordings for their cell type and columnar position relative to the chosen whisker stimulus (de Kock et al., 2007; Egger et al., 2020). The peak conductance for synaptic inputs from each excitatory cell type was fitted to *in vitro* and *in vivo* unitary post-synaptic potential amplitudes, as described in (Egger et al., 2020). Depending on the network embedding, L5PTs receive different fractions of excitatory and inhibitory inputs, which may not maintain functional EI balance (Landau et al., 2016). We therefore included a scaling factor for evoked inhibitory synaptic strength, which we constrained to fit empirically observed PW response probabilities (scaling factors ranged from 0.79 to 1.56 depending on the biophysically detailed model).

Simulations were performed using Python 2.7, dask (Dask Development Team, 2016) and NEURON 7.4 (Hines and Carnevale, 1997). We simulated 1000 whisker stimulus trials for the PW and each of the 8 surround whiskers (SWs), for all 7 biophysically detailed models in each of the 81 network embedding locations. The database also contained trials for one 2nd SW (E2), which was however not included into analysis. This results in 810000 simulation trials per biophysically detailed model.

4.2.2 Reduced model inference and simulations

We performed reduced model inference for each biophysically detailed model separately. We split the multi-scale model simulation dataset for each biophysical model into a training and test dataset (split ratio: 70 % to 30 %, respectively). Synapses were binned based on their time point of activation (1 ms bins) and soma distance (50 micron bins). We excluded trials of the recent AP category (in which there was an AP in the last 50 ms) for inference of the spatiotemporal filters. Spatial and temporal filters were constructed as a weighted sum of

basis functions f_i and g_j. The soma distance dependent weight w_z and the time dependent weight w_τ of a synapse are given by

$$w_\tau = \sum_i a_i f_i(\tau)$$

$$w_z = \sum_j b_j g_j(z)$$

where a_i and b_j are free parameters.

We used raised cosine functions (Weber and Pillow, 2016) as basis functions. The temporal basis functions f_i were:

$$f_i(\tau) = \frac{1}{2}\cos(k \cdot log[\tau + c] - \phi_i) + \frac{1}{2}$$

for τ such that k·log(t + c) ϵ [ϕ_i – π, ϕ_i + π] and 0 elsewhere. Values used were k = 3, c = 5, and φ ϵ [3, 12]. Analogously, the spatial basis functions g_i were:

$$g_j(z) = \frac{1}{2}\cos\big(k \cdot log[z + c] - \phi_j\big) + \frac{1}{2}$$

for z such that k·log(z + c) ϵ [ϕ_i – π, ϕ_i + π] and 0 elsewhere. Values used were k = 2, c = 1, and φ ϵ [1, 11].

Using the spatiotemporal filters, we compute a weighted sum of active synapses at a time point t, in the following referred to as weighted net input (WNI). The WNI is computed as follows:

$$WNI(t) = \sum_{\tau,z} n_{z,t-\tau,E} \cdot w_{\tau,E} \cdot\ w_{z,E} + \sum_{\tau,z} n_{z,t-\tau,I} \cdot w_{\tau,I} \cdot\ w_{z,I}$$

where, $\tau \in$ [0, 80 ms] is the time before t, $z \in$ [0, 1300 µm] is the distance of the synapse from the soma, and $n_{z,t-\tau}$ is the number of active synapses at a given 1 ms time and 50 micron soma distance bin. w_τ and w_z are the temporal and spatial filter, respectively. n, w_τ and w_z are split by synapse type, as indicated by the subscripts E (excitatory) and I (inhibitory).

To adjust the free parameters (a_i and b_j for excitatory and inhibitory input) which determine the shape of the spatial and temporal filters, we use a gradient-free optimisation method (COBYLA, implemented in SciPy: (Virtanen et al., 2020)) to maximize the area under the receiver operating characteristic curve (AUROC) between the WNI and AP output for a selected 1 ms time bin. The 1 ms time bin selected for the optimisation is in the following referred to as the inference time point $t_{inference}$. For each biophysical model, we performed this optimization for $t_{inference} \in [0,25]$ ms post whisker stimulus. We then manually selected one inference time point for each biophysical model which resulted in reduced models that generalised well to other time points based on visual inspection of the peri-stimulus time histograms (PSTHs). The selected inference time point of each selected model is visualized in Fig. S4. Finally, we normalized the spatial and temporal filters such that the peak of the temporal filter is 1 for excitatory synapses and -1 for inhibitory synapses, and the value of the spatial filter for excitatory synapses at spatial bin 0 is 1.

In order to calculate the spiking probability for a given WNI value, we used these filters to calculate the WNI for all biophysical model simulation trials from the training dataset at $t_{inference}$, and recorded whether a AP occurred at this time bin or not. The trials were then binned by WNI. The AP probability for each bin corresponds to the proportion of trials that produced an output AP in the biophysically detailed model simulation (Fig. 8C). Bins with few data points were combined to ensure a minimum of ten data points per bin. Linear interpolation was used to find the spiking probability corresponding to any WNI value, and WNI values greater/smaller than values seen in the biophysical model simulation were assigned the highest/lowest spiking probability respectively. We refer to this as the “nonlinearity” function.

In order to estimate the effect of recent APs on spiking probability, WNI values were plotted against the time since the previous AP (Fig. 8D). Then, a boundary was drawn to best separate spiking from non-spiking trials by dropping points with the lowest 5 % of all WNI values (to remove the effect of outliers) and then drawing a boundary along the minimum remaining spiking WNI values. This boundary was normalized such that its value is zero for time to previous AP $\rightarrow \infty$ by subtracting the offset. We found this two-step procedure (first estimate spatiotemporal filters based on a dataset in which recent APs are filtered out, second determine the post AP penalty based on these filters) to be more robust than a joint inference of both. WNI with penalty applied maintains a high AUROC over all stimulus periods (Fig. 8F), while the uncorrected WNI's AUROC drops after the sensory stimulus due to a high proportion of trials in which the biophysical model is refractory.

When using the reduced models to generate spiking output from synaptic input, we first apply the spatiotemporal filters to binned excitatory and inhibitory synaptic input at each prediction time point. The resulting WNI value is looked up in the nonlinearity to determine the spiking probability. According to this probability, a random number generator is used to determine whether an AP is produced at this prediction time point or not. If an AP is produced, the WNI of subsequent time points is penalised according to the post AP penalty.

4.2.3 Analysis

'AP probability' is used to denote the probability that an AP is evoked within a 1 ms time bin. 'Response probability' denotes the probability that one or more APs are generated in the 25 ms following a whisker stimulus. 'Accuracy' is the percentage of trials in which the multi-scale and reduced model agree on whether or not an AP is elicited in a 25 ms window following the whisker stimulus. We used the SciPy function 'metrics.roc_auc_score' to compute the AUROC score, which by default uses the full range of values (i.e. thresholds are automatically chosen such that the false positive rate varies between 0 % and 100 %). Distributions of response probabilities were compared with the two-sided Kolmogorov-Smirnov test using the SciPy function 'scipy.stats.ks_2samp'.

4.3 Thalamocortical wiring specificity

Author contributions: Arco Bast*, Jason M. Guest*, Rieke Fruengel, Rajeevan T. Narayanan, Christiaan P.J. De Kock, Marcel Oberlaender. *These authors contributed equally. JMG, RTN & CDK performed experiments. MO provided supervision and wrote the manuscript. AB developed the biophysically detailed L5PT models, analysis pipelines for experimental and simulation data, a semi-automated method for AP detection with manual proofreading and high performance computing routines for simulation and analysis, performed analysis of electrophysiological and anatomical data to constrain multi-scale models, extracted APs for all anaesthetised experiments and performed cell type-specific and distance-dependent removal manipulation simulations. RF performed a detailed characterisation of CaAP shapes, categorised and investigated non sensory-evoked CaAPs, performed manipulations to test the effect of TC facilitation and increased prestimulus activity, and reanalysed experimental data from awake animals with manual spike proofreading and behaviour categorisation. AB & RF analysed manipulated simulations, performed manipulations to test the effect of decreased prestimulus activity, quantified the voltage in different dendritic compartments and analysed experimental data from airpuff and optogenetics experiments.

Corresponding publication: Bast*, A., Guest*, J.M., Fruengel, R., Narayanan, R.T., de Kock, C.P.J. & Oberlaender, M. (in preparation) ‘The Origins of Burst Firing in Sensory-evoked Cortical Output’.

The Methods section is in part adapted from the manuscript. Figures and Supplementary Materials are adapted from the manuscript where noted in the figure legends.

4.3.1 In vivo experiments

Experiments were performed by JMG, CDK and RN.

Details of experimental materials and methods can be found in Bast *et al.* (2021). In brief, for electrophysiology experiments in anaesthetised rats, the animals received a 700 ms airpuff stimulus to caudally deflect the whiskers at 2.5 second intervals, with 20-30 trials per recording. APs were recorded using an extracellular loose patch amplifier, which was low-

pass filtered (300 Hz) to measure the local field potential (LFP). Trials were assigned to the response types depending on their activity within the first 50 ms after stimulus onset: 3 AP bursts were defined as 3 APs occurring within 30 ms, 2 AP bursts as 2 APs within 10 ms. Trials that fulfilled both criteria were assigned as 3 AP bursts. The same criteria were used for classifying optogenetic responses, active touch responses, and the simulations. Some of the rats had an adeno-associated virus (AAV) expressing channelrhodopsin (ChR2) injected into VPM thalamus, which was later used to view TC axonal boutons to identify TC synapse locations onto L5PT dendrites (combined with immunolabelling with VGlut2). These AAV-injected rats were also used for an optogenetics experiment, with TC terminals expressing ChR2 optically stimulated by 10 ms light pulses to the cortical surface. Stimuli were given at 100 ms, 500 ms and 2.5 second intervals, with 20-100 trials of each. APs and LFP were recorded the same way as for the airpuff stimulus. Following electrophysiological recordings, the neurons were filled with biocytin, and later reconstructed – these reconstructed L5PT dendrite morphologies were used to generate biophysically detailed models for this study.

For electrophysiology experiments in awake animals, rats were habituated to head fixation before the recording experiment. On the recording day, rats were anaesthetised and targeted loose-patch recordings were made using intrinsic optical imaging. Passive whisker stimulation and receptive field mapping were used to confirm intrinsic optical imaging results. The whiskers were then clipped apart from the PW or a single SW, and anaesthesia was terminated. The rats were head fixed, and an object was positioned 2 cm lateral from the whisker pad and anterior relative to the whisker set point, to ensure that touches were the result of whisker protraction. Touch events and whisker angle were determined manually by video recording. Following electrophysiological recordings, the neurons were filled with biocytin, and later reconstructed.

4.3.2 Multi-scale model simulations

The work in this section was performed by AB.

Simulations were performed using Python 3.8, dask (Dask Development Team, 2016) and NEURON 7.8 (Hines and Carnevale, 1997). First, we determined L5PT simulation

configurations that result in simulations with AP rates that are within the range observed across L5PTs during periods preceding passive whisker stimuli. For the 612 network-embedded L5PT models (i.e., 68 multi-compartmental models for 3 morphologies and 9 embeddings into the network model, respectively) we simulated prestimulus activity by activating neurons in the network model that are presynaptic to the L5PTs with the spontaneous firing rates that we recorded in anesthetised animals across layers and for different excitatory and inhibitory types, and in VPM thalamus. We distributed 5000 inhibitory synapses uniformly across the dendritic tree, and activated them with different multiples of the *in vivo* observed firing rates (i.e., 0.25 to 3.75 times the *in vivo* rates, in 0.25 steps). For each of these 9180 L5PT simulation configurations (i.e., 68 L5PT models * 9 network embeddings * 15 EXC/INH ratios), we simulated 48 trials of 1245 ms duration, of which we discarded the first 245 ms as initialisation of the simulations. 2241 of the 9180 configurations predicted ongoing AP rates as observed for L5PTs *in vivo* (i.e., >0 Hz and ≤11 Hz). Second, we determined which of these 2241 L5PT simulation configurations result also in simulations with 1 AP, 2 AP and 3 AP burst rates that are within the respective ranges observed across L5PTs during the onset response evoked by passive whisker stimuli. For this purpose, we simulated 445 ms of prestimulus activity as described above, and then activated neurons in the network model that are presynaptic to the L5PT models by generating Poisson spike trains for each presynaptic neuron based on the empirically measured PSTHs of the respective types. We systematically tested whether our simulation results are robust with respect to different stimuli, different timings and ratios of sensory-evoked inhibition relative to excitation. Stimuli: principal whisker, three whiskers within the same row, or three whiskers within two adjacent rows, respectively. Timings: latencies as observed *in vivo*, inhibition shifted by 1 or 2 ms to earlier time points. Ratios: we activated inhibitory neurons with different multiples of the *in vivo* observed firing rates (i.e., 0.1 to 1.0 times the *in vivo* PSTHs, in 0.1 steps). For each of these 161352 L5PT simulation configurations (i.e., 2241 models with realistic prestimulus rates * 3 stimuli * 3 EXC/INH timings * 8 EXC/INH ratios), we simulated 20 trials to select all configurations that predict sensory-evoked responses within the *in vivo* observed ranges (i.e., ≥1 AP probability: >0 %, 2 AP burst probability >0 % and ≤40 %, 3 AP burst probability ≥0 % and ≤25 %). 67424 of the 161352 configurations predicted pre- and

post-stimulus activity as we had observed for L5PTs *in vivo*, out of which 22850 configurations contained sensory-evoked bursts. Third, we selected these 22850 L5PT simulation configurations and repeated the simulations, this time 200 instead of 20 trials per model configuration. Moreover, in the first coarse selection step, we had accepted 3 AP burst probabilities up to 25 %. Now, we only selected those L5PT simulation configurations with 3 AP burst probabilities up to 13 %, which is the maximal value observed in anesthetised animals. The simulations hence identified 20359 L5PT simulation configurations that predict pre- and post-stimulus activity, including fast sensory-evoked bursts with 2 and 3 APs, consistent with the *in vivo* data. These comprised configurations for all three morphologies (most superficial BP: 12589, middle BP: 3298, deepest BP: 4472 models), all nine network locations, all 68 biophysical parameter sets, all 3 whisker stimuli, all inhibitory timings, 7/8 post-stimulus EXC/INH ratios, and 13/15 pre-stimulus EXC/INH ratios.

4.3.3 Multi-scale model manipulations

Next, we used these 20359 L5PT simulation configurations to test the effect of various manipulations. To this end, we replayed the control simulations, but modified the synaptic input in various ways. For the cell type removal manipulations, we removed the sensory-evoked synaptic input from either all intracortical excitatory cell types, VPM, layer 2/3 pyramidal neurons (L2/3PNs), spiny neurons in layer 4 (L4SPs), layer 5 intratelencephalic neurons (L5ITs), L5PTs, L6CCs, or VPM and L6CCs. All subsequent analyses were performed for models from the morphology with the most superficial BP. To investigate distance-dependent impacts of TC synapse distributions, we selected models with triplet probability >= 2 %, and simulated 1000 trials for each of them. We replayed the simulations 20 times, each time removing 50 TC synapse activations in 200 µm intervals with increasing soma distance ranging from 0-200 to 950-1150 µm. To investigate the effect of decreased ongoing activity, we selected all simulation trials with a triplet response in the control scenario and then removed 50 synapse activations from intracortical excitatory cell types in the 20 ms window before TC onset, from different dendritic compartments (basal, trunk, apical tuft). If in a trial there were less than 50 synapse activations, all active synapses were removed. To investigate the effect of increased ongoing activity, we selected all L5PT simulation

configurations with triplet probability ≥1 %, and simulated 200 trials for each where the *in vivo* recorded ongoing firing rates of intracortical excitatory cell types were scaled by a factor from 0.5 to 2 throughout the simulation. We scaled the firing rates of either superficial (L2-4), deep (L5-6), or all (L2-6) types. As a proof of principle for how increased ongoing activity could increase the rate of triplets, we show results from models that had a triplet rate of at least 35 % with a scaling factor of 1.7 applied to all types. Finally, plasticity has been reported for VPM synapses to some pyramidal neurons in the deep layers of barrel cortex, with ~1/3 of the synapses showing facilitation on the second stimulus in a series (synapse class 1C, described by Viaene, Petrof and Sherman (2011)). To investigate whether synaptic facilitation in the TC-L5PT pathway could account for the observed transition from singlets to bursts in our optogenetics experiment, we repeated all simulations with an increased synaptic weight for all TC synapses (2.5 times the control value, as reported in Viaene, Petrof and Sherman (2011) for the second stimulus).

4.3.4 Categorising sensory- vs non sensory-evoked CaAPs

We characterised the onset time (time of crossing -30 mV), peak voltage and width of CaAPs (time between up and down crossing of -30 mV) in trials which are abolished (no longer cross the -30 mV threshold) by removing intracortical sensory-evoked input (i.e., sensory-evoked CaAPs) compared to those which remain (i.e., non sensory-evoked CaAPs). Comparing the properties of these two categories of CaAPs, we defined sensory-evoked CaAPs as having peak voltages >-20 mV, widths >8.9 ms and onset times >6 ms after TC onset. In contrast, non sensory-evoked CaAPs had peak voltages >-15 mV, widths >4 ms and an onset time before TC onset. We filtered out trials with non sensory-evoked CaAPs from our analyses up to section 5.3.3.

5 Results

5.1 Sparse connectivity enables efficient information processing in cortex-like neural networks

5.1.1 Training and manipulating cortical column ANNs

A first attempt at investigating the relevance of cortical structural properties for information processing was made by making an ANN model of a cortical column. I took the connectome of one column (C2) in the multi-scale model and extracted a binary cell-to-cell connectivity matrix. I constructed an ANN where each node represented one neuron from the cortical column (23087 nodes/neurons). Nodes representing inhibitory neurons were initialised with all negative outgoing weights, and nodes representing excitatory neurons were initialised with all positive outgoing weights. Only nodes representing neurons which were synaptically connected in the multi-scale model were connected by a trainable weight, resulting in an average connection probability of 0.14 between nodes. There is no way to split neurons from the cortical column into distinct layers with strictly feedforward connectivity between them due to the high number of recurrent and feedback connections. Therefore, the architecture of the ANN had a single recurrent hidden layer which housed all nodes from the cortical column, an input layer which represented sensory input from VPM (the primary thalamus of the whisker system), and an output layer which performed the classification based on outputs of nodes representing L5PTs (see Fig. 1 in Methods section). I trained such a “cortical column” network on the MNIST handwritten digit recognition task (Lecun et al., 1998), a popular benchmark machine learning task. This “cortical column” network successfully learned to classify MNIST in five epochs to 97.4 % accuracy. Having established that a cortical column network can be trained to perform classification tasks, the challenge was now to disentangle which features of cortical connectivity were relevant for the performance of the ANN.

To determine whether the specific connectivity from the barrel cortex was relevant for function, I shuffled the connectivity matrix from the barrel cortex while keeping all other parameters the same. This results in a connectivity matrix with the same total number of

synaptic connections, but otherwise random connectivity. Interestingly, out of 5 ANNs with randomly shuffled connectivity matrices, 3 trained successfully, but were slightly slower than the true cortical column ANN, and 2 failed to train at all in the allotted training time (Fig. 2A – note that these plots show the test loss, not accuracy, where a lower loss means better performance). I found that the gradients associated with weights in the hidden layer, and therefore also the changes in weights, were vanishingly small in the two networks which did not train (Fig. 2B). At the time of training these networks, I did not identify the reason for this failure to learn. In retrospect, it is plausible that they fail for the same reason as the Dale's principle networks which I train later: an excitation-inhibition imbalance caused by the comparatively small fraction of inhibitory nodes, which in some random initialisations may be better mitigated by connectivity than in others.

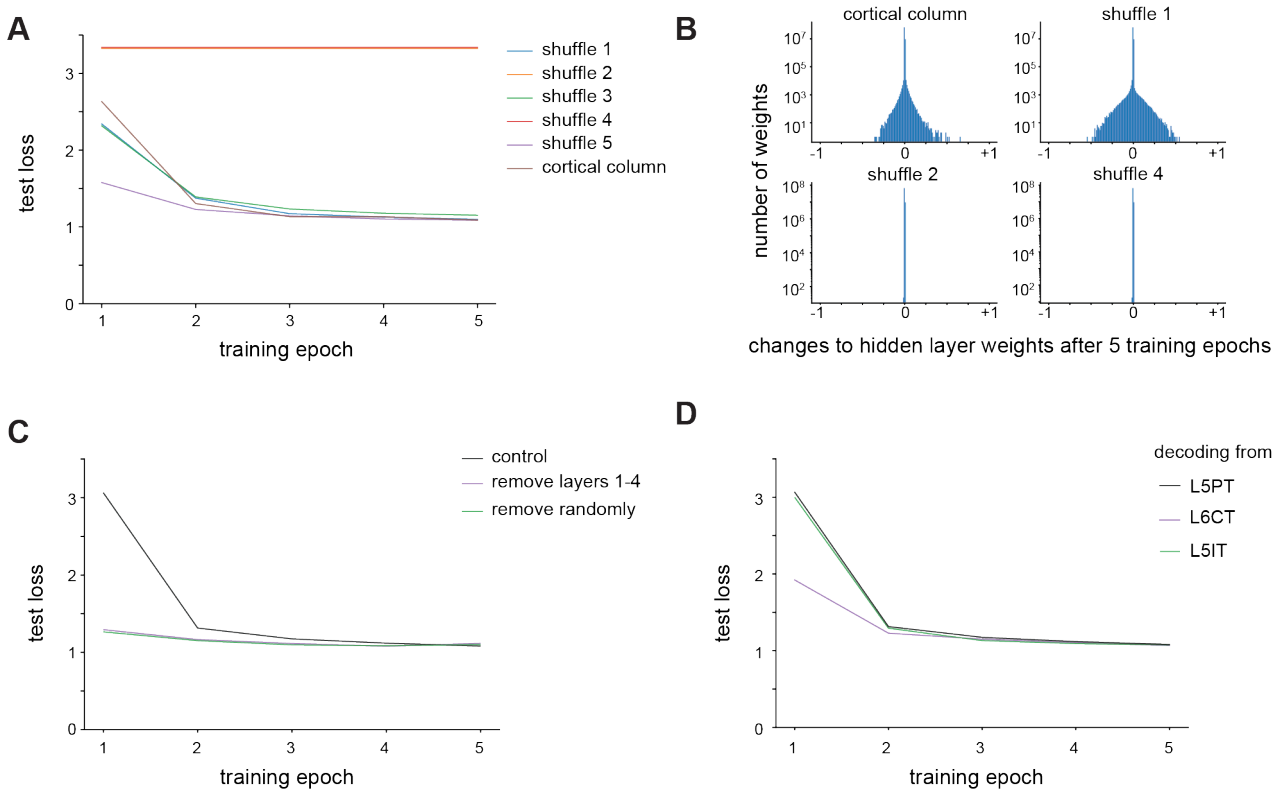

Figure 2. Manipulations of cortical column ANNs. A: Test loss comparing shuffled connectivity matrix to control cortical column ANN. B: Hidden layer weight changes after 5 epochs of training, on control cortical column and shuffled ANNs. C: Test loss comparing control cortical column ANN to an ANN with nodes representing neurons in cortical layers 1-

4 removed, or the equivalent number of nodes randomly removed. D: Test loss comparing decoding from L5PTs to decoding from L5Its or L6CTs.

I hypothesised that perhaps specific subnetworks were important for information processing in the cortical column. In fact, it had been reported previously that densely connected ANNs contain sparse subnetworks which, when trained from scratch, achieve comparable performance to the full dense network with shorter training time (Frankle and Carbin, 2019). Perhaps such particularly important subnetworks also existed in the cortical column. As a first, drastic manipulation, I removed all cells from layers 1-4, leaving only layers 5-6 in the ANN (8788 nodes/neurons) and their respective input and output layer connectivity. I expected that such a network would be inferior at learning the task. Surprisingly, the ANNs with only L5-6 neurons performed better than the whole cortical column ANN (Fig. 2C). I noted that removing nodes by layer was changing the patterns of connectivity within the ANN, but also its size and overall connection probability. As a control, I randomly removed the same number of nodes as are in L1-4 from across all cortical layers, and repeated the training process – this had the same effect (Fig. 2C), suggesting that the change in performance was due to a change in size or perhaps connection probability of the network, and not due to the loss of some specific connectivity circuits.

Instead of changing the connectivity in the hidden layer, I next investigated whether perhaps information is distributed in the cortical column ANN such that L5PTs are the optimal output cell type. To this end, I modified only the output layer connectivity: instead of passing outputs from L5PTs to the linear decoder layer, I took outputs from all layer 5 intratelencephalic neurons (L5IT, 1575 nodes) or all layer 6 corticothalamic neurons (L6CT, 3359 nodes), both cell types which are close to L5PTs in terms of their cortical depth, but have rather different connectivity. Surprisingly, decoding from L5ITs resulted in equivalent performance to decoding from L5PTs, while decoding from L6CTs even improved performance (Fig. 2D).

It became clear that performing specific manipulations on the cortical column ANN's connectivity would not be a practical approach to isolating how structural features affect a network's performance.

5.1.2 Sparse connectivity enables time- and data- efficient training of large and recurrent networks

To systematically investigate the effect of network size, sparsity and recurrence on training, I generated ANNs with random connectivity and different connectivity parameters which made them more or less cortex-like (Fig. 3). I began by training RNNs with different numbers of hidden layer nodes and different connection probabilities between these nodes on MNIST (Fig. 4A). The lower the connection probability, the fewer nodes were connected by a trainable weight, and therefore the sparser the network. The input and output layers were fully connected. To avoid interpreting artefacts from a particular initialisation, I repeated each set of network connectivity parameters 10 times, using different random weight initialisations.

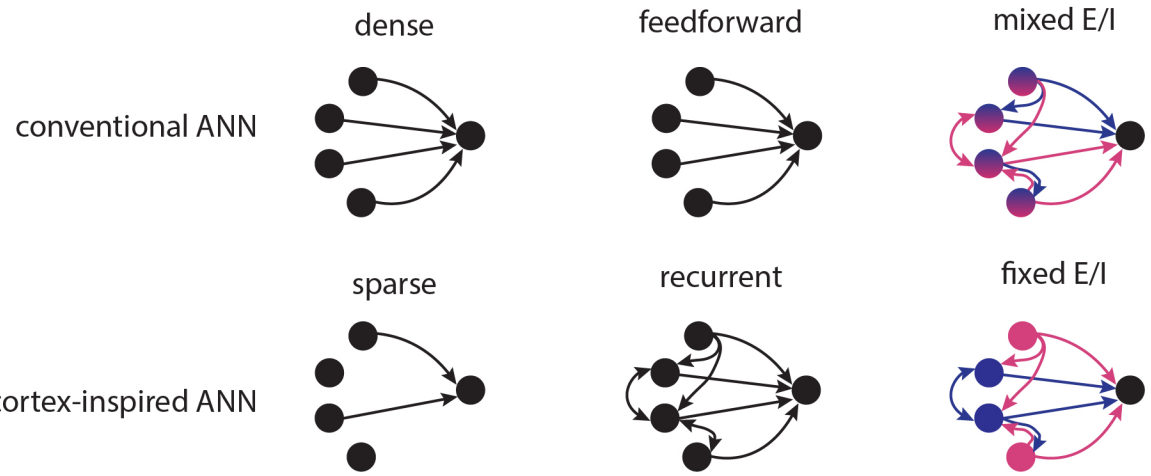

Figure 3. Schematic illustration depicting ANNs with conventional or cortex-inspired structural properties. Top row, left to right: conventional ANNs are densely connected, feedforward and have nodes with mixed excitatory and inhibitory weights. Bottom row, left to right: cortex-inspired ANNs are sparsely connected, have recurrent connectivity and have nodes with fixed excitatory or inhibitory weights.

My first observation was that networks with different connectivity parameters differ in their performance most notably during the early stages of training (Fig. 4B). To compare the performance of networks with limited training time, I therefore evaluated the single epoch accuracy, i.e., the performance on the testing dataset after the network has seen each example in the training dataset only once. I find that sparse connectivity facilitates time-limited

learning in large and recurrently connected networks, but confers no benefit and is rather detrimental in small networks (Fig. 4C). After just a single epoch of training, large sparse networks (with 20,000 nodes and a connection probability of 10 %) attained a mean test accuracy of 90.3±2.0 % while their densely connected counterparts only reached 77.5±7.3 % (for all values see Table S1). This disparity in performance is abolished by further training, with sparse and dense networks reaching comparable test accuracies after around 80 training epochs ($p > 0.05$, KS-test).

Next, I assessed whether these findings generalise to another, more challenging benchmark image recognition dataset, CIFAR10, which uses full-colour RGB images (Krizhevsky, 2012). Similarly to results on MNIST, large sparse networks outperform their dense counterparts on the CIFAR10 dataset as well (Fig. 4D). After a single training epoch, large sparse networks attained a mean test accuracy of 29.7±1.5 %, while large densely connected networks only reached 17.4±3.9 % (for all values see Table S2). On the CIFAR10 dataset, while large sparse networks initially outperform their dense counterparts and attain a higher maximum accuracy, they begin to overfit after around 40 epochs, and their test performance eventually drops below that of dense networks.

To confirm that these findings are not due to some artefact related to performing image classification with recurrent networks, I evaluated the networks on a native time series dataset, Sleep-EDF sleep-stage classification from EEG recordings (Kemp et al., 2000). Once again, I find that large sparse networks outperform their dense counterparts, reaching 50.7±1.8 % accuracy after one training epoch while dense networks of the same size only attained 10.3±0.0 % accuracy (Fig. 4E, for all values see Table S3). I therefore conclude that sparse connectivity facilitates training when training time is limited in large and recurrent networks, and that this effect is not dataset-specific.

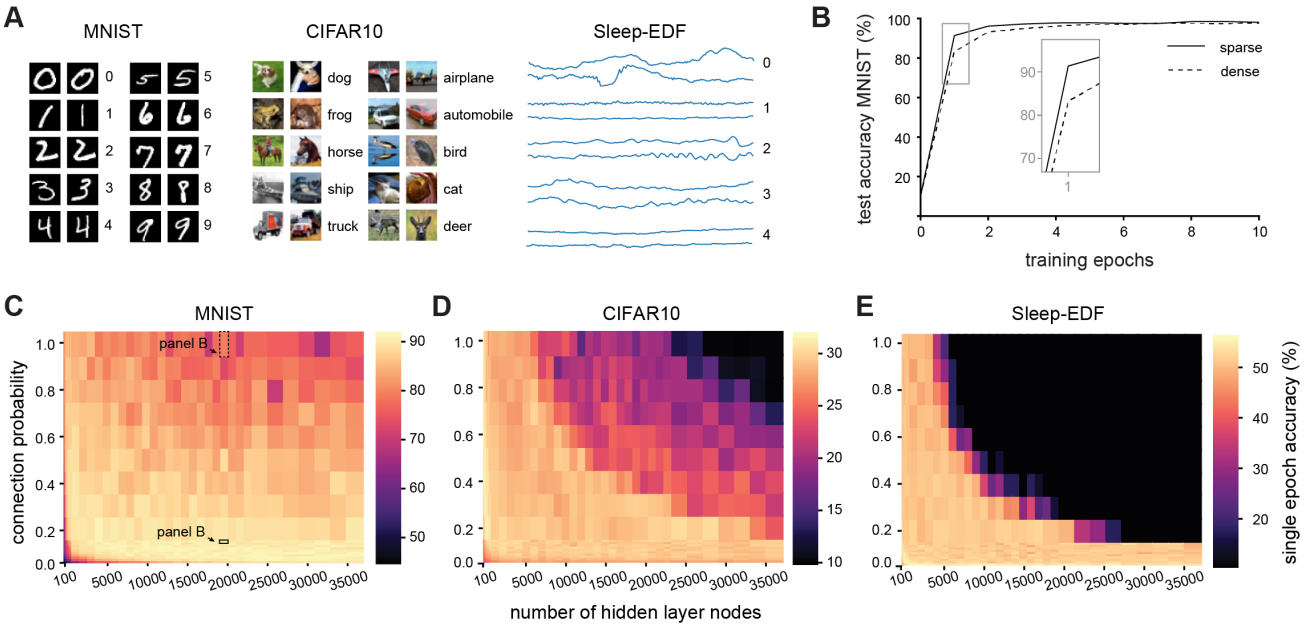

Figure 4. Effect of network size and sparsity on time-limited training. A: Examples from each of the three classification datasets used in this study. B: Test accuracy over 10 training epochs of a densely (connection probability = 1) and sparsely (connection probability = 0.1) connected network with 20000 hidden layer nodes on the MNIST dataset. C: Test accuracy after one training epoch on the MNIST dataset for networks with different numbers of hidden layer nodes and connection probabilities between nodes in the hidden layer (mean over 10 initialisations). D: Test accuracy after one training epoch on the CIFAR10 dataset for networks with different numbers of hidden layer nodes and connection probabilities between nodes in the hidden layer (mean over 10 initialisations). E: Test accuracy after one training epoch on the Sleep-EDF dataset for networks with different numbers of hidden layer nodes and connection probabilities between nodes in the hidden layer (mean over 10 initialisations).

I next investigated how this finding depends on other network properties. I find that the advantage conferred by sparse connectivity is limited to networks with a recurrent architecture (Fig. 5A). In feedforward networks, sparse connectivity has very little effect on the performance of large networks, and is typically detrimental in small networks (Fig. S2). I had so far constructed networks using the ReLU activation function (see Fig. S1B), with which nodes require a threshold of net excitation received before any output is produced. This is reminiscent of neuronal integration, whereby neurons require net excitatory synaptic input to surpass a certain threshold before an action potential response is elicited, and otherwise produce no output. I therefore next tested whether our finding still holds when using an

activation function whose properties are very different from neuronal integration, e.g., the symmetrical hyperbolic tangent function, which can produce positive and negative outputs. I find that with this less biologically realistic activation function, sparse connectivity no longer improves learning efficiency in large networks (Fig. 5B). The tanh activation function has been reported to cause vanishing gradients (Ven and Lederer, 2021). To investigate whether this may explain my results, I repeated training and recorded the gradients during the first epoch of training in large sparse networks (10000 hidden layer nodes, connection probability 0.1). Surprisingly, networks with the tanh activation function have slightly larger gradients than those with the ReLU activation function (mean gradient magnitude 2.0×10^{-5} for tanh network and 6.0×10^{-6} for ReLU network), indicating that the failure of large networks using tanh is not simply due to vanishing gradients.

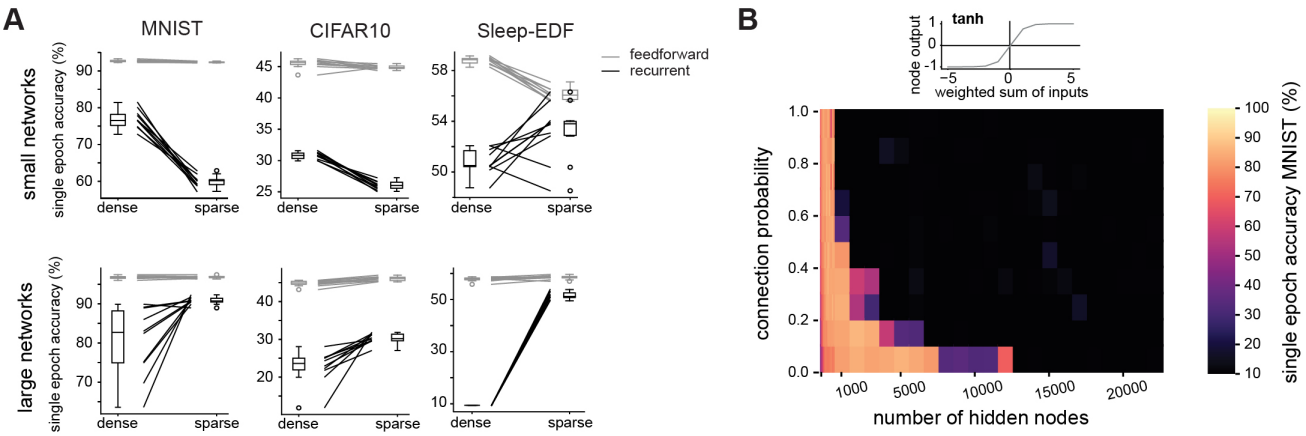

Figure 5. The benefit of sparse connectivity is dependent on a recurrent network architecture and a single-cell activation function with a threshold. A: Comparison of test accuracy between densely (connection probability = 1) and sparsely (connection probability = 0.1) connected networks in small networks (1000 hidden layer nodes, top row) and large networks (10000 hidden layer nodes, bottom row). Nodes were either placed in a single hidden layer with recurrent connectivity (black lines) or in two hidden layers connected by feedforward connections (grey lines). B: Test accuracy after one training epoch on the MNIST dataset for networks with different numbers of hidden layer nodes and connection probabilities between nodes in the hidden layer when nodes use the hyperbolic tangent activation function instead of ReLU.

Having established that sparse connectivity can improve network performance in large and recurrent networks when training time is limited, I next investigated whether the same is true

with data limitations. To this end, I trained networks on a reduced training set from the MNIST dataset, where only a subset of samples from each class was used during training (50, 100, 500 or 1000 samples per class – results for 100, 500 and 1000 samples per class, as well as for the full dataset are found in Fig. S3). The test dataset remained unaltered. Training time was not a constraint in these experiments, so all networks were trained for 50 epochs and I recorded the best test accuracy attained during this training period. I find that in large (20000 hidden nodes) recurrent networks, sparse connectivity becomes more advantageous the more restricted training data are: while on the full training dataset of 60000 examples, sparse and dense networks attain comparable maximum test accuracies (99.0 ± 0.1 % vs 98.5 ± 0.1 %), their performance diverges the more restricted the training dataset is, with respective test accuracies of 70.1 ± 10.2 % and 40.3 ± 7.2 % when trained on only 500 examples (50 per class, Fig. 6A). Meanwhile, in small networks (500 hidden nodes), I again see that sparse connectivity is detrimental to the training process, with test accuracies only reaching 57.0 ±1.3 % in sparse networks but 66.0±4.0 % in dense networks when trained on the most restricted training dataset. Repeating this experiment on the CIFAR10 dataset, large sparse networks outperform their dense counterparts on all subsets of the training data, with test accuracies of 54.9 ± 0.1 % vs 42.1 ± 1.8 % on the full training dataset, and 24.0 ± 0.1 % vs 17.6 ± 1.5 % on the most restricted dataset (Fig. 6B). My findings therefore suggest that in large and recurrent networks, as found in the cortex, sparse connectivity enables the network to learn efficiently in terms of both training time and training data.

5.1.3 Sparsely connected networks form distributed, robust representations

To determine how inputs are represented by networks with different connectivity properties, I recorded the activation values of nodes in the hidden layer when presented with all examples from the test dataset. To assess whether activity is sparse or dense, I first recorded whether a node produces a zero or non-zero activation in the final time step, which is passed to the output layer (Fig. 7A). I find that in large networks (10000 hidden nodes), already after a single training epoch inputs are sparsely represented by the hidden layer activity in both sparsely and densely connected networks, with a majority of hidden layer nodes sending zero-activations to the output layer in response to any given image (Fig. 7B). Surprisingly, I

find that in sparsely connected networks more nodes contribute non-zero outputs to the classification than in densely connected networks (11.2 ±1.5 % vs 3.0 ±1.0 % after one training epoch). When looking at the activation values, not just whether they are zero or non-zero, I note that the average magnitude of hidden layer activations is larger in densely connected networks and smaller in sparsely connected networks (Fig. 7C). I calculated the mutual information between the outputs of 10000 randomly chosen pairs of hidden layer nodes for a sparse and a dense network. I find that the mutual information is significantly higher between nodes in the sparsely connected network (mean 0.019±0.065) compared to densely connected networks (mean 0.005±0.029, KS-test $p = 1.03 \times 10^{-44}$). A larger proportion of nodes contributing to the classification with smaller activations suggests that large sparse networks employ a more distributed, consensus-based coding strategy than dense networks.

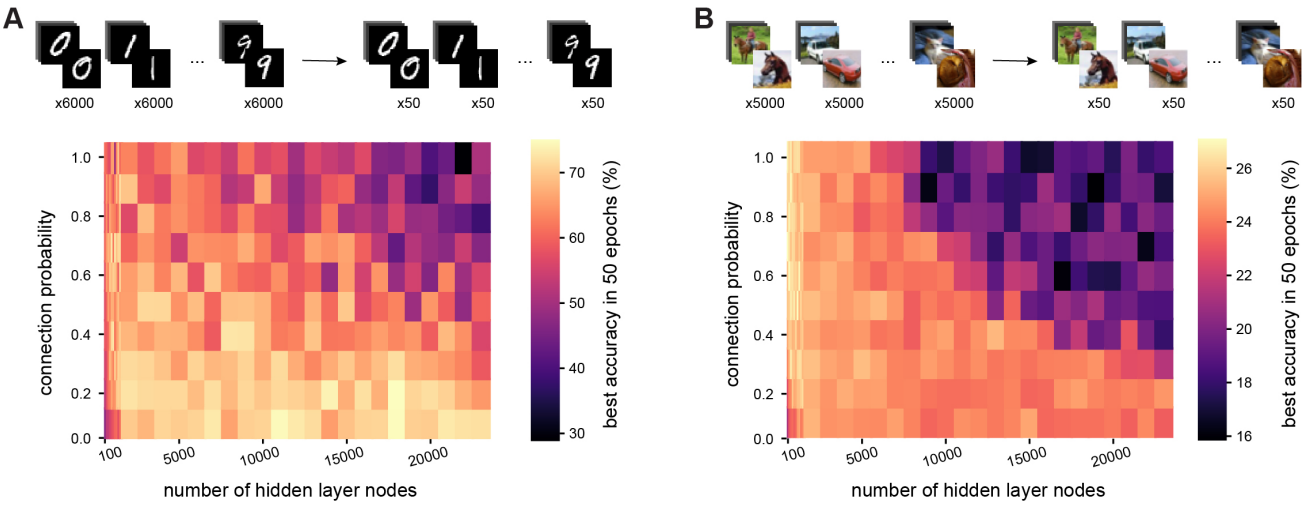

Figure 6. Sparse connectivity enables data-efficient training of large and recurrent networks. A: Test accuracy after one training epoch on a reduced version of the MNIST dataset (50 examples per class instead of 6000) for networks with different numbers of hidden layer nodes and connection probabilities between nodes in the hidden layer. B: Test accuracy after one training epoch on a reduced version of the CIFAR10 dataset (50 examples per class instead of 5000) for networks with different numbers of hidden layer nodes and connection probabilities between nodes in the hidden layer.

Such a distributed code may have advantages in terms of robustness to noise, as the output of any individual node may be less important for the final classification. To test this, I set the output to the classification layer of a proportion of randomly selected nodes to zero for each image at test time and assessed the networks' performance (Fig. 7D). Indeed, I find that sparsely connected networks outperform their dense counterparts at all noise levels (Fig. 7E). For instance, when 50 % of all nodes' outputs are set to zero after one epoch of training, the test performance of sparse networks is barely affected (from 91.1 % to 90.9 %), while the performance of dense networks suffers more (from 88.1 % to 76.3 %). This discrepancy is even more noticeable at higher dropout levels, with sparse networks' performance only dropping by 3.1 %, while dense networks lose 33.3 % in accuracy when 90 % of nodes are zeroed out. With more training, dense networks gradually become more robust to low levels of dropout noise, but even after 10 training epochs they are still outperformed by sparse networks at high dropout levels.

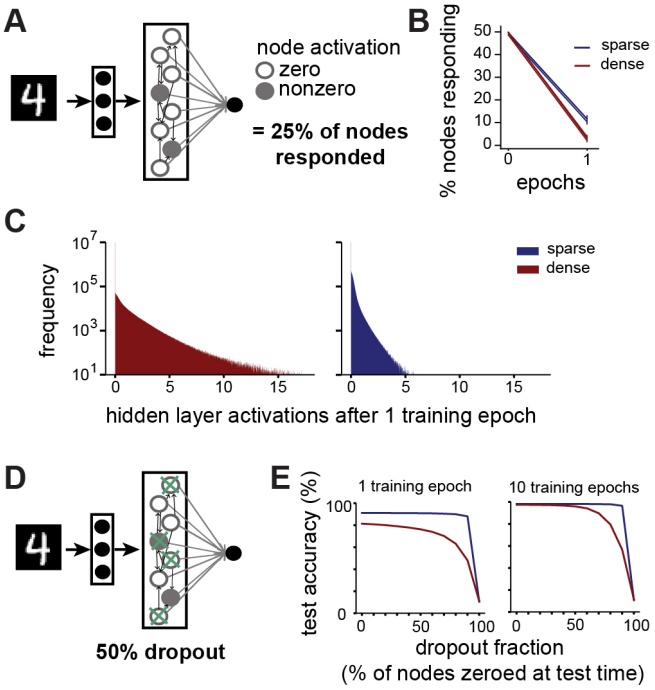

Figure 7. Sparsely connected networks form distributed, robust representations. A: I recorded how many nodes in the hidden layer send zero and nonzero activations to the output layer in response to all images from the MNIST test dataset (N = 10000). B: Change in percentage of nodes sending a nonzero activation to the output layer in the first training epoch

on MNIST. C: Distribution of activation values sent from hidden layer nodes to the output layer by large networks (10000 hidden layer nodes) with dense (connection probability = 1, red) or sparse (connection probability = 0.1, blue) connectivity in response to all images from the MNIST test dataset after one training epoch. D: I performed a dropout experiment where the output from a randomly selected fraction of nodes in the hidden layer was set to zero at test time. E: Effect of different dropout fractions on test accuracy in large networks (10000 hidden layer nodes) with dense (connection probability = 1, red) or sparse (connection probability = 0.1, blue) connectivity after 1 training epoch (left) and 10 training epochs (right).

5.1.4 Sparse connectivity facilitates efficient training in networks with fixed excitatory and inhibitory nodes

How do these sparse representations form during training? When recording the changes in hidden layer weights during training, I observe a greater tendency for negative weights to increase their magnitude, and positive weights to change their sign and become negative, than vice-versa (Fig. S4A). This leads to a sparse representation as the weighted sum of inputs to any given node is more likely to be negative, resulting in an output of zero after the ReLU activation function is applied. This extent of sign reversal of weights is biologically unlikely, as each neuron generally transmits the same set of neurotransmitters to all of its post-synaptic partners, and cannot change this for individual connections (Dale's principle). Therefore, I constructed and trained networks which obey Dale's principle, with fixed excitatory and inhibitory nodes, reminiscent of biological neuronal cell types (Fig. 8A). I chose the proportion of inhibitory nodes which has been reported experimentally for somatosensory cortex (Meyer et al., 2011) (11.5 % inhibitory, 88.5 % excitatory). Inhibitory nodes were initialised with random, all negative outgoing weights, and excitatory nodes with random, positive outgoing weights. During training, if any weight's sign would be reversed, its value was set to zero instead.

In line with reports from others, I note that the training of large (10000 nodes), densely connected networks is indeed severely impeded by the constraints of Dale's principle (Fig. 8B). I observe a delay of tens of training epochs at the start of the training process, during which the network's performance remains around chance level. However, the sparser the

connectivity, the shorter the delay before a network's performance shows improvement. The length of this delay in training of densely connected networks increases with the size of the network, and no delay was observed at any tested network size for sparsely connected networks with a connection probability of 0.1 (Fig. 8C).

I investigated why sparse networks outperform their dense counterparts under these conditions. When examining the activations sent from the hidden layer to the output layer, I find that at initialisation, activations from the hidden layer nodes of densely connected networks are highly correlated with each other, whereas those from sparsely connected networks are less correlated (Fig. 8D). This implies that in densely connected networks, all nodes initially produce very similar outputs, which would make it more challenging to develop meaningful, distinguishable outputs through training.

The gradient of a weight represents the rate of change of the loss function with respect to that weight, providing a measure of how modifying the weight will influence the network's overall error. Therefore, recording the gradients during training allows us to observe the shape of the error landscape. I find that during early training epochs (with the exception of the first epoch), densely connected networks have very small gradients associated with their weights, while the gradients in sparsely connected networks are much larger (Fig. 8E). This suggests that densely connected networks, unlike their sparse counterparts, become stuck on a plateau in the error landscape, i.e., a region with a high error and small gradients. This plateau is difficult to leave via gradient descent, and therefore causes a delay in learning.

To verify that this is indeed a feature of the weight initialisation, and not some other aspect of training, I initialised a network with a sparse weight matrix (i.e., with 90 % of weights starting at zero), but then allowed all weights to be modified during training regardless of their starting value, as would be the case for a dense network. These "sparse-to-dense" networks all start training without a delay like standard sparse networks, but their learning rate soon slows and they take longer to reach their peak performance, similar to networks with more dense connectivity (Fig. S4B).

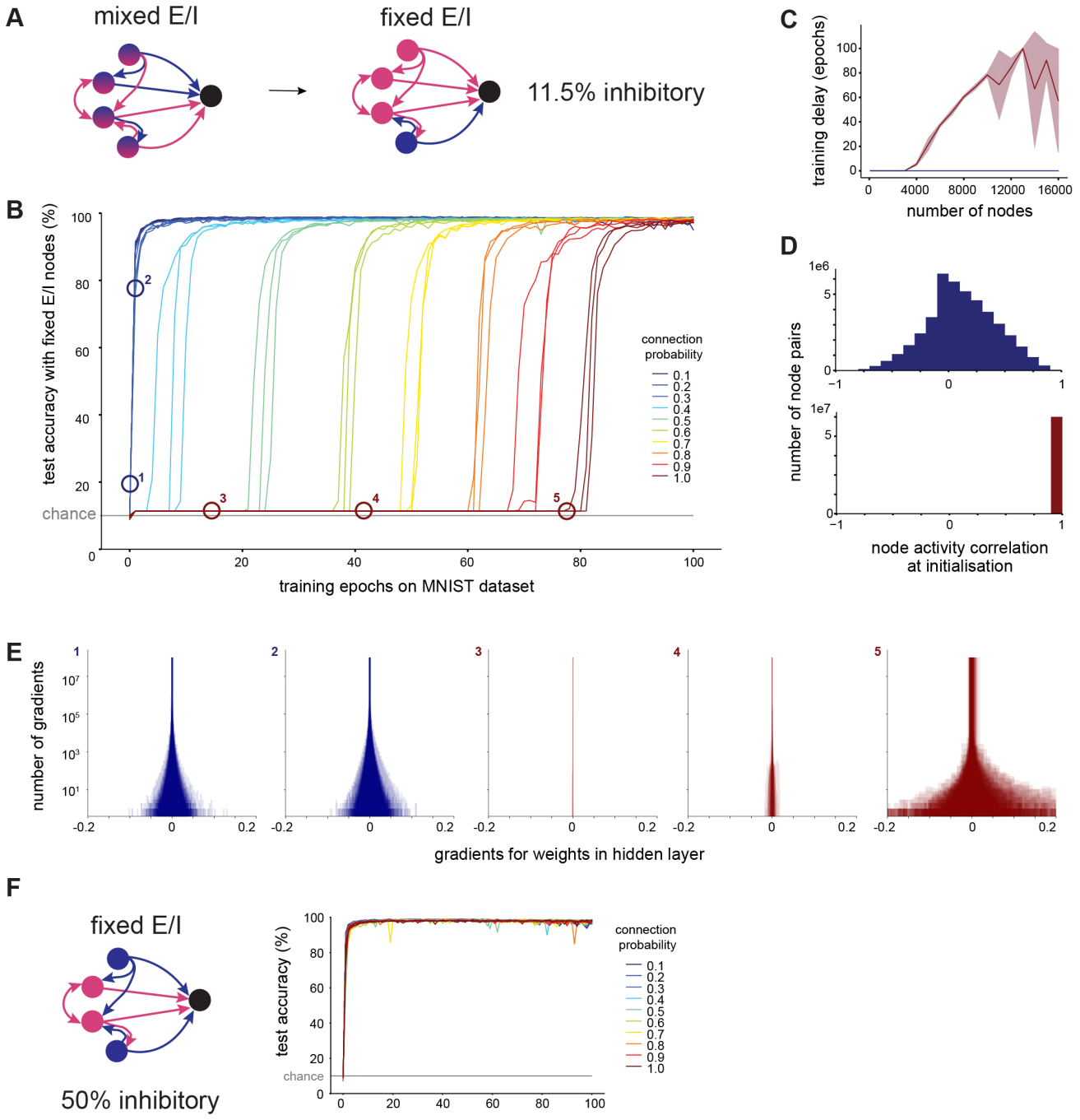

Figure 8. Sparse connectivity facilitates efficient training in networks with structural E/I imbalance seen in cortex. A: I constructed recurrent ANNs where each node was fixed to either excitatory (only positive outgoing weights) or inhibitory (only negative outgoing weights). I set 11.5 % of nodes to be inhibitory, as reported in measurements from somatosensory cortex. B: Test accuracy over time in large networks (10000 hidden layer nodes) trained on MNIST with fixed excitation and inhibition and different hidden layer

connection probabilities. C: Training delay (number of epochs before the network's performance exceeds chance level) as a function of the number of nodes in the network's hidden layer, for sparse (connection probability = 0.1, blue) and dense (connection probability = 1, red) connectivity. Shaded area shows standard deviation. D: Distribution of Pearson R correlation values between hidden layer node activations in response to all MNIST test dataset images, at initialisation (before any training), for sparse (connection probability = 0.1, blue) and dense (connection probability = 1, red) connectivity with 10000 hidden layer nodes. E: Distributions of gradients associated with hidden layer weights for sparse (connection probability = 0.1, blue) and dense (connection probability = 1, red) networks with 10000 hidden layer nodes. Numbers correspond to labels in panel B. F: I repeated the same experiment with fixed excitatory and inhibitory nodes, but now set 50 % of nodes to be inhibitory (balanced E/I). Test accuracy over time is shown for networks with 10000 hidden layer nodes.

The hidden layer activations in densely connected networks with a learning delay were not only highly correlated, but also very large, which led me to speculate that these networks may be experiencing an excitation-inhibition (E-I) imbalance due to the low proportion of inhibitory nodes. To test this, I repeated the training of Dale-compliant networks with a biologically unrealistic, balanced proportion of excitatory and inhibitory nodes (50 % excitatory, 50 % inhibitory). I found that these balanced networks did not experience a delay in training, regardless of their density (Fig. 8F). This confirms that densely connected networks with a biologically plausible fraction of inhibitory nodes experience an E-I imbalance which prevents them from learning efficiently, and that this imbalance is mitigated by sparse connectivity.

5.1.5 Beyond random connectivity: degree sequences and network motifs

The investigations in this section were performed by SH and EQ, and supervised by RF.

Finally, we investigated the effect of higher-order features of cortical connectivity on the training and performance of ANNs. It has been shown that fully connected, recurrent network motifs (Fig. 9A) are overrepresented in cortex compared to their expected prevalence in a randomly connected network (Udvary *et al.*, 2022, Fig. 9B). To investigate the effect of recurrent loops on network function, we constructed and trained ANNs with varying numbers of recurrent network motifs. To this end, we began with a randomly connected ANN (connection probability = 0.14 to match the barrel cortex column) and replaced the connectivity between randomly sampled triplets which were previously unconnected ('empty'

motif, Fig. 9A) by a fully recurrent motif. To maintain the overall number of edges in the network, we compensated for this addition of edges by removing edges from sparse feedforward motifs. Our modified connectivity matrices differ substantially from random connectivity, and are more similar to cortical connectivity in terms of network motif occurrences (Fig. 9C). A clear effect on performance when modifying the occurrence of recurrent loops was only seen in the smallest tested networks, with 100 hidden layer nodes (Fig. 9D). Similar to previous results, the largest effect was seen early in training, with all networks reaching comparable performance after ten training epochs. When trained on MNIST, networks built from only recurrent loops achieved a test accuracy of 65.39 ± 0.87 % after the first and 80.61 ± 2.47 % after the second training epoch, while networks with random connectivity attained 61.13 ± 1.75 % and 72.92 ± 3.96 % accuracy in the first two epochs. This difference was only statistically significant in the first training epoch (t-statistic = 3.07, p = 0.037). Networks with increased recurrent loops fell between these results in terms of training performance. In larger networks with 1000 and 10000 nodes, there was no significant effect of changing the occurrence of recurrent loops. This is generally consistent with our earlier finding that dense connectivity is beneficial in very small networks: although the overall connection probability of the network remains unchanged, we increased the number of fully connected recurrent loops, which are the most densely connected motif. It is therefore plausible that, even without changing the overall number of edges, it is still beneficial in small networks to create small, local areas of dense connectivity. The inclusion of Dale's principle, by making a fraction of all nodes inhibitory as described previously but otherwise keeping everything the same, did not lead to any additional observations of interest, nor did training on reduced training datasets.

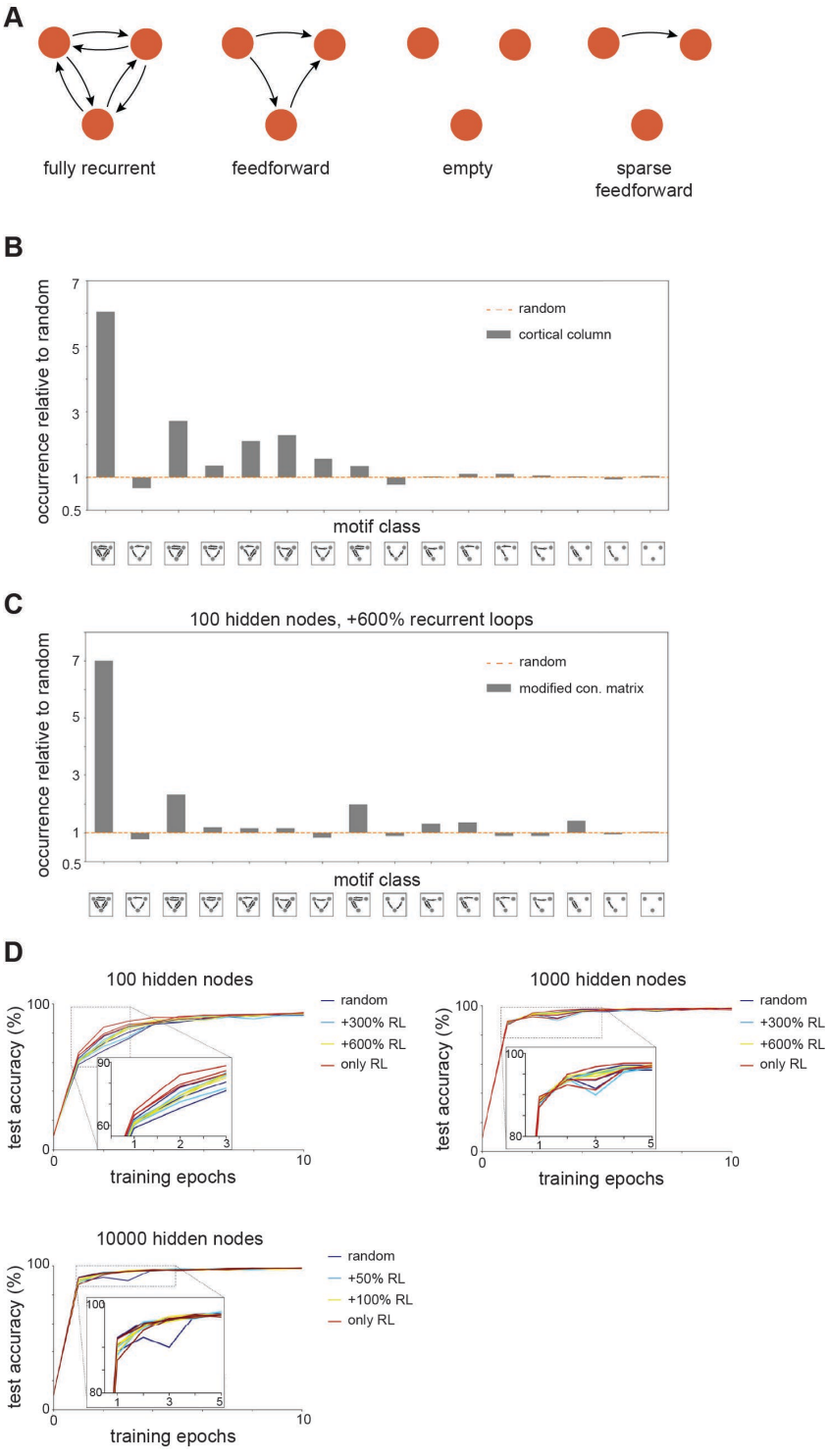

Figure 9. Effect of recurrent triplet network motif occurrences on ANN performance. A: Examples of triplet network motif classes relevant to this study. B: Occurrence of triplet motifs in a barrel cortex column (C2) from the multi-scale model, compared to in a random network.

C: Occurrence of triplet motifs in a network with 100 hidden layer nodes, modified to have a 600% increase in fully recurrent loops, compared to in a random network. D: Test accuracy of ANNs of different sizes, with connectivity matrices containing different amounts of recurrent loops, training on MNIST for 10 epochs. Random: randomly generated connectivity matrix. +x% RL: occurrence of fully recurrent loops increased by x%, overall connection probability maintained. Only RL: connectivity matrix constructed entirely of recurrent connections, overall connection probability maintained.

In cortical networks, the in- and out-degree of a neuron (i.e., the number of incoming and outgoing synaptic connections a neuron forms) are correlated (Landau et al., 2016; Udvary et al., 2022). We quantified the in- and out-degrees for all neurons in the C2 column of the barrel cortex model (Fig. 10A). Compared to a randomly generated connectivity matrix with the same connection probability, cortical networks have more broad degree distributions. We construct ANNs which preserve these cortical degree sequences in several ways: we sample the in- and out-degree for each node separately from the in- and out-degree distributions (uncorrelated), we sample in- and out-degree from the same neuron, which is randomly chosen for each node, preserving correlations between in- and out-degree (correlated), and we construct an ANN adhering to Dale's principle as before, and sample the in- and out-degree for each node from a randomly chosen neuron of the same type (excitatory or inhibitory). We tested networks with 500, 5000 and 10000 hidden layer nodes. We trained all networks for 100 epochs on MNIST and CIFAR10 – the first 10 training epochs on the CIFAR10 dataset are visualised here (Fig. 10B). We observed no notable differences between any networks of one size, with networks preserving the cortical in- and out-degrees not differing significantly from random networks in their performance.

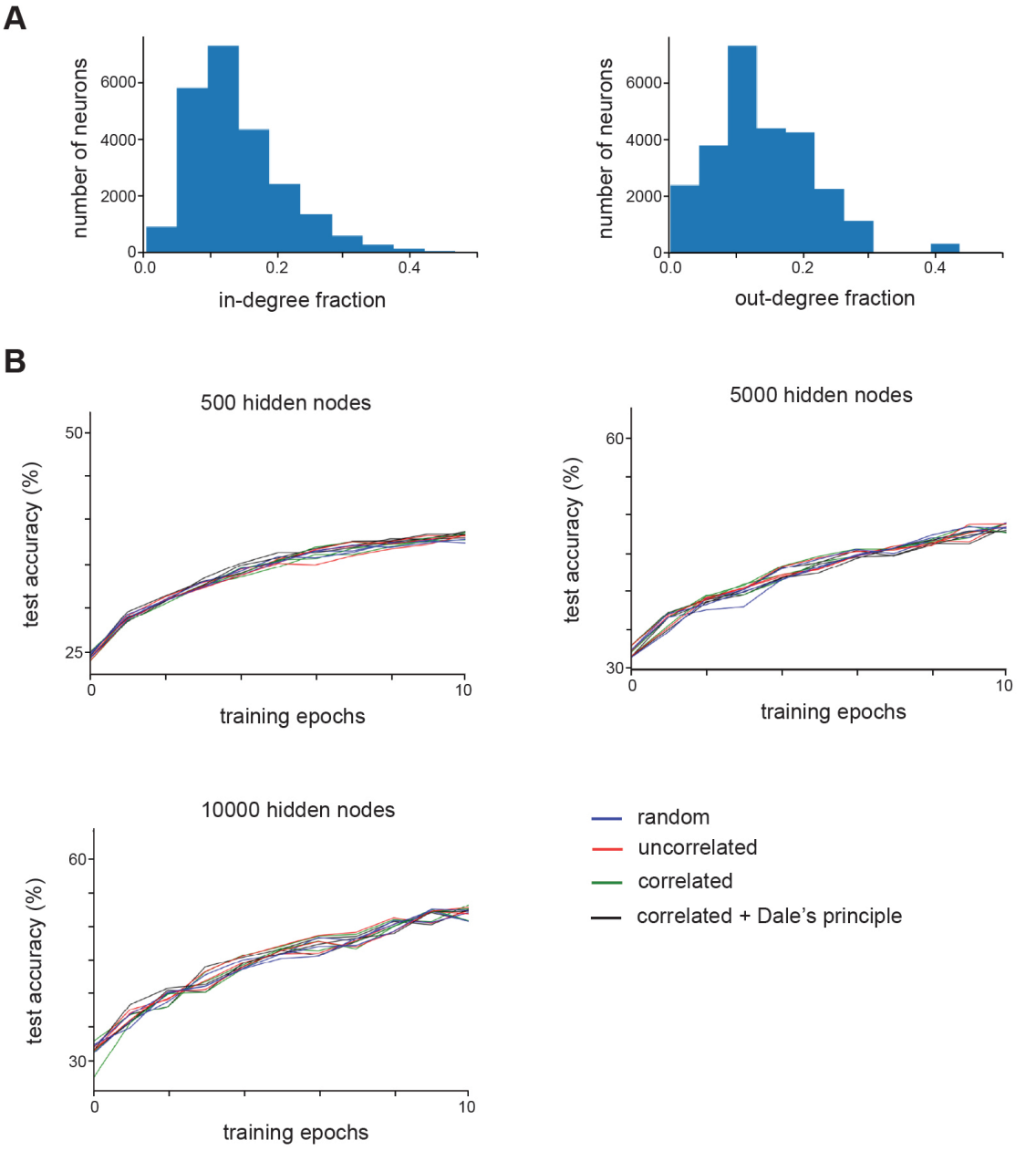

Figure 10. Effect of degree distribution preservation on ANN performance. A: In- and out-degree distributions of the C2 column of the barrel cortex model. B: Test accuracy of ANNs of different sizes on the CIFAR10 dataset, with different methods for transferring degree sequences from the cortical network. Random: randomly generated connectivity matrix. Uncorrelated: in- and out-degree sampled independently for each node from distributions in panel A. Correlated: in- and out-degree for each node taken from a randomly sampled neuron, preserving degree correlations. Correlated + Dale's principle: Dale's principle network constructed as before, then in- and out-degree for each node taken from a randomly sampled neuron of the corresponding type (excitatory/inhibitory).

5.2 Network-neuron interactions underlying sensory responses of layer 5 pyramidal tract neurons in barrel cortex

5.2.1 Reduction of the multi-scale models into analytically tractable models

The starting point for this study was the multi-scale model of rat barrel cortex, which we extended to account for cell-to-cell and trial-to-trial variability. We performed simulations of L5PT responses to passive deflections of either the somatotopically aligned “principal whisker” (PW) or one of the eight surround whiskers (SW) (Fig. 11A). Previously, the multi-scale model only included a single L5PT model, so in order to account for cell-to-cell variability we extended it with models from five different L5PT dendrite morphologies (Fig. 11B), with a total of seven biophysically detailed models with different biophysical compositions (Table S4), reflecting different densities and dendritic distributions of active conductances. All of these L5PT models reproduced the cell type’s characteristic electrophysiological properties (Fig 11H & Table S5). *In vivo,* the L5PTs from which the dendritic morphologies were reconstructed had the typical broad and heterogeneous receptive fields (Fig. 11C). We embedded the L5PT models into the barrel cortex model, which provides realistic estimates for which neurons in the primary thalamus (VPM) and barrel cortex are synaptically connected to the L5PT (Fig. 11D) and the locations of these synapses on the dendrite (Fig. 11E). We placed each L5PT model at 81 different locations in and around the C2 column (Fig. S5). Activating each neuron with cell type-specific activity measured experimentally for this stimulus condition (de Kock *et al.*, 2007, Fig. S6) results in different empirically well-constrained spatiotemporal synaptic input patterns (Fig. 11F) for each morphology, location of the cell and configuration of active neurons in the network. These input patterns down to the level of single synapse activity (Fig. 11G) are consistent with those observed experimentally via dendritic spine imaging (Varga et al., 2011). Our simulations reproduce the characteristic fast responses (Fig. 11I) and broad and heterogeneous receptive fields (Fig. 11J). The distribution of response probabilities closely matches for all whisker stimuli (p-values ranging between 0.07 and 0.92) and the means of these distributions, i.e., the ‘mean receptive field’, is significantly correlated (R value 0.76, $p = 0.017$).

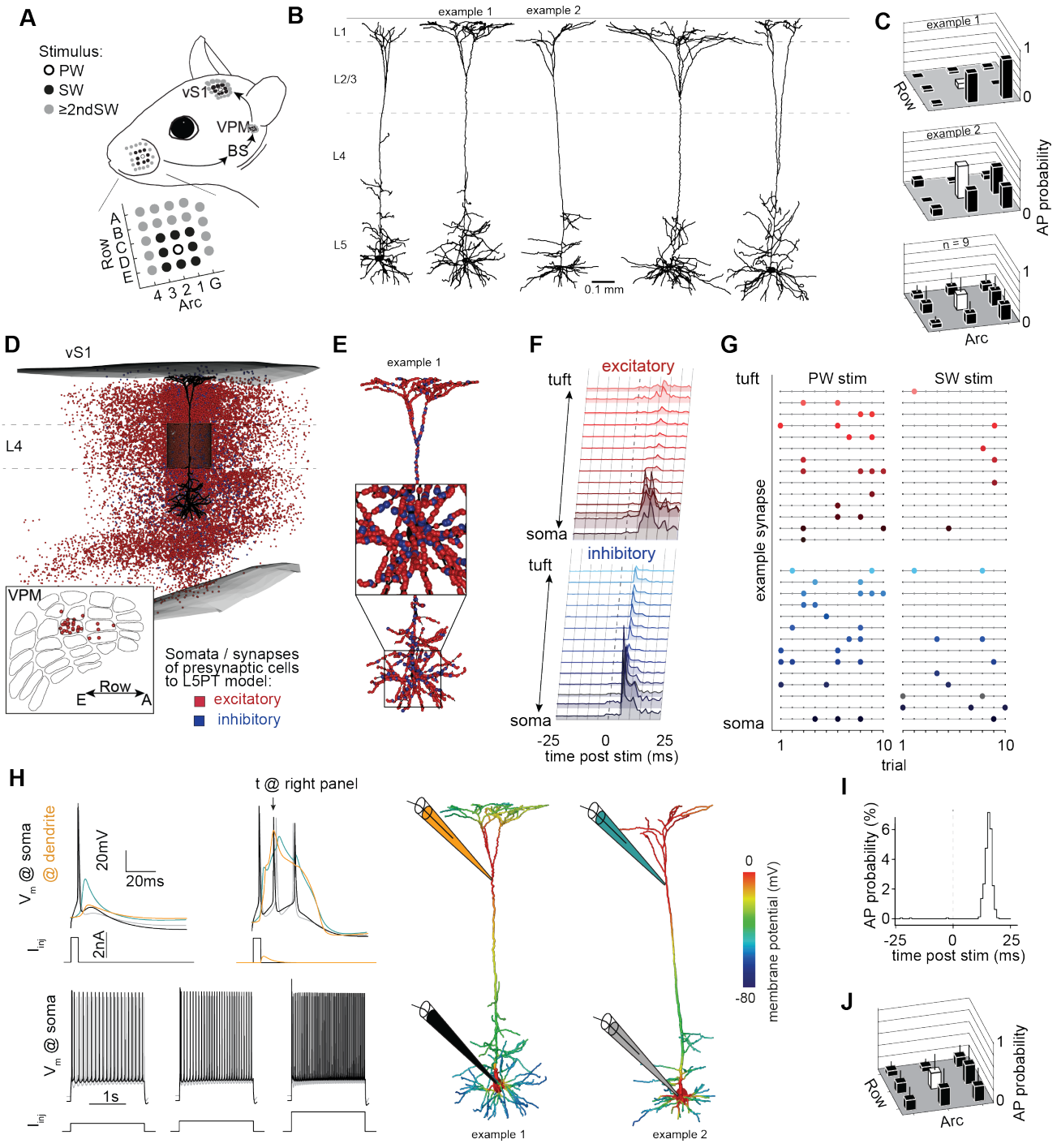

Figure 11. Biophysically detailed multi-scale model of whisker deflection evoked responses in cortical pyramidal tract neurons. A: Sensory-evoked signal flow: stimuli of single whiskers (which are arranged in 'arcs' and 'rows' on the animal's snout) are transmitted to the brainstem (BS), from there to the VPM thalamus, and from there to the primary sensory cortex of the vibrissal system (barrel cortex, vS1). This pathway is somatotopically organized,

with barreloids in VPM and barrels in cortex corresponding to the respective whiskers. B: In vivo labeled L5PT dendrite morphologies used in this study. C: Corresponding receptive fields to passive single whisker touch, measured in vivo (upper panels). Average receptive field across 9 in vivo recorded L5PTs (lower panel). Error bars are std. D: Network model of rat barrel cortex and VPM provides anatomically realistic estimates of which neurons are connected to a L5PT embedded into the network. In this study, the simulated neurons are located in the C2 column of vS1, thus we refer to the somatotopically aligned C2 whisker as the 'principal' whisker, and the adjacent whiskers as surround whiskers. Red and blue markers denote soma locations of presynaptic excitatory and inhibitory somata, respectively. E: Synapse distribution originating from the neuron shown in Panel B. F: Spatiotemporal input pattern to L5PT: combining the anatomical constraints with empirical measurements of the activity of different presynaptic populations (Fig. S6) provides spatiotemporal input patterns that the L5PT can receive after sensory stimulation. G: Trial-to-trial activity of example synapses matching the soma distance from panel F for a principal whisker (C2) and surround whisker (D2) stimulus. H: Biophysically detailed multi-compartmental L5PT models reproduce the cell type's characteristic electrophysiology (left panel), i.e. back propagation of APs (upper left), dendritic Ca-APs and somatic burst firing (upper right), as well as regular firing properties (lower row). Right panel: biophysically detailed neuron morphologies at the moment of a dendritic Ca AP. I: Simulated response to principal whisker touch. J: Simulated receptive fields across morphologically and biophysically diverse L5PT multi-compartmental models across 81 network embeddings capture broad and heterogeneous receptive fields.

The challenge now, despite these highly realistic and detailed simulations, remains in understanding the effective computation performed by the system. The first difficulty in understanding the computation performed by L5PTs in the multi-scale model arises from the biophysically detailed models themselves: although mechanistically accurate and able to realistically reproduce *in vivo* responses, these models are based on a series of partial differential equations, which are not analytically tractable. This makes it almost impossible to understand the computation being performed by a given L5PT to transform synaptic input into AP output. The second challenge arises from the many sources of variability in the multi-scale model: from trial-to-trial variability in synaptic activations to cell-to-cell variability in morphological and biophysical properties, this variability makes it challenging to extract general computational principles from the model. We here begin by addressing the first challenge, and reducing the biophysically detailed models into analytically tractable models which explain the input-output computation at the single cell level.

The model reduction approach was developed using one of the biophysically detailed models (morphology from example 1 in Fig. 11B). In a dataset comprising 1000 simulation trials from each of the 81 network embeddings and 9 whisker stimuli, we searched for features of the synaptic input which are predictive for the generation of an AP in a given millisecond (hereafter referred to as the prediction time point, Fig. 12A). We found that a weighted count of active inhibitory vs excitatory synapses, where each synapse activation is weighted by its distance from the soma and its time relative to the prediction time point, can predict AP output well, with an AUROC score of 0.949 (Fig. 12B). According to these spatial and temporal weights, which we hereafter refer to as the spatiotemporal filter, the contribution of both excitatory and inhibitory synapses to AP output decays with increasing soma distance, reaching approximately zero at around 500 µm. In the temporal dimension, the contribution of excitatory and inhibitory synapses has a time course resembling excitatory and inhibitory postsynaptic potentials (EPSPs and IPSPs). Excitatory and inhibitory synapses contribute the most to AP output if they were active around 4ms and 9ms before the prediction time point, respectively. In the following, we refer to this weighted count of active synapses as the weighted net input (WNI). The probability of observing an AP in the biophysical model increases with the WNI (Fig. 12C).

We revisited trials which are misclassified by the WNI, and found that misclassifications are most frequent when there was an AP a few milliseconds before the prediction time point. This is because the biophysically detailed model is less excitable for a short period following an AP, reflecting the time constants of involved ion channels. Therefore, even if the WNI is high the probability of another AP is low. To account for this spike history-dependent property of the biophysically detailed model, we determined the boundary separating trials with an AP from those without an AP and with a recent AP based on WNI and time since the previous AP (Fig. 12D). Subtracting the value of this boundary as a post AP penalty from the WNI results in an improved AP prediction accuracy (AUROC 0.990). Thus, three features are sufficient to predict AP output in the biophysically detailed model: the spatial and temporal distribution of synaptic input, and the time since the previous AP.

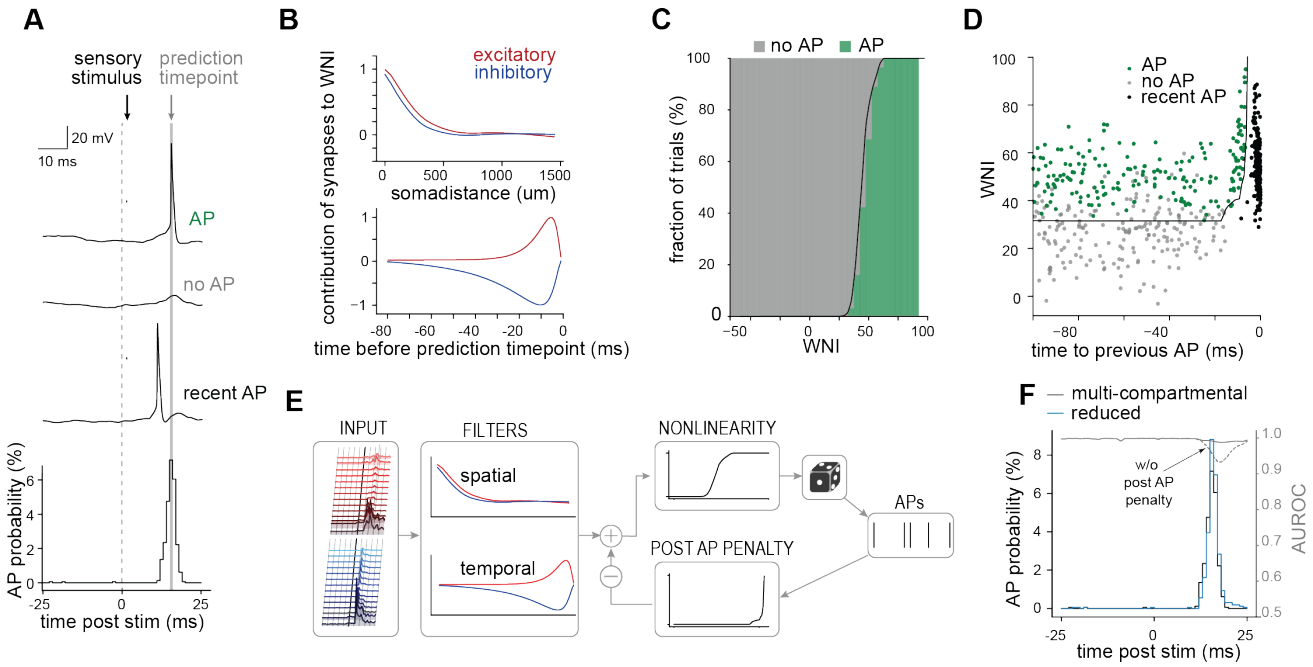

Figure 12. Input-output computation of L5PTs upon single whisker deflections. A: Exemplary responses of the multi-compartmental model with respect to the prediction time point (the time point for which the occurrence of an AP is to be predicted) for the three relevant response categories 'AP', 'no AP', and 'recent AP' (AP was elicited shortly before the prediction time point). B: Spatiotemporal input filter that best separates AP and no AP trials assigns strong weight to proximal synapses (top) active in a short time window before the prediction time point (bottom). C: Nonlinear relationship between WNI and AP probability. WNI represents the 'drive' a neuron receives; the higher the WNI the higher the probability an AP will be generated. D: Weighted net input–the input filtered by the spatiotemporal filter–separates AP and no AP trials, but not 'recent AP' trials, which can be distinguished based on a second measure, 'time to previous AP'.E: Reduced model structure. APs are generated stochastically based on the AP probability (output of the nonlinearity). If an AP is generated, subsequent APs become less likely due to the post AP penalty, which is subtracted from the WNI. This reduced model directly relates AP output to synaptic input and previously generated APs in the simulated in vivo condition. F: The reduced model's responses match the biophysically detailed model across many trials (close PSTH match) and on the and on the single trial level (high AUROC score across all time points). Without the post AP penalty, the AUROC score drops during the sensory-evoked response.

Based on these three features, we can therefore describe the input-output computation of the biophysically detailed model under this stimulus condition in an analytically tractable model. In order to predict AP output from synaptic input, we assemble the spatiotemporal filters, the nonlinear relationship between WNI and AP probability, and the post AP penalty into a

generalised linear model (GLM, Fig. 12E). We applied the GLM to predict AP output at time points ranging from 25 ms before to 25 ms after the onset of the whisker stimulus. Even though the GLM was inferred from the peak of AP responses (see Methods), it maintained a high AUROC score before and after this time point (Fig 12F). On a single-trial level, the model was able to predict whether or not an AP was elicited in the 25 ms after the stimulus with an accuracy of 96.7 % (Table S6), and the timing of APs within this window was accurate to within 1.1 ± 1.6 ms (Fig. S7). Thus, the GLM accurately predicts APs throughout the entire time interval, and the post-stimulus time histogram (PSTH) predicted by the GLM hence closely matches with the PSTH of the multi-compartmental model (Fig 12F).

5.2.2 Input-output computation is robust to morphological and biophysical diversity

Is this input-output computation a general property of L5PTs, despite cell-to-cell variability in morphology or biophysics? To address this question, we repeated model reduction on each of the seven biophysically detailed models, which represent five different dendritic morphologies. We find that the input-output computation is qualitatively preserved across L5PTs with different morphologies and biophysical properties (Fig. 13A): all L5PTs count active excitatory and inhibitory synapses, assigning more importance to proximal synapses active shortly before the prediction time point, and with a similar penalty to AP probability following an earlier AP. Each reduced model achieved a high accuracy throughout the sensory-evoked response (Fig. 13B), with highest accuracies generally during the peak response (AUROC median/min/max: 0.97/0.93/0.998). At the single trial level, the reduced models were able to predict whether or not an AP was elicited in the 25ms after the stimulus with a mean accuracy of 90.9 ± 3.9 % (min 85.0 %, max 96.7 %, confusion matrices for all models are provided in Table S6) and the timing of APs was predicted accurately to within a mean of 2.0 ms (min 1.1 ms, max 3.1 ms, Fig. S7).

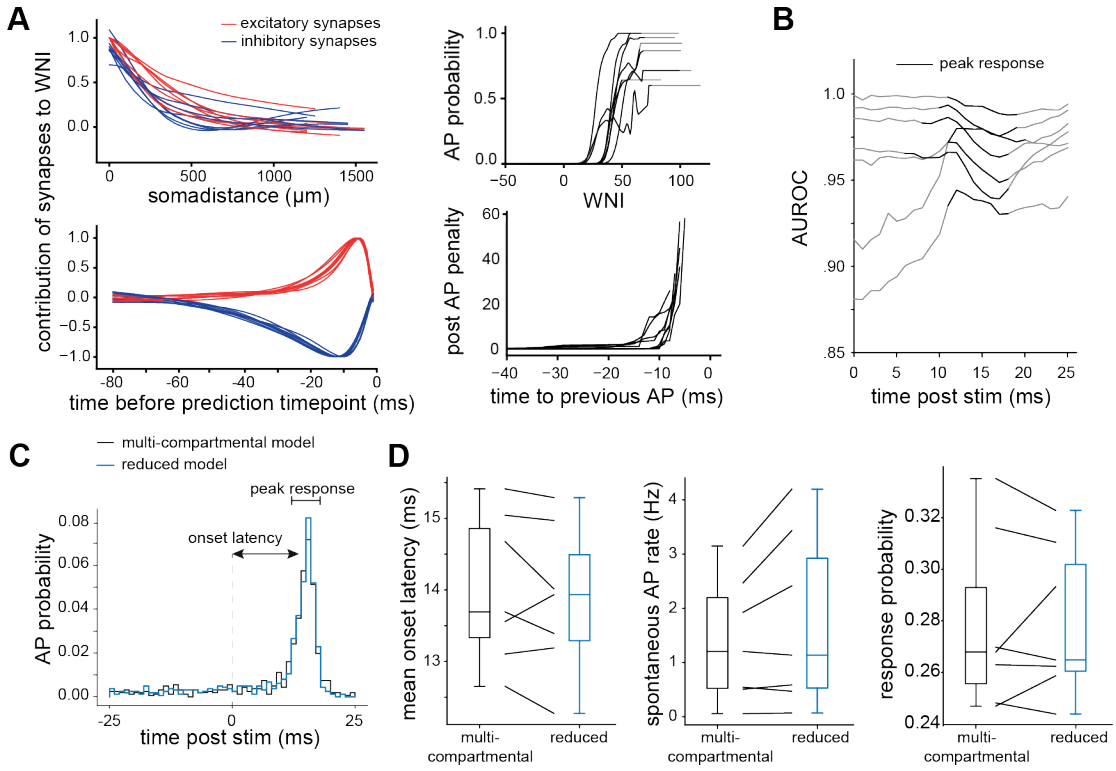

Figure 13. Input-output computation is robust to morphological and biophysical diversity. A: Reduced models inferred on the different multi-compartmental models are qualitatively similar, with similar temporal and spatial filters, nonlinearity and post AP penalty. B: All models have high AUROC scores, specifically during the sensory-evoked (peak) response. C: We quantify latency, spontaneous AP rate (before the stimulus) and response probability for each pair of multi-compartmental and reduced model. D: Comparing response properties between multi-compartmental and corresponding reduced model shows close match.

While the input-output computation is qualitatively preserved across L5PT models, the reduced models are not identical. For example, the decay of the spatial filter with soma distance differs slightly between reduced models, which we find largely reflects diameter differences of the apical trunk dendrite (Fig. S8). To what extent do these variations in the reduced model capture real, small differences in the input-output computation across L5PTs? We compared the predicted PSTHs between each biophysically detailed model and its corresponding reduced model and found that the slight differences between their shapes are matched well by the reduced model (Fig. S7). To quantify this similarity, we compared the time to maximum response (latency), spontaneous AP rate before the stimulus, and response

probability between biophysically detailed models and corresponding reduced models (Fig. 13C). Each reduced model generated responses which matched those of the corresponding biophysically detailed models in all of these properties, while preserving the considerable variation across them (Fig. 13D). Thus, the small variations in the reduced models account for the differences in the input-output computation between the biophysically detailed models.

5.2.3 Linear reduced models capture effects of nonlinear mechanisms

It is somewhat surprising that the reduced models are able to accurately predict AP output in the biophysically detailed models with such a linear model structure. This suggests that either the synaptic input patterns arising from passive whisker deflections in the anaesthetised animal do not strongly activate nonlinear dendritic mechanisms in L5PTs, or that these nonlinear mechanisms are active but do not greatly influence the prediction of AP output. To determine whether nonlinear dendritic mechanisms are active during our stimulus condition, we recorded currents through voltage-gated sodium and calcium channels, and through N-methyl-D-aspartate (NMDA) and α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptors during a simulation of a PW stimulus with a biophysically detailed model (example morphology 2). We recorded these currents for 16 locations across the basal and apical dendrites, including the primary bifurcation point (BP) of the apical trunk (Fig. 14A).

We find that nonlinear dendritic mechanisms are frequently active in our simulated experimental condition. For example, we find calcium-mediated depolarisations which exceed -30 mV at the primary BP in 18 % of simulation trials (Fig. 14B middle). However, this calcium influx generally does not lead to bursts of APs at the soma (of 179 trials with calcium depolarisation: 13 % no AP, 83 % 1 AP, 4 % burst with 2 APs), and the reduced model was able to predict a response equally well for trials with and without a calcium depolarisation (86 % vs 90 %, $p = 0.07$). In addition to calcium transients, we frequently observe peaks of sodium influx in the distal dendrites (Fig. 14E). NMDA currents dominate the synaptic conductance in the trunk and tuft dendrites (Fig. 14C, D & F). However, despite generally being considered a nonlinear conductance, they do not seem to impact the prediction accuracy of the reduced model: we compared trials with low NMDA conductance (33rd

percentile and below) and high NMDA conductance (66th percentile and above) in the tuft dendrites and found no significant difference for the accuracy of the reduced models (88 % vs 91 %, p = 0.17). Thus, active nonlinear synaptic and dendritic mechanisms are active during simulations of passive single whisker deflection, but either their impact on somatic AP output is low in this condition, or the reduced models are able to capture their impact on the effective input-output computation.

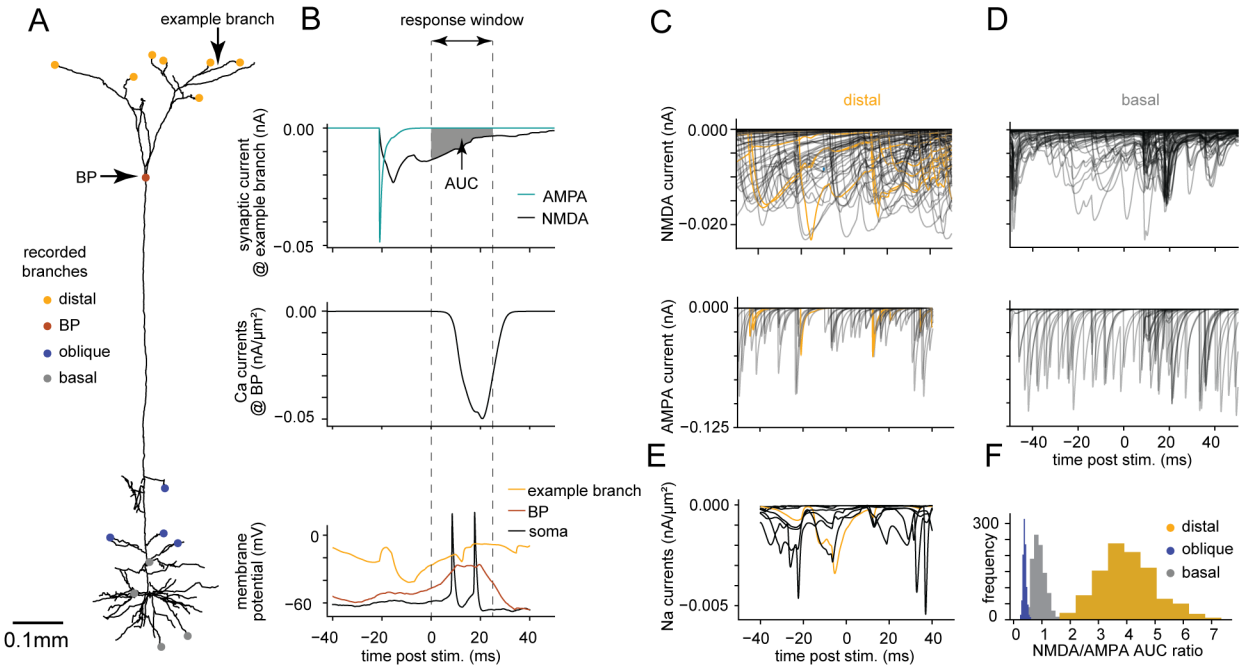

Figure 14. Active nonlinear mechanisms in simulations of passive whisker deflection with biophysically detailed multi-compartmental models. A: example morphology 2 (from Fig. 11), for which we simulated 1000 PW stimuli while recording synaptic AMPA and NMDA currents, ion currents and the transmembrane potential from the marked branches. B: example trial with a 2 AP burst response at the soma, a calcium-mediated depolarization at the primary bifurcation point (BP), and AMPA and NMDA currents of an example synapse at the example branch. We quantified the charge exchanged through the AMPA and NMDA receptors of each synapse during the 25ms window following the whisker stimulus ('response window'). C: synaptic currents of all synapses recorded on distal branches. Yellow lines are synapses on the example branch. D: as C, but for basal dendrites. E: Sodium currents recorded at distal dendritic branches. The yellow line is the example branch. F: Ratio of NMDA/AMPA area under the curve (AUC) across 1000 simulation trials.

To investigate this, we re-simulated trials for each of the 9 whisker stimuli for a biophysically detailed model (example morphology 2) but either reduced the NMDA conductance by 50 % or completely removed NMDA conductance in all synapses. Reducing NMDA conductance reduced sensory-evoked responses to all whiskers (Fig. 15A & B), indicating that NMDA contributes strongly to AP output. We find that reduced models inferred from these modified biophysically detailed models show changes to their spatial and temporal filters (Fig. 15C). According to these reduced models, NMDA increases the influence of distal synapses (Fig. 15D) and widens the integration time window for synapses (Fig. 15E). Interestingly, reduced models for biophysically detailed models with reduced or absent NMDA have a better prediction accuracy than reduced models for the control biophysical model (Fig. 15F). These findings imply that in the case of NMDA, the reduced models are able to capture the influence of this nonlinear conductance on the effective input-output computation to an extent, where it is reflected in the shape of the spatiotemporal filter, but there is some small residual effect which the linear reduced models cannot account for.

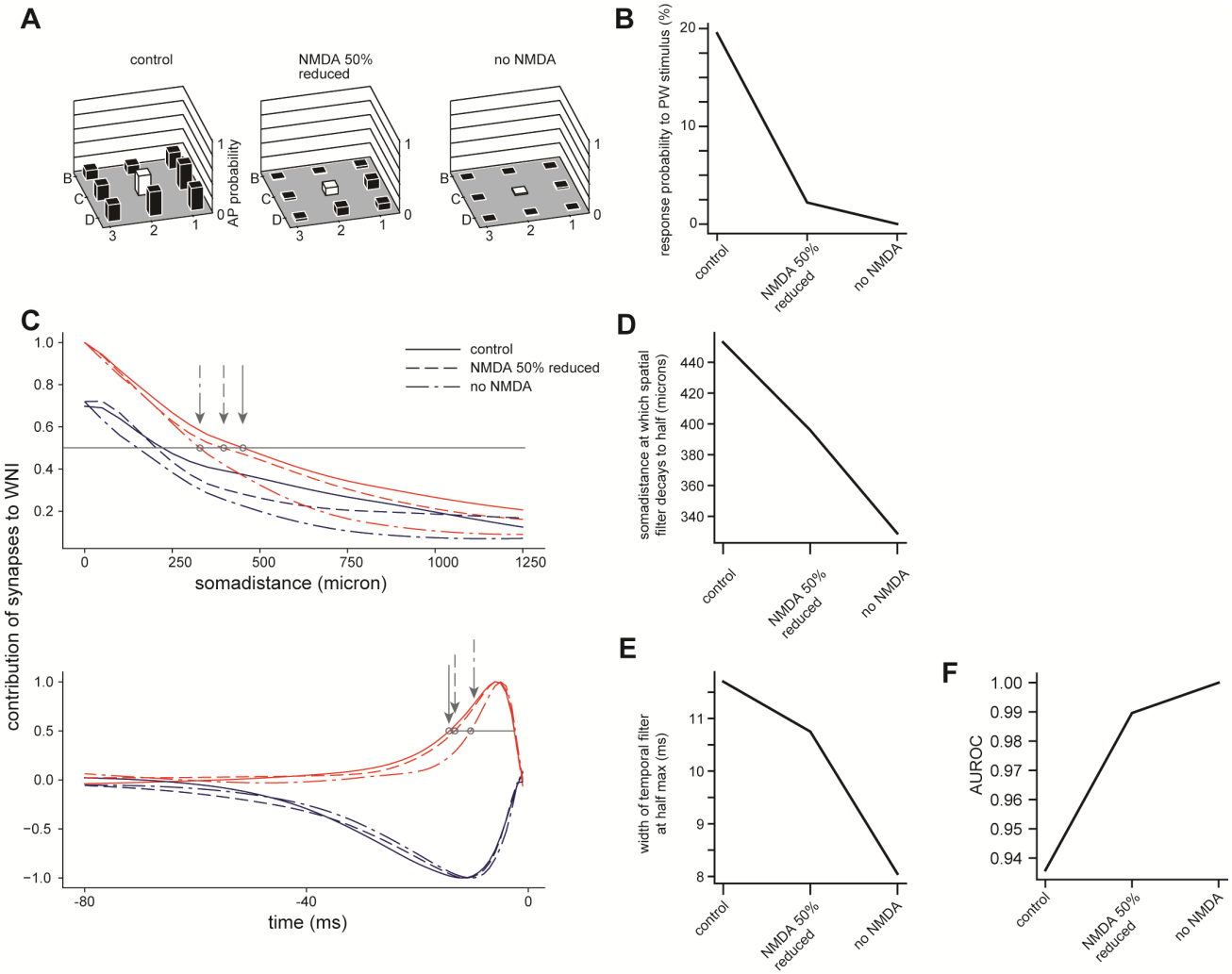

Figure 15. Influence of NMDA on sensory-evoked responses and input-output computation. A: Simulated receptive fields to passive single whisker touch for one L5PT multi-compartmental model across 81 network embeddings, with NMDA conductance set to 100 %, 50 % and 0 % of control value. B: Response probability to a principal whisker stimulus by NMDA conductance. C: spatial and temporal filters (red: excitatory synapses, blue: inhibitory synapses) inferred from multi-compartmental models with different NMDA conductances. D-E: width of spatial and temporal filters by NMDA conductance. F: AUROC score of GLM by NMDA conductance.

5.2.4 Contribution of different input pathways to sensory responses

Which presynaptic neuron populations contribute to sensory-evoked responses? The reduced models have so far only distinguished between excitatory and inhibitory inputs, however the multi-scale model contains information about the location and cell type of the

neuron from which each synapse originates. To separate the contribution of each presynaptic population to L5PT output, we calculated each cell type's contribution to the WNI individually by applying the spatiotemporal filters to synaptic input from one cell type at a time (Fig. 16A). We first analysed L5PTs at the centre of the column somatotopically aligned with the PW defined in this study (C2). By looking at the contributions of different presynaptic populations to WNI, it becomes clear that they contribute in two distinct ways: by their spontaneous activity preceding the stimulus ('baseline WNI'), and/or by an increase in activity following the stimulus ('sensory-evoked WNI'). During periods of spontaneous activity, other L5PTs are the largest contributor to the baseline WNI – this is consistent with their high spontaneous firing rates and high pairwise connectivity (de Kock et al., 2007; Hay and Segev, 2015). Subtracting the baseline WNI to isolate the sensory-evoked contributions of different populations, we find that the primary contributor to sensory responses in the case of a PW stimulus is direct thalamic input from VPM (Fig. 16B). This is different for a SW stimulus (Fig. 16C), where the contribution of VPM is minimal and the primary source of excitatory drive is L6CCs. The reduced models thereby predict that the contributing thalamocortical and intracortical pathways depend on the stimulated whisker.

Given that the identity of the whisker stimulus relative to the location of the L5PT (somatotopically aligned or not) affects the contributing input pathways, does the location of the L5PT within a single column also have an effect? We computed the relative WNI contribution of VPM and L6CC, the two primary contributors to sensory-evoked WNI for a PW stimulus, with the L5PT at 81 different embedding locations in and around the C2 column (Fig. 16D). We find that VPM input dominates only in the centre of the barrel column, and that towards the edge of the barrel and in the septa, L6CC is the most important input pathway for sensory-evoked responses. Thus, for most L5PTs in the column, L6CC is the major contributor to sensory-evoked excitation. This is consistent with an earlier finding, which showed that experimentally inactivating L6CCs results in narrow receptive fields, largely abolishing responses to SWs and greatly reducing the response to the PW (Egger et al., 2020). The effect of this manipulation is reproduced by the reduced models (Fig. 16E), indicating that the reduced models capture the origin of broad receptive fields in L5PTs.

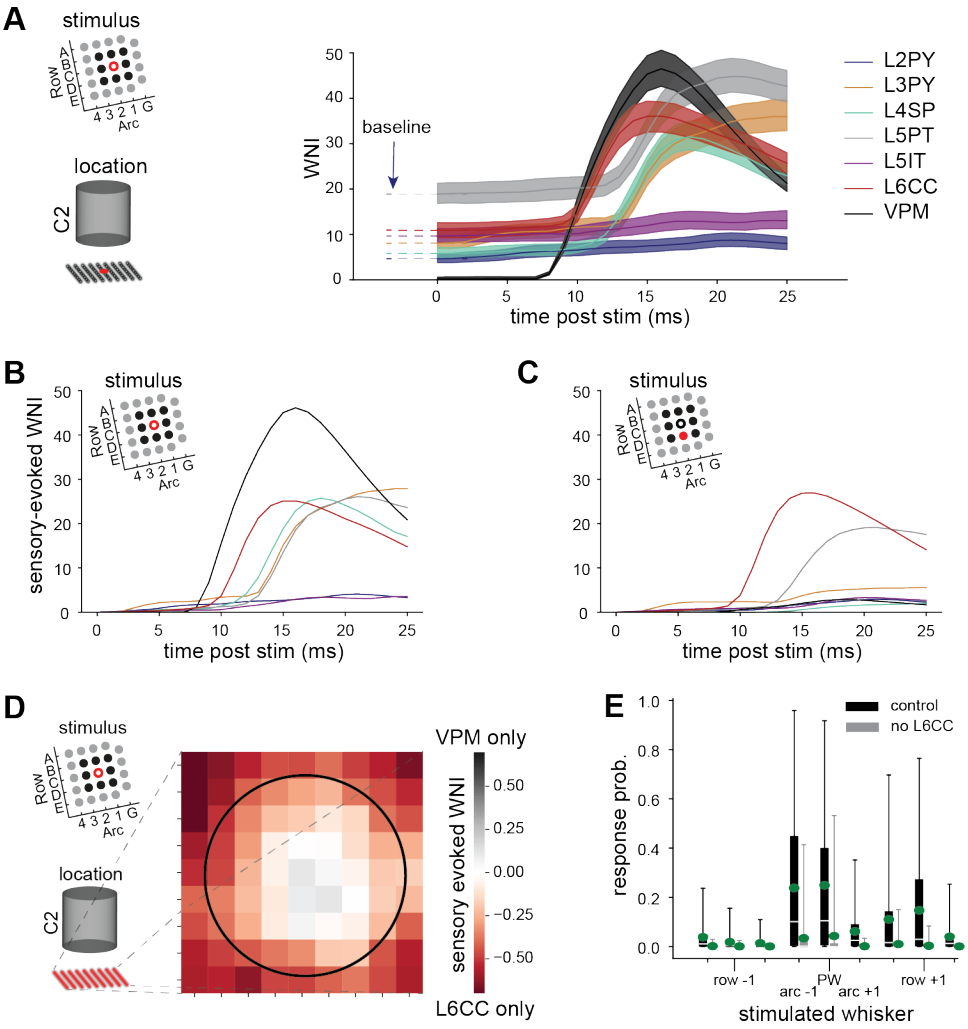

Figure 16. Reduced models predict contribution of input pathways to sensory responses. A: Absolute contribution of presynaptic populations to WNI following a PW stimulus to a model located at the centre of the C2 column: pyramidal neurons in L2/3 (L2PY, L3PY), spiny neurons in L4 (L4SP), intratelencephalic neurons in L5 (L5IT), L5PT, corticocortical neurons at the L5/6 border (L6CC), and relay cells in the ventral posterior medial nucleus of thalamus (VPM). Despite the lack of sensory-evoked responses by L4PY, L6CT and L6INV (Egger et al., 2020), their contributions were considered in the overall WNI calculations (Figs 11 and 12). B: Sensory-evoked contribution (i.e. absolute contribution minus baseline for each input pathway) of presynaptic populations to WNI following a PW stimulus. C: Sensory-evoked contribution of presynaptic populations to WNI following a SW stimulus. D: Contribution of the main input pathways–VPM and L6CC–depending on the soma location of the L5PT model in a 9x9 grid across the C2 column for a PW stimulus. The black circle denotes the C2 column border. E: Comparison between model responses to 9 different whisker stimuli (PW and 8 SW) under control conditions and when removing

sensory-evoked input from L6CC. Removing evoked L6CC activity attenuates responses, in particular to surround whiskers.

5.2.5 Network vs neuron contributions to receptive field variability

How can the large variability in L5PT responses and receptive field shapes arise from such small variability in input-output computation? To address this question, we repeated simulations of all whisker stimuli and network embedding locations with the reduced models, which we had previously simulated with the biophysically detailed models. The reduced models closely match the response probability to a given whisker stimulus of their corresponding biophysically detailed model (Fig. 17A, Pearson correlation coefficient = 0.97). Consequently, the receptive fields of reduced and biophysically detailed models are also highly similar (Fig. 17B & C, Pearson R between receptive fields 0.96 ± 0.01). The sensory responses of the reduced models also closely match *in vivo* measurements in terms of response probability to different whiskers (Fig. 17D) and large variability in the shape of the receptive field between L5PTs (Fig. 17E).

As the reduced models accurately capture the trial-to-trial and cell-to-cell variability in sensory-evoked responses, we now use them to disentangle the contributions of neuronal and network properties on receptive field shape. To this end, we computed the similarity between receptive field shapes from reduced models representing different biophysical properties, dendritic morphologies, and different embedding locations in the network model (Fig. 17F). We find that a change in network embedding, and therefore in synaptic connectivity, has the largest effect on receptive field shape (i.e., results in the least correlated receptive fields). The dendritic morphology has a lesser impact on receptive field shape, and the biophysical properties have little effect. The reduced models thereby suggest that variations in network input are the primary cause of receptive field variability between L5PTs. Overall, the reduced models predict that under the experimental condition of passive single whisker deflection in anaesthetised rats, the input-output computation is largely preserved across L5PTs despite their dendritic and biophysical variability, and that the cell-to-cell and trial-to-trial variability in sensory-evoked responses primarily arises from variability in network input.

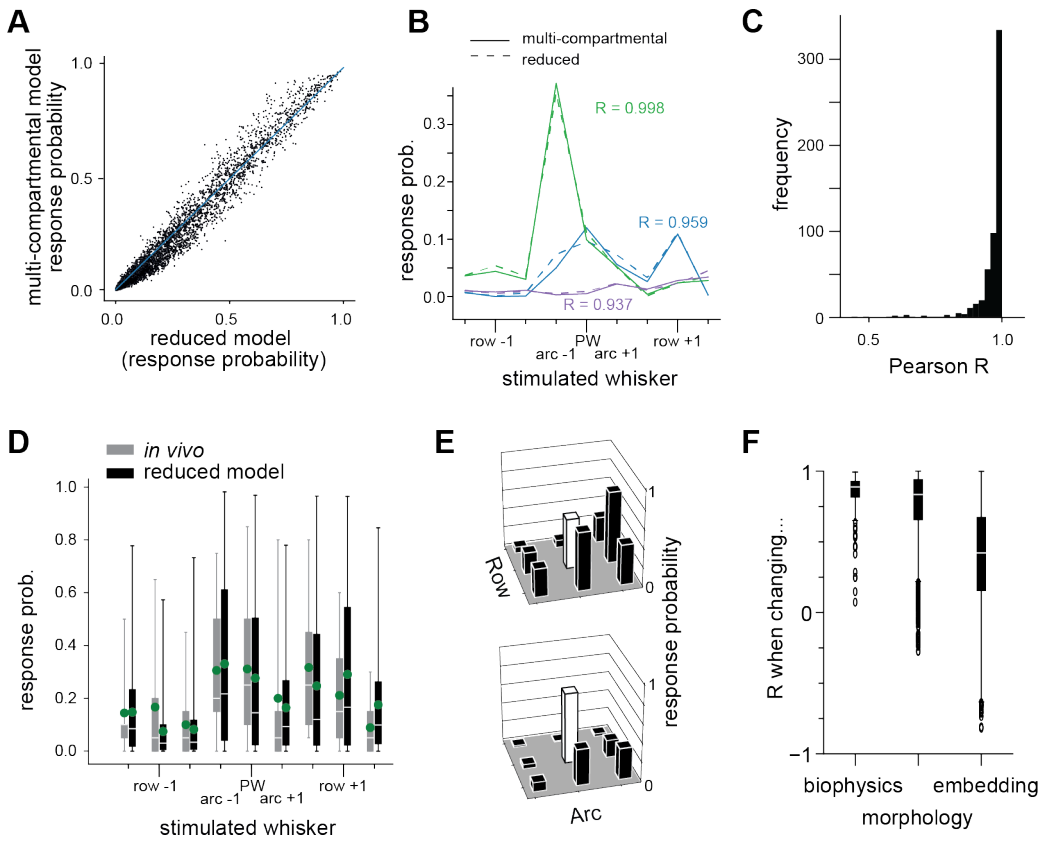

Figure 17. Reduced models predict origins of receptive field variability. A: Comparison of responses of 7 different biophysically detailed models and their corresponding reduced models to 9 different whisker stimuli (PW and 8 SW) in 81 different network embedding locations. Response probability is the probability that one or more APs are generated 0- 25ms after the sensory stimulus. B: Comparison of exemplary receptive field shapes shows close match between biophysically detailed and reduced model. C: Quantification of receptive field similarity for all cell positions and biophysically detailed models. D: Comparison between in vivo and reduced model responses to 9 different whisker stimuli (PW and 8 SW). Green dots represent the mean response probability. E: Exemplary receptive fields of reduced models. F: Influence of biophysics, morphology and cell position on receptive field shape, quantified by computing the correlation coefficient between receptive fields if one of these properties is changed.

5.3 The origin of bursts in sensory-evoked cortical output

5.3.1 Characterisation of calcium potentials in the distal dendrites

We began with a database of 660960 L5PT simulation configurations, comprising 68 different biophysically detailed models from 3 dendritic morphologies, which were selected to represent the full biophysical parameter space, and which were embedded in 9720 different network configurations of the barrel cortex model. We simulated either a single PW stimulus or a multi-whisker stimulus including the PW and 2 SWs. Of these, we selected the 20359 L5PT simulation configurations whose ongoing and sensory-evoked firing rates matched *in vivo* observations, and which had sensory-evoked somatic burst firing.

In order to quantify the occurrence of CaAPs and measure the effect of manipulations, we need a definition of "CaAP". In earlier analyses, a CaAP was defined as depolarisation at a recording site just below the primary bifurcation point in the distal dendrite which crosses the threshold of -30 mV. However, this is a rather coarse definition which fails to capture the clear differences in dendritic responses depending on somatic responses (Fig. 18A). Notably, trials in which the somatic response is a triplet have a high prevalence of broad, sustained distal dendritic depolarisation. Other somatic responses tend to be associated with shorter dendritic depolarisation, which although they cross the -30 mV threshold do not seem to represent a sustained calcium influx. Therefore, we do not consider these as CaAPs, and set a new definition of CaAP as having a minimum width above the -30 mV threshold, and a minimum peak depolarisation, with both values depending on its onset time (values are based on a systematic characterisation of CaAP properties, which is discussed in detail in Methods and section 5.3.3).

Using this refined definition of a CaAP, we compare the co-occurrence of CaAPs and different somatic responses. We find that a triplet is almost always associated with a CaAP (Fig. 18B), but that vice-versa, observing a CaAP does not allow prediction of the somatic response (Fig. 18C), as singlet and doublet trials, which represent the majority of simulation trials, may also have CaAPs.

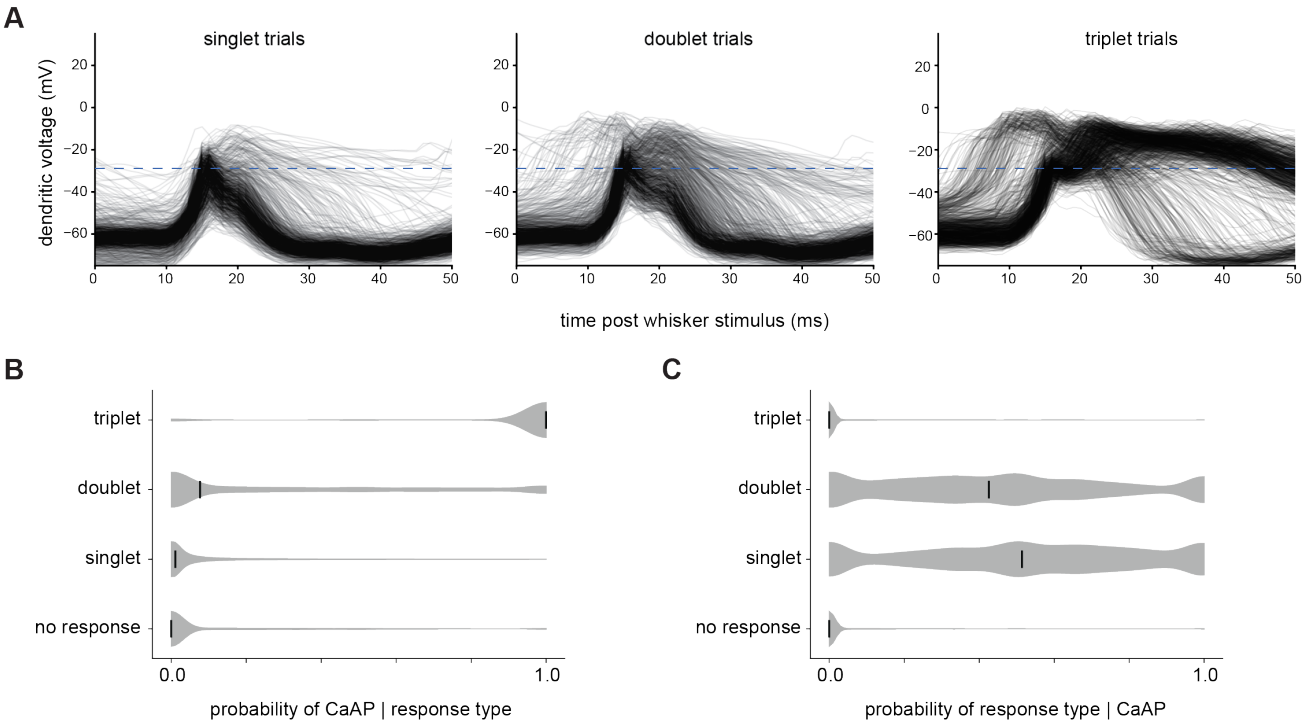

Figure 18. Relationship between dendritic and somatic response. A: 1000 randomly chosen dendritic voltage traces for trials with different somatic responses. Horizontal dashed blue line indicates the previous threshold for a CaAP at -30mV. B: Probability of observing a CaAP depending on somatic response. C: Probability of observing different somatic responses given that a CaAP occurred.

5.3.2 The network origin of CaAPs and triplets

Which inputs on the dendrite are necessary for CaAPs, and therefore triplets? Is the wiring specificity of direct TC inputs to the primary BP functionally relevant for these responses? We selectively removed direct sensory-evoked TC input from different sections along the dendritic tree by soma distance to investigate the location-dependent effect of TC input on CaAPs and triplets. We see that for the somatic response, removing TC input to the primary BP or to the basal dendrites greatly reduces the occurrence of triplets (Fig. 19A). Doublets are largely independent of TC input to the distal dendrite, and rather rely on input to the basal dendrites. CaAPs, on the other hand, are primarily driven by input to the primary BP (Fig. 19B).

So is direct TC input sufficient to drive CaAPs and triplets? If we remove all other sensory-evoked input from intracortical cell types, triplets are completely abolished (Fig. 19E). Likewise, CaAPs are absent, with only a small dendritic depolarisation remaining (Fig. 19C&D). Direct TC input is therefore clearly not sufficient to account for bursts in sensory-evoked output. Given its apparently small contribution to dendritic depolarisation, is direct TC input really necessary for triplets? We next removed all direct TC input, leaving only sensory-evoked input from intracortical (IC) cell types. This manipulation also all but abolished triplets (Fig. 19G), as well as a majority of CaAPs (Fig. 19F). Breaking down the contribution of intracortical excitatory input by cell type, we see that removing sensory-evoked input from any source reduces or abolishes bursts (Fig. S9). We therefore conclude that both direct sensory input from VPM, and indirect sensory-evoked input from intracortical cell types are necessary for sensory-evoked bursts.

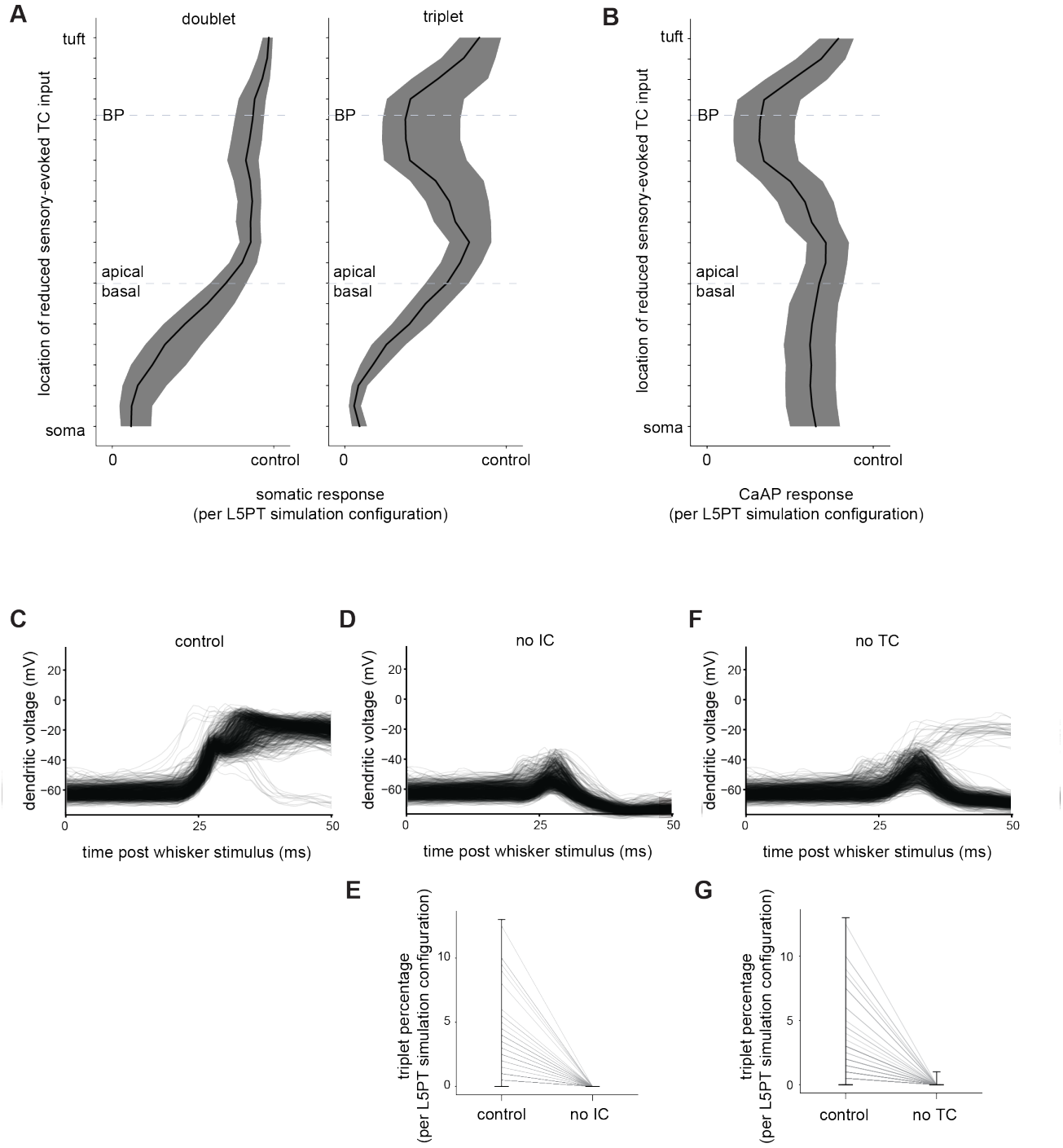

Figure 19. Network origin of triplets and CaAPs. A: Effect of removing direct TC input at different dendritic locations on doublets and triplets. B: Effect of removing direct TC input at different dendritic locations on CaAPs. C: Example dendritic voltage traces from triplet trials. D: Example voltage traces from the same trials, but with sensory-evoked intracortical input removed. E: Effect of removing sensory-evoked intracortical input on triplet rate across L5PT

simulation configurations. F: Example dendritic voltage traces from the same trials as panels C&D, but with sensory-evoked TC input removed. G: Effect of removing sensory-evoked TC input on triplet rate across L5PT simulation configurations.

In addition to sensory-evoked excitatory input, it had been suggested by results from an optogenetics experiment (Fig. 20A, for description of experiment see Methods) that pre-stimulus input may influence the prevalence of triplets. Activating TC synapses optogenetically at 100ms intervals resulted in persistent network activity from intracortical cell types, reflected in a persistently depolarised LFP (Fig. 20C), which was absent in experiments with longer inter-stimulus intervals (Fig. 20B). In this high frequency stimulus experimental condition, we observed an increase in triplets during the course of the experiment, which was not observed in lower frequency stimulation protocols. The more depolarised the LFP, the higher the probability of a triplet response (Fig. 20D&E).

To test the effect of pre-stimulus inputs in the multi-scale model, we replayed existing simulation trials but decreased pre-TC input from intracortical cell types to the basal, trunk and tuft compartments. This was done by removing 50 synapse activations at random from the corresponding compartment in the 20 ms before TC onset (example voltage traces in Fig. 21A). This manipulation reduced the occurrence of triplets (Fig. 21B), but interestingly had only a minor effect on CaAPs (Fig. 21C). It should, then, be possible to do the inverse: increase pre-stimulus inputs to increase the rate of triplets. As a proof of concept, we demonstrate in 3 L5PT simulation configurations that increasing ongoing AP rates of intracortical cell types can create a triplet in a trial which previously did not have one (Fig. 21D). For these 3 simulation configurations, the probability of a triplet response increases with increased ongoing activity, consistent with results from the optogenetics experiment (Fig. 21E&F). However, this manipulation did not have the same effect in all simulation configurations, the reason for which would require further investigation and likely a comprehensive characterisation of the activity of intracortical cells by cell type during this experiment, which was beyond the scope of the present analysis.

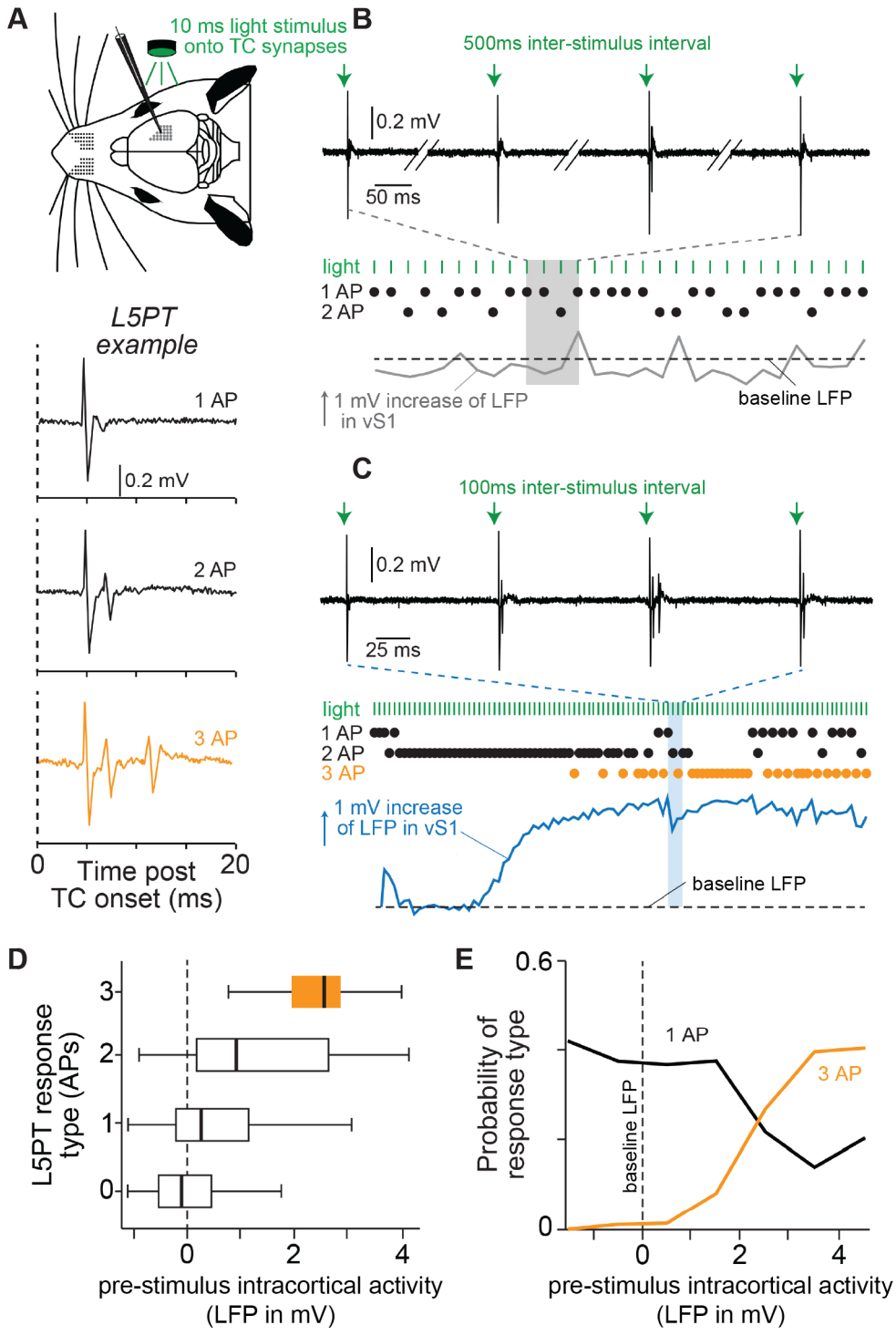

Figure 20. Optogenetics experiment suggests role of pre-stimulus inputs for triplets.
A: Above: schematic of the optogenetic experiments in anaesthetised rats. Below: Example

trials from the L5PT whose responses are seen in panels B&C, showing three TC-evoked response types. B: Somatic APs recorded during 4 example stimuli with a 500 ms inter-stimulus interval. We estimated ongoing intracortical activity by the LFP amplitude in barrel cortex at the onset of each light stimulus, compared to the baseline before the first stimulus. C: As B, for stimulus protocol with 100 ms inter-stimulus intervals, showing increased intracortical ongoing activity. D: Distribution of somatic response types by pre-stimulus activity. E: Probability of singlets and triplets by pre-stimulus activity. *Figure adapted from* (Bast et al., 2021).

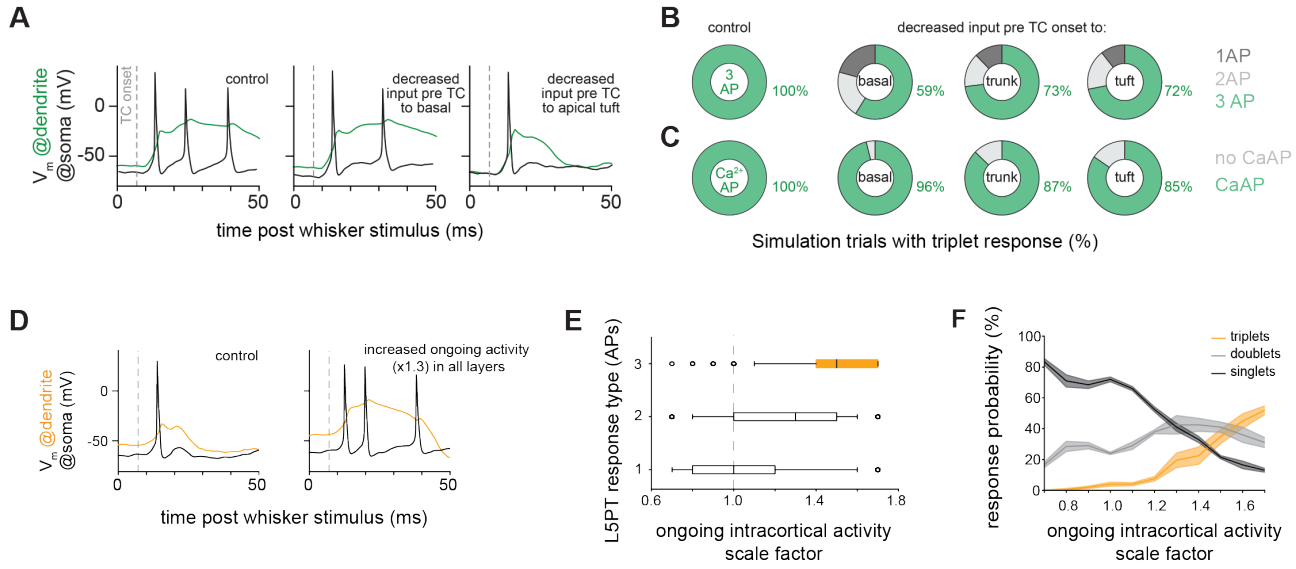

Figure 21. Simulations to investigate the effect of pre-stimulus inputs on triplet rate. A: Simulation examples show how decreased pre-TC input to basal or apical dendrites affects a triplet trial. B: Effect of decreased pre-TC input on somatic responses. C: Effect of decreased pre-TC input on CaAPs. D: Simulation examples show how increased pre-TC input, achieved by scaling the ongoing AP rates of intracortical excitatory cell types, affects a trial which previously did not have a triplet. E: Analogous to 20D from optogenetics experiment, distribution of somatic response types by pre-stimulus activity in 3 L5PT simulation configurations. F: Analogous to 20E, probability of singlets, doublets and triplets by pre-stimulus activity in 3 L5PT simulation configurations. *Panels A-C adapted from* (Bast et al., 2021).

An alternative explanation was proposed for the increased triplet rate in the optogenetics experiments: facilitation of TC synapses as a result of high frequency stimulation. This has been reported experimentally for a subtype of TC synapses, which were observed to be facilitating for the second stimulus in a series (Viaene et al., 2011). We performed simulations of a single stimulus with increased strength of TC synapses mirroring the EPSP amplitude

ratio observed experimentally at an inter-stimulus interval of 100 ms (2.5x increase in uPSP amplitude, Fig. 22A). These simulations showed that it is possible to create triplets by increasing TC synaptic strength (Fig. 22B). However, according to the aforementioned study, subsequent stimuli in a series had a depressing effect. We therefore extrapolated the effect that synaptic strength changes would be expected to have on triplets over a series of 40 stimuli (Fig. 22C). We show that if our findings were due to synaptic facilitation, the proportion of triplet responses would increase sharply at the start of the experiment and then decay. This is not consistent with observations during the optogenetics experiments, where the triplet rate increases during the course of the experiment, mirroring the LFP depolarisation, and triplets were never seen at the second stimulus. This indicates that TC facilitation does not play a significant role in increased bursts responses in this experiment.

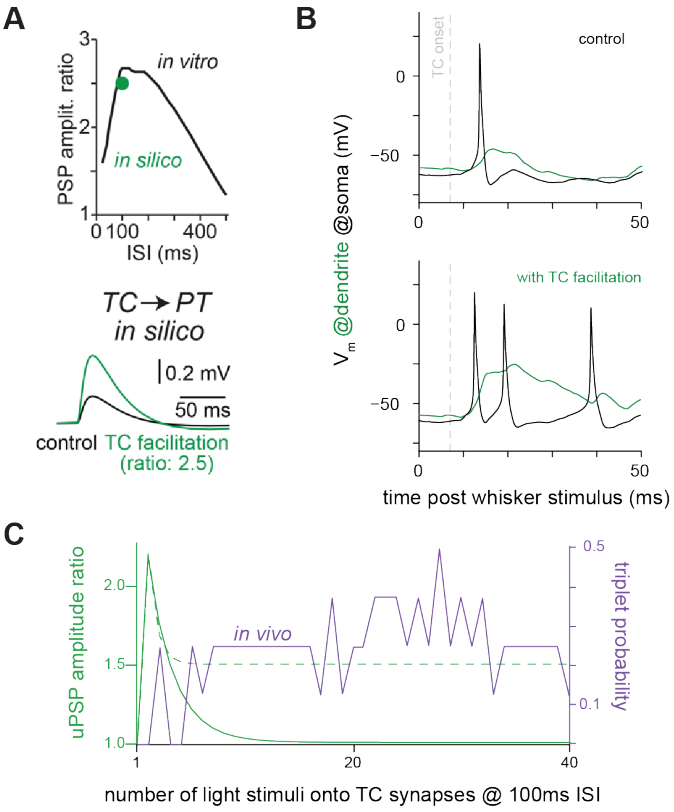

Figure 22. Investigation of TC synapse facilitation as an explanation for optogenetics results. A: Upper: PSP amplitude ratio at different inter-stimulus intervals from (Viaene et al., 2011), lower: we selected a facilitation ratio of 2.5 for our simulations, which corresponds to

the ratio observed at 100ms ISI, as in our optogenetics experiment. B: Simulation example shows that TC facilitation can create a triplet in a trial which did not previously have one. C: Expected effect of TC facilitation (green), extrapolated to 40 stimuli, compared to *in vivo* results from optogenetics experiment (purple). Solid green line: assume that the synaptic strength decays back to baseline, dotted green line: do not assume that the synaptic strength decays back to baseline. *Figure adapted from* (Bast et al., 2021).

5.3.3 Early triplets and CaAPs have different origins

We noticed that in a small subset of trials, the CaAP was not abolished by removing TC or intracortical sensory-evoked input (Fig. 23A). This implies that these CaAPs are entirely initiated by pre-stimulus input, a hypothesis which is supported by the fact that they begin to depolarise, on average, 15.4 ms before the onset of sensory-evoked activity. In order to quantify the occurrence of these non sensory-evoked CaAPs, we separated trials where the CaAP remains when intracortical sensory-evoked input is removed from those in which it is abolished. We then compared the distribution of onset time, peak voltage and width of the CaAPs between these two categories (Fig. 23B). We note that non sensory-evoked CaAPs are earlier, with a higher peak and shorter width, than sensory-evoked CaAPs, and set definitions for classifying CaAPs into these two categories accordingly (values in Methods). We find non sensory-evoked CaAPs in approximately 3.78 % of all trials, compared to 7.85 % of trials with sensory-evoked CaAPs.

How does the network origin of non sensory-evoked triplets differ from sensory-evoked ones? We find that for non sensory-evoked triplets, direct TC input is only relevant to the basal dendrites, unlike sensory-evoked triplets which are also affected by input to the distal dendrite (Fig. 23C). To test our hypothesis that non sensory-evoked CaAPs are driven by ongoing intracortical activity before the stimulus, we repeated the manipulation of decreasing pre-stimulus input to various dendritic compartments (Fig. 23D&E). Decreasing pre-TC input to any dendritic compartment almost completely abolished triplets in these trials. Abolished triplets overwhelmingly become doublets, with the first spike of the triplet absent when pre-TC input is reduced. The occurrence of these CaAPs is reduced by manipulating input to the basal and trunk dendrites, and completely abolished by reduced pre-stimulus input to the tuft.

We therefore conclude that these non sensory-evoked CaAPs are indeed driven by pre-stimulus input, but generating a triplet at the soma additionally requires sensory-evoked input.

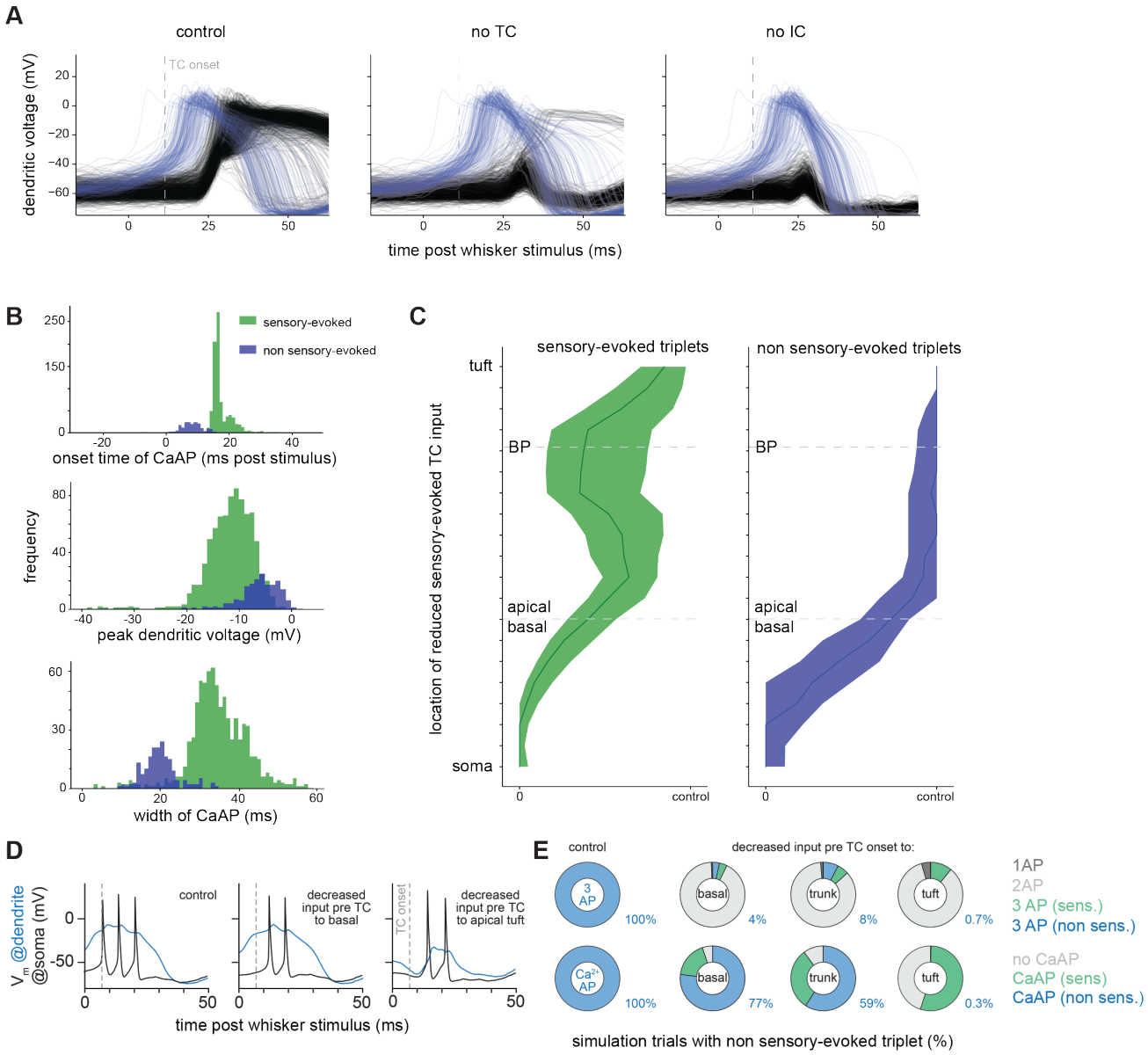

Figure 23. Characterisation of non sensory-evoked CaAPs and triplets. A: Dendritic voltage traces from 1000 randomly chosen triplet trials in control, no sensory-evoked thalamocortical input (no TC) and no sensory-evoked intracortical input (no IC) manipulations. Blue lines indicate trials which remain depolarised above -30mV in the no IC manipulation (non sensory-evoked CaAPs). B: Distribution of CaAP properties between sensory- and non sensory-evoked CaAPs. C: Effect of removing direct TC input at different

dendritic locations on triplets with sensory- and non sensory-evoked CaAPs. D: Simulation examples show how decreased pre-TC input to basal or apical dendrites affects a triplet trial with a non sensory-evoked CaAP. E: Effect of decreased pre-TC input on somatic responses and CaAPs in trials which had a non sensory-evoked CaAP in the control condition. *Panels D&E adapted from* (Bast et al., 2021).

5.3.4 Triplets distinguish awake from anaesthetised animals

The overall prevalence of triplet responses is low in anaesthetised animals (3% triplets in response to whisker deflection, N = 25 L5PTs). It has been reported that during active touch in awake animals, the distal dendrites of L5PTs in barrel cortex are highly active (Xu et al., 2012), and activity in L5PT dendrites in motor cortex has been found to be significantly higher in awake than anaesthetised animals (Murayama and Larkum, 2009). If dendritic activity is specifically relevant for triplets, as our results have suggested so far, we would therefore expect to see an increased rate of triplets, but not other responses, when comparing experimental data from anaesthetised to awake animals. Indeed, we find that compared to passive whisker deflections in anaesthetised rats, triplets are significantly increased in awake animals performing active whisker touches on a stationary object (ANOVA with multiple comparison: $p=0.02$), while overall responses ($p = 0.48$) and doublets ($p = 0.45$) are not (Fig. 24A). Finally, we compared the occurrence of different responses depending on the behaviour of the awake animal (Fig. 24B). We find that the rate of singlets ($p = 0.65$) and doublets ($p = 0.08$) remains unchanged, regardless of whether the animal is quiescent, actively whisking or touching the object. Triplets, on the other hand, seem to specifically encode active touch, with an increase in triplets upon touch ($p = 0.02$) that is not seen during other periods of whisking. This corroborates our finding that doublets and triplets have different origins, with triplets reflecting additional activity in the distal dendrites.

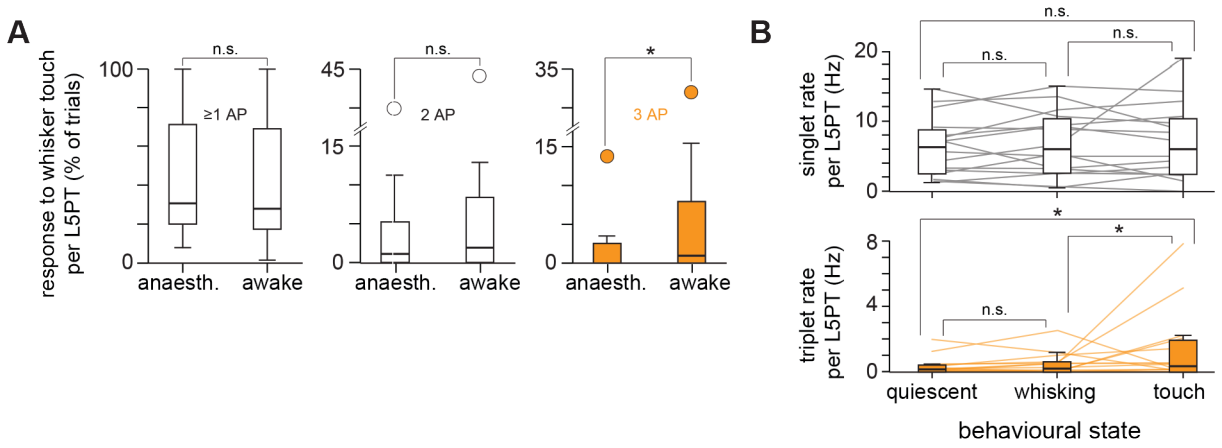

Figure 24. Comparison of anaesthetised and awake responses. A: Probability of any response, doublet response and triplet response to whisker touch in anaesthetised animals (passive whisker deflection) or awake animals (active whisker touch). B: Singlet and triplet rate in awake animals when the animal was quiescent, actively whisking without touching the object, or immediately after a whisker touched the object.

6 Discussion

Using a multi-scale model based on data, I explored structure-function relationships in cortex. At a macro-scale, I used ANNs with cortical features to show that sparse connectivity enables efficient learning in large and recurrent networks. Using a model reduction approach to explain the computation performed by L5PTs, we show that the input-output computation is largely preserved across morphologically and biophysically diverse neurons, and we found that the main source of variability in sensory responses and receptive field shapes is synaptic input. Finally, we investigate the relevance of subcellular scale wiring specificity in the TC-L5PT pathway, finding that this, combined with intracortical sensory-evoked excitation and in certain circumstances prestimulus input, is necessary for bursts observed in sensory-evoked responses.

Comparing our investigations at different levels of detail, it becomes evident that there is a trade-off between biological/mechanistic realism (and therefore how likely it is that our results apply to the system we are trying to understand) and interpretability, i.e. how much understanding we can actually gain. A highly detailed model of a complex system is just as complex a system, and will not necessarily *per se* help in understanding a given phenomenon. Nonetheless, the highly detailed multi-scale model was a necessary starting point for all of our investigations, to ensure that any simplifications or manipulations we performed are still relevant to the biology, and that we aren't just making top-down assumptions.

Using data to gain understanding has presented a general challenge in neuroscience (Jonas and Kording, 2017). Even as large, multimodal datasets become increasingly available from different animals performing different tasks, under the influence of various pharmaceutical, surgical or genetic manipulations, the question remains how to use this data. A popular assumption – or perhaps hope – seems to be "keep on mapping and it will eventually get cognitive" (Gomez-Marin, 2021): keep on collecting data, and an understanding of the brain will naturally follow. Of course, the collection of high-quality experimental data is necessary to expand and validate our understanding of the brain, but it would be short-sighted to suggest

that it will in itself be sufficient. It was already noted over a decade ago that “as data sets grow and become more complex, it will become more and more difficult to analyze and extract conclusions” (Sejnowski et al., 2014). Thanks to advances in experimental techniques and a huge investment of time and resources, we are gradually approaching an era of comprehensive big data neuroscience. For example, in the relatively simple nervous system of *C. elegans*, the connectome has been available for many years (White et al., 1986), and simultaneous whole-brain imaging of neuronal activity has been demonstrated more recently (Nguyen et al., 2016). Thus, the major challenge will become: how can we take rich, complex datasets and use them to inform our understanding of the brain? Here, we show that using extensive data to construct a detailed, biologically realistic multi-scale model allows us to systematically disentangle the contributions of neuron and network features to cortical function.

6.1 Sparse connectivity enables efficient information processing in cortex-like neural networks

I generated and trained ANNs constrained by interpretable features of cortical networks in order to disentangle the effect of structural properties on network function. I find that sparse connectivity is a prerequisite for efficient learning when the network adheres to certain other properties of cortical networks: large recurrent networks, even more so when nodes are either excitatory or inhibitory like cortical neurons. For biological context, the smallest computational unit of the cortex is often reported to be the cortical column, consisting of 10,000 - 20,000 neurons, with connectivity between 10 - 30 % (Meyer et al., 2013). Therefore, the large and sparse networks I investigated here share the same parameters as this elementary computational unit of the cerebral cortex. My finding that sparsity may be beneficial to information processing in cortical networks is mirrored by a recent study about CA3 hippocampus, which showed that the large, sparse networks found in the human hippocampus have a higher memory capacity than the small, dense networks found in mouse hippocampus (Watson et al., 2025).

From a machine learning perspective, many attempts have been made to achieve high performance with sparse networks due to their potential for lower computational and memory demands. A significant approach to generating performant sparse networks is pruning, which begins by training a densely connected network and then removing (pruning) edges which are deemed unimportant throughout the training process, to eventually obtain a sparse network. Pruning approaches are able to generate sparse networks whose performance does not significantly differ from dense networks for both feedforward (Han et al., 2015) and recurrent (Narang et al., 2017) networks, but still require starting with a densely connected network. A related method is rewiring, which begins with a sparse network, and then allows existing edges to be pruned and new edges to be formed during training, while maintaining a constant total number of edges. This was shown to produce performance equivalent to densely connected networks for feedforward architectures (Dettmers and Zettlemoyer, 2019), and in fact resulted in sparse networks which outperformed their dense counterparts for recurrent architectures (Liu et al., 2021). Here I show that, given the right conditions, a sparse recurrent network can outperform its dense counterparts even with fixed sparse connectivity.

Given the extent of non-random wiring observed in the cortex (Song et al., 2005; Udvary et al., 2022), it is somewhat surprising that sparse networks with random connectivity perform as well as they do here, and that I find no notable effects of preserving higher-order connectivity features from cortical networks, like network motif occurrences and degree sequences. There is evidence for random connectivity in some biological neural networks, like olfactory inputs to the mushroom body in *Drosophila* (Caron et al., 2013). This aligns with computational models like the Liquid State Machine, which posits that a sufficiently large population of randomly interconnected neurons can generate a diverse enough set of input representations, even in the absence of learning, for a simple downstream classifier to learn to distinguish between inputs (Maass, 2011). There is some evidence that the same may be true in ANNs: Frankle and Carbin (2019) find that sparse subnetworks exist in randomly initialised densely connected ANNs which, when trained in isolation, achieve at least equivalent performance to the full-sized dense network with less training. This finding suggests that within randomly initialised networks, there exist structures which are inherently

well-suited for learning and/or performing the task at hand, which could explain why my large, sparse networks perform well even with randomly initialised connectivity.

Our results suggest that the integrative properties of nodes (i.e., their activation function) are relevant for determining whether a network will benefit from sparse connectivity. However, a linearly weighted sum followed by a ReLU activation function by no means reproduces the complex input-output computations performed by real neurons. For example, unlike the nodes in these ANNs, synaptic conductance and therefore the activation of biological neurons are stochastic (Rusakov et al., 2020). There are strong parallels between stochastic activity and the practice of node dropout during training in machine learning, as both result in the absence of activity from a varying subset of neurons/nodes. Dropout has been shown to reduce overfitting (Srivastava et al., 2014), a problem which I observed in our large, sparse networks on the CIFAR10 dataset. Furthermore, dropout has been shown to decrease the susceptibility of ANNs to adversarial attacks (Dapello et al., 2021). Adversarial attacks are small changes made to e.g. pixels in an image, which are generally imperceptible to humans but cause a complete misclassification of the example by the ANN – this indicates that ANNs often rely on non-robust features for classification (Ilyas et al., 2019), a key distinction from biological neural networks, which may be resolvable by including dropout. Overall, this suggests that incorporating additional biological details on a single-cell scale may also be fruitful.

Sparsely and densely connected networks differ in their node activity, and therefore in their representations of input. I found that in sparsely connected networks, more nodes participate in the classification than in densely connected networks. It has previously been shown that correlations in in-degrees (i.e., the number of synaptic inputs received by a neuron from various presynaptic populations) are a mechanism to compensate for heterogeneous inputs and enable balanced state dynamics, where a majority of nodes are able to contribute to signal processing (Landau et al., 2016). Here, I show that sparse connectivity could be another way to facilitate a balanced state and reduce quiescence. However, by itself sparsity is insufficient to account for the broad representation typically seen experimentally in cortical

recordings. Biological neural networks may make use of specific synaptic plasticity mechanisms in order to equalise excitation-inhibition ratios and thereby regulate the activity of individual neurons (Xue et al., 2014). I used global backpropagation of error and gradient descent-based optimisation to train our networks, which cannot account for local synaptic plasticity mechanisms. The impact of more local learning rules should therefore also be the topic of further studies.

Why do sparsely connected networks perform better than densely connected networks under these conditions? My analysis of node activations in Dale's networks shows that the outputs of nodes in densely, but not sparsely connected networks are highly correlated at initialisation, a relationship which has also been reported in spiking neural networks (Pernice et al., 2011). I find that this is likely due to an excitation-inhibition imbalance here, the effects of which are mitigated by sparse connectivity. Correlations in activity between neurons in biological neural networks are thought to be relevant for information processing (Shadlen and Newsome, 1998; Averbeck et al., 2006) and learning (Bi and Poo, 2001). However, the extremely high correlations in densely connected networks here suggest a high redundancy in information across nodes, and therefore a very limited effective capacity of the network. This could make it more difficult for the dense network to develop distinguishable representations of its inputs, and thereby explain the delay in learning. It has also been observed that very wide and very deep ANNs often contain many nodes with nearly identical representations, a phenomenon attributed to overparameterisation (Nguyen et al., 2021). Since dense networks have more parameters than sparse networks, this overparameterisation may be a contributing factor to my findings.

It has been speculated that innate behavioural abilities and rapid learning in animals may be facilitated by specific wiring properties in neural circuits that emerge during development (Zador, 2019). As the information capacity of the genome is orders of magnitude too small to encode the connection between each pair of neurons explicitly, it was suggested that wiring rules may underlie the formation of neural circuits in the developing brain. According to this, a bulk of the 'learning' – or more precisely, meta-learning - in biological neural networks would

thereby take place on the timescale of evolution, and changes to the network structure would be implemented in the genome in the form of wiring rules. It is indeed plausible that such wiring rules could inform biological neural networks, as exemplified by a study proposing rules connecting different innexins to form gap junctions in the nervous system of *C. elegans* (Kovács et al., 2020), or another showing that just the structural composition of the neuropil is sufficient to explain a large portion of the non-random connectivity observed in the rat somatosensory cortex (Udvary et al., 2022). To further support this, I here find that even a 'wiring rule' as simple as sparse connectivity can facilitate efficient information processing.

In summary, I show that the degree of connectivity in a cortex-like recurrent network is critical, conditional on the network's size, for efficient information processing.

6.2 Network-neuron interactions underlying sensory responses of layer 5 pyramidal tract neurons in barrel cortex

We developed an approach to reduce a detailed, multi-scale model of L5PTs in rat barrel cortex into analytically tractable models. The reduced models reveal that, in the experimental condition of passive whisker deflection in anaesthetised animals, just three features are sufficient to explain the transformation of synaptic input into AP output: the count of excitatory and inhibitory synapses active relatively close to the soma and in a recent time window, and the time since the previous AP. We show that this input-output computation is qualitatively preserved across L5PTs with diverse morphological and biophysical properties. Consistent with an earlier study (Egger et al., 2020), we predict that the broadness of receptive fields in L5PTs arises from a horizontal spread of sensory-evoked excitation by L6CC neurons. The variability in trial-to-trial and cell-to-cell responses is predicted to arise primarily from variations in synaptic innervation from the network.

A major advantage of the reduced models is their interpretability, which makes the input-output computation easily understandable. Of course, other interpretable single neuron models are already popularly used, like leaky integrate-and-fire models, which are highly interpretable and computationally efficient. However, these models are based on

assumptions about neuronal computation, and have no representation of the constituents of a biological neuron. Our reduced models, on the other hand, start with empirical morphological and electrophysiological measurements. The neuronal computation is *derived* by exposing a biophysically detailed neuron model to synaptic input constrained by *in vivo* measurements of cellular activity in the specific experimental condition. The reduced models can therefore also account for biological variability in dendritic properties, and can explain their effect on input-output computation. They achieve this at a similar computational cost to leaky integrate-and-fire models, and thereby open the door to large-scale simulations of neuronal activity with high biological realism.

However, the reduced models presented here are likely only applicable to the specific experimental condition of passive anaesthetised single whisker deflection. Other recent studies have developed reduced biophysically detailed models which are valid under a wide range of synaptic input conditions, like simplified conductance-based models (Amsalem et al., 2020; Wybo et al., 2021), stacks of linear-nonlinear units (Ujfalussy et al., 2018) or deep artificial neural networks (Beniaguev et al., 2021). However, all of these models are difficult to understand – for example, Beniaguev, Segev and London (2021) find that a 7-layer convolutional neural network is necessary to represent the computational complexity of a single L5PT with NMDA synapses. Instead, the reduced models compromise on generalisability to different experimental conditions in favour of being accurate in a single experimental condition but highly interpretable.

Generating reduced models for other experimental conditions would likely require a different model structure. It has been shown that dendritic compartmentalisation can change depending on synaptic input conditions (Wybo et al., 2019), making it unlikely that the reduced models described here will generalise well to other experimental conditions. For example, performing passive whisker deflections in an awake animal (Takahashi et al., 2020) or allowing the animal to actively move its whiskers to touch objects (De Kock et al., 2021) changes L5PT output from predominantly single APs to high-frequency bursts of APs, which are mediated by active calcium conductances in the distal dendrite. The number of APs in a

response is not well captured by the reduced models here (Fig. S10). Generating reduced models which account for this phenomenon would require a multi-scale model which includes well-constrained spatiotemporal synaptic input patterns and captures the variability in output in this new experimental condition. Finding the new minimal description for the input-output computation under this condition would then be an iterative process, and is likely to require additional information about synaptic input or dendritic properties.

While our reduction approach currently relies on the multi-scale model, the same approach could be applied to future experimental data. It is plausible that large-scale electron microscopy and population-level voltage imaging could provide the required resolution and volume of data to derive an input-output computation. A reduction approach could thereby facilitate interpretation of experimental data, and help to disentangle how neuron and network features contribute to observed activity.

In summary, we show that the input-output computation performed by L5PTs is surprisingly independent of their morphology and biophysical properties. We show that the broadness of L5PT receptive fields originates from innervation by horizontally extensive L6CC axons, and that the heterogeneity of L5PT receptive fields arises from the diversity of the neuropil, due to which each L5PT receives inputs from a different population of presynaptic neurons.

6.3 The origin of bursts in sensory-evoked cortical output

We investigated the origin of bursts in sensory-evoked output of L5PTs, and the relevance of highly specific thalamocortical innervation to the distal dendrites for these responses. Using manipulations to synaptic input in the multi-scale model, we found that bursts are driven by a combination of thalamocortical and intracortical sensory-evoked input, as well as ongoing activity before the sensory stimulus.

We show that doublets and triplets have different origins: while doublets only require TC input to the basal dendrites, distal TC innervation is necessary for triplet bursts, which are almost always associated with a CaAP. I identify an exception to this finding with non sensory-evoked CaAPs, which represent approximately 1/3 of CaAPs across our L5PT simulation

configurations. Non sensory-evoked CaAPs are initiated by pre-stimulus ongoing activity, and these trials therefore do not require distal TC input to produce a triplet response. Nonetheless, the majority of burst responses require the distal dendritic input provided by the wiring specificity of the thalamocortical input pathway. Even though, in isolation, this thalamocortical excitation only has a minor effect on dendritic depolarisation, and is insufficient to cause a CaAP on its own, the fast and reliable activation of distal TC terminals is necessary to provide enough dendritic depolarisation to allow subsequent intracortical sensory-evoked inputs to cause a CaAP and somatic burst firing.

We find that ongoing activity present before the sensory stimulus can modulate sensory-evoked responses. Which information might be encoded in what we have so far described as “ongoing” activity? One component is likely to be the spontaneous, cortex-wide oscillations which are observed in anaesthetised and quiet awake animals (Mohajerani et al., 2010). Surprisingly, in L2/3 of barrel cortex, it has been shown that neurons are *less* responsive to sensory stimuli when they are depolarised by spontaneous waves of activity (Petersen et al., 2003). The shorter, more spatially localised sensory-evoked responses when the cortical region was in the ‘UP state’ (more depolarised) were speculated to improve temporal and spatial discrimination of inputs. Conversely, we here predict that in L5PTs, increased pre-stimulus depolarisation leads to a stronger response with a higher likelihood of burst firing. This is consistent with the experimental observation that dendritic activity and somatic output appear to be strongly correlated in L5PTs *in vivo* (Palmer et al., 2012). In addition to spontaneous oscillations, non sensory inputs to L5PTs likely contain motor information: L5ITs, which provide input in particular to the distal dendrites of L5PTs, significantly increase their activity in response to whisking (De Kock and Sakmann, 2009). I find that non sensory activity can affect the burst rate, most strongly via non sensory-evoked CaAPs. Therefore, L5PTs in barrel cortex could also integrate information from the motor-sensory pathway like whisker motion and position, which may be relevant for interpreting sensory information.

We find that triplet bursts specifically encode active touch in awake animals, with an increase in sensory-evoked triplet responses in awake animals compared to anaesthetised animals.

This is consistent with the finding that anaesthesia disrupts signalling in the apical dendrites (Suzuki and Larkum, 2020), as we predict that calcium activity in the distal dendrites is necessary for burst firing in L5PTs. Disrupting this distal activity by anaesthesia would therefore also impair bursts in cortical output. Furthermore, this likely explains why calcium-mediated activity in the distal dendrites is correlated with conscious perception and perceptual thresholds (Takahashi et al., 2016), and thereby clarifies the role of dendritic calcium signalling in conscious perception: excitatory thalamocortical feedback from POm enhances sensory processing in the barrel cortex (Mease et al., 2016). This is consistent with the Global Neuronal Workspace model (Dehaene and Changeux, 2011), which posits that the basis of conscious processing is an amplification of relevant sensory activity which leads to an “ignition” of a large-scale prefronto-parietal network. However, overcoming the strong depression of POm synapses to kickstart the amplification of sensory activity requires a synchronous volley of bursts from L5PTs (Mease et al., 2017). Such a volley of bursts requires calcium activity in the distal dendrites of multiple L5PTs, so that sensory-evoked activity can elicit triplet responses. The TC-L5PT connectivity which we revealed here is perfectly set up to provide synchronous input to the distal and basal dendrites of L5PTs across a whole column. In combination with sensory-evoked intracortical activity, and with the possibility of encoding non sensory information as well, this highly specific TC wiring enables L5PTs in barrel cortex to respond to sensory input with a synchronous volley of bursts and thereby ignite conscious sensory processing.

In summary, we used the highly detailed multi-scale model and targeted manipulations to reveal the functional implications of thalamocortical wiring specificity, showing that it is necessary for synchronous bursts in cortical output.

7 Abstract

Using a multi-scale model based on extensive experimental data, I explored structure-function relationships in cortex. At a macro-scale, I used ANNs with cortical features to show that sparse connectivity enables efficient information processing in large and recurrent networks, as are found in cortex. I propose that cortex-inspired ANNs may be a useful tool to disentangle the relevance of other properties of cortical connectivity and neuronal function for information processing in a network. Incorporating more biological detail on the neuron and network level, we used a model reduction approach to explain the computation performed by L5PTs following a sensory stimulus. We showed that the input-output computation is largely preserved across morphologically and biophysically diverse neurons, and found that the main source of variability in sensory responses and receptive field shapes is differences in synaptic input. Finally, we investigated the relevance of subcellular scale wiring specificity in the TC-L5PT pathway. We found that this wiring specificity, combined with intracortical sensory-evoked excitation and in certain circumstances prestimulus input, is necessary for bursts observed in sensory-evoked responses which may ignite conscious sensory processing.

8 List of Figures

Figure 1. Cortical column ANN performing MNIST image classification..... 17

Figure 2. Manipulations of cortical column ANNs..... 34

Figure 3. Schematic illustration depicting ANNs with conventional or cortex-inspired structural properties. 36

Figure 4. Effect of network size and sparsity on time-limited training..... 38

Figure 5. The benefit of sparse connectivity is dependent on a recurrent network architecture and a single-cell activation function with a threshold..... 39

Figure 6. Sparse connectivity enables data-efficient training of large and recurrent networks. 41

Figure 7. Sparsely connected networks form distributed, robust representations. 42

Figure 8. Sparse connectivity facilitates efficient training in networks with structural E/I imbalance seen in cortex. 45

Figure 9. Effect of recurrent triplet network motif occurrences on ANN performance..... 48

Figure 10. Effect of degree distribution preservation on ANN performance. 50

Figure 11. Biophysically detailed multi-scale model of whisker deflection evoked responses in cortical pyramidal tract neurons..... 52

Figure 12. Input-output computation of L5PTs upon single whisker deflections..... 55

Figure 13. Input-output computation is robust to morphological and biophysical diversity. . 57

Figure 14. Active nonlinear mechanisms in simulations of passive whisker deflection with biophysically detailed multi-compartmental models. 59

Figure 15. Influence of NMDA on sensory-evoked responses and input-output computation. 61

Figure 16. Reduced models predict contribution of input pathways to sensory responses. 63

Figure 17. Reduced models predict origins of receptive field variability..... 65

Figure 18. Relationship between dendritic and somatic response. 67

Figure 19. Network origin of triplets and CaAPs..... 69

Figure 20. Optogenetics experiment suggests role of pre-stimulus inputs for triplets. 71

Figure 21. Simulations to investigate the effect of pre-stimulus inputs on triplet rate. 72

Figure 22. Investigation of TC synapse facilitation as an explanation for optogenetics results. 73

Figure 23. Characterisation of non sensory-evoked CaAPs and triplets. 75

Figure 24. Comparison of anaesthetised and awake responses..... 77

9 Supplementary Material

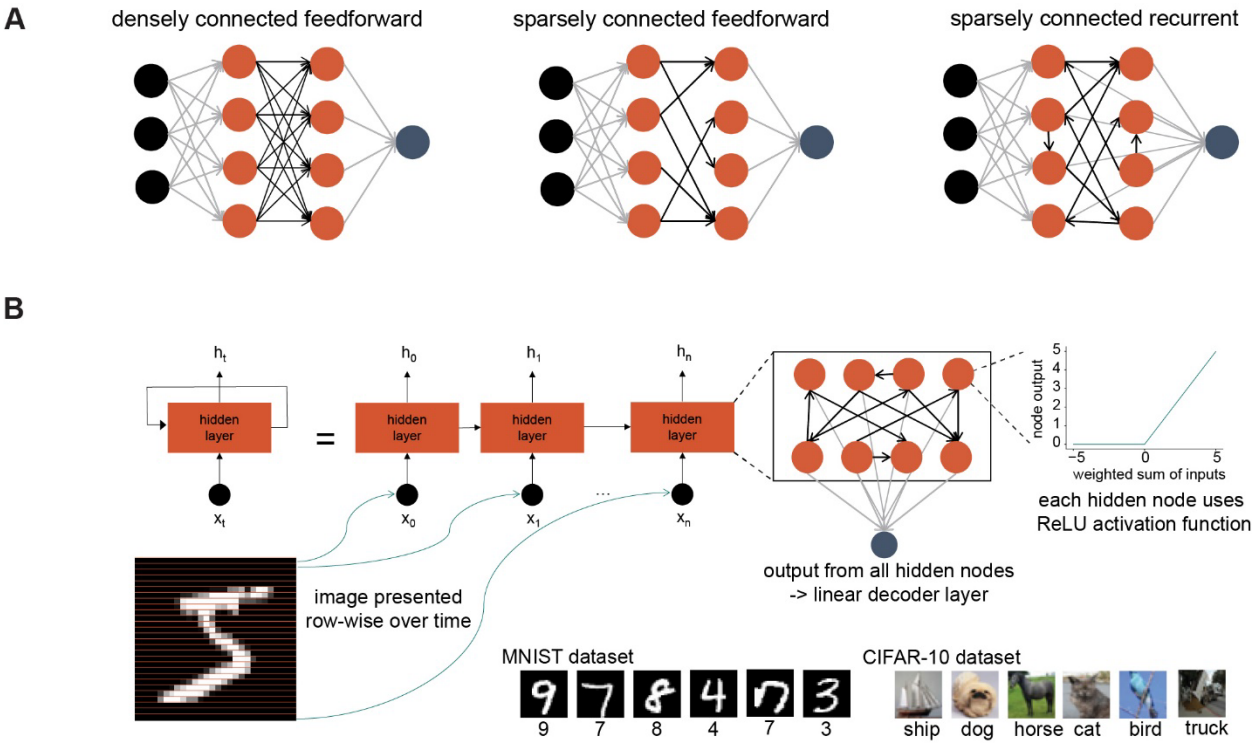

Figure S1. A: Schematic illustration of ANNs with different connectivity properties. Densely connected feedforward: nodes in the hidden layer are arranged into 2 layers. All nodes in the first hidden layer are connected to each node in the second layer by a trainable weight. Sparsely connected feedforward: nodes in the hidden layer are arranged into 2 layers. Some of the nodes in the first layer are connected to some of the nodes in the second layer by a trainable weight, according to the connection probability. Sparsely connected recurrent: nodes in the hidden layer are arranged into a single layer. Within this recurrent layer, any node can be connected to any other node with a trainable weight. The number of connections changes according to the connection probability. B: Modifying image classification tasks (MNIST & CIFAR10) for use with recurrent networks. The image was split row-wise, and at each time step the next row of the image was provided as input to all nodes in the recurrent layer.

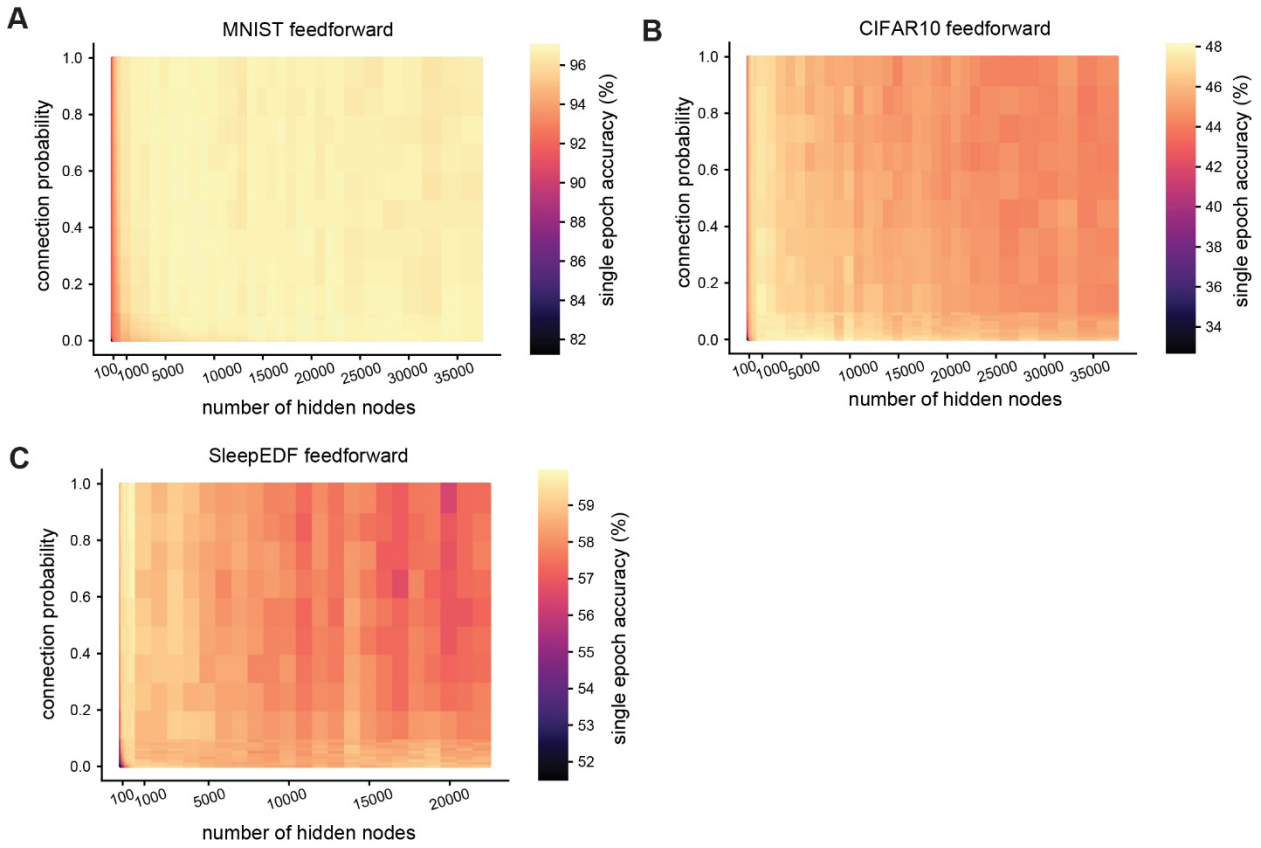

Figure S2. Sparsity confers minimal benefit in feedforward networks. A: Test accuracy after one training epoch on the MNIST dataset for feedforward networks with different numbers of hidden layer nodes and connection probabilities between nodes in the hidden layers. B: Test accuracy after one training epoch on the CIFAR10 dataset for feedforward networks with different numbers of hidden layer nodes and connection probabilities between nodes in the hidden layers. C: Test accuracy after one training epoch on the Sleep-EDF dataset for feedforward networks with different numbers of hidden layer nodes and connection probabilities between nodes in the hidden layers.

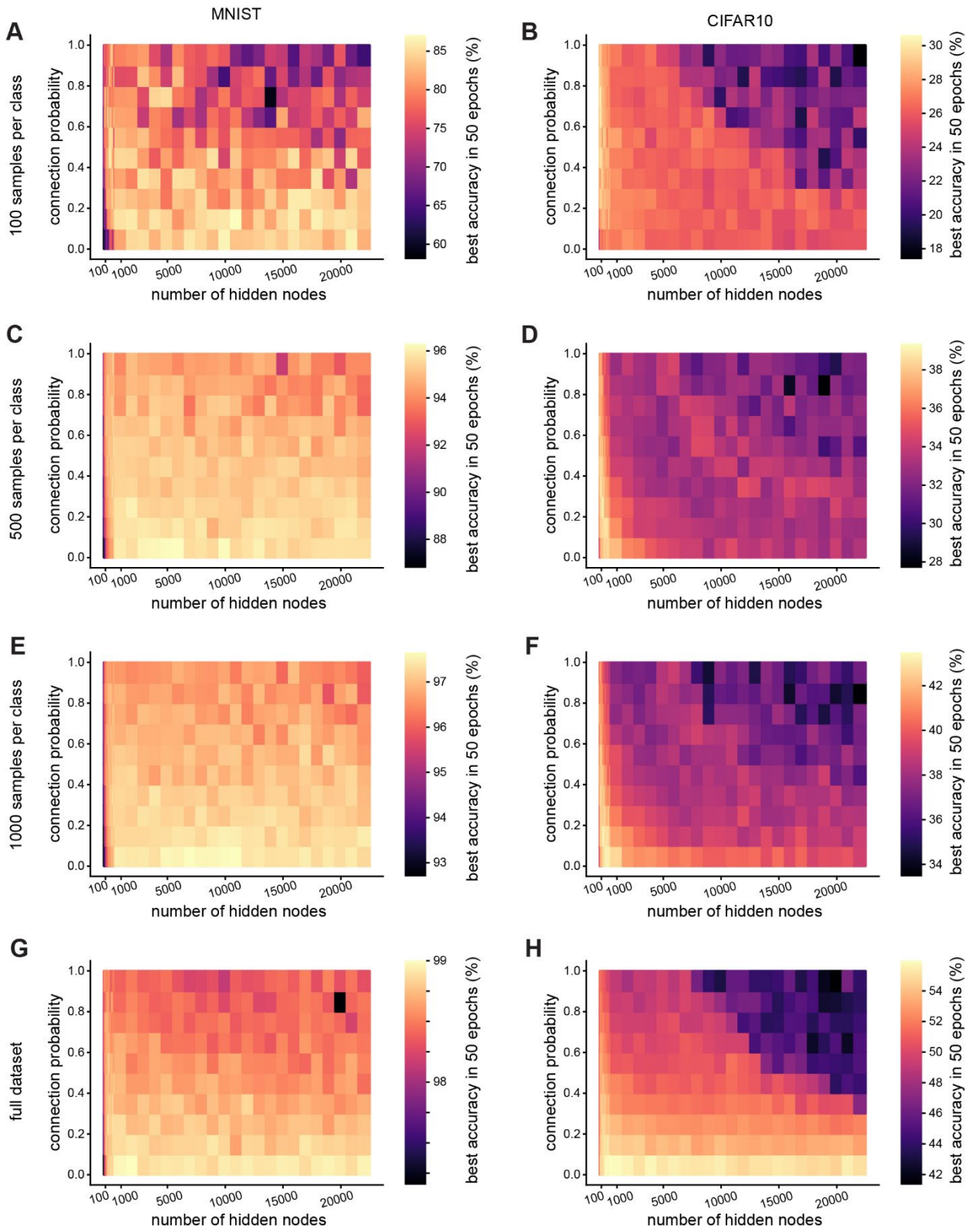
MNIST
CIFAR10
A
B
C
D
E
F
G
H
100 samples per class
500 samples per class
1000 samples per class
full dataset
connection probability
number of hidden nodes
best accuracy in 50 epochs (%)
100
1000
5000
10000
15000
20000

Figure S3. Reduced training data supplement. A-H: Test accuracy after one training epoch on for recurrent networks with different numbers of hidden layer nodes and connection probabilities between nodes in the hidden layers. Left column: MNIST dataset, right column: CIFAR10 dataset. Rows top to bottom: training dataset reduced to 100 examples per class, training dataset reduced to 500 examples per class, training dataset reduced to 1000 examples per class, full training dataset (6000 examples per class for MNIST, 5000 examples per class for CIFAR10).

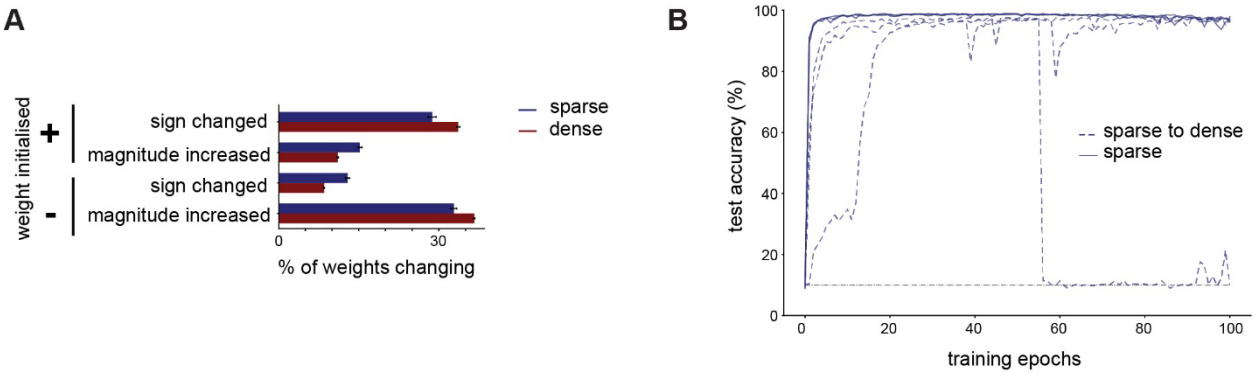

Figure S4. Supplement for Dale's principle figure. A: Changes in weights after one epoch of training a network with 10000 hidden layer nodes with dense (connection probability = 1, red) or sparse (connection probability = 0.1, blue) connectivity on the MNIST dataset. We illustrate weights whose sign changed (from positive to negative or vice versa), and those whose magnitude increased without changing sign. B: Test accuracy over time in large networks (10000 hidden layer nodes) trained on MNIST with fixed excitation and inhibition (11.5 % inhibitory nodes). We compare networks which were initialised with a connection probability of 0.1. In sparse networks, any weight that was initialised at zero according to the connection probability was not trainable, and could not change during training. In sparse to dense networks, although the network was initialised with a sparse connectivity matrix, all weights were trainable, and could therefore change from zero, making the network less sparse during training.

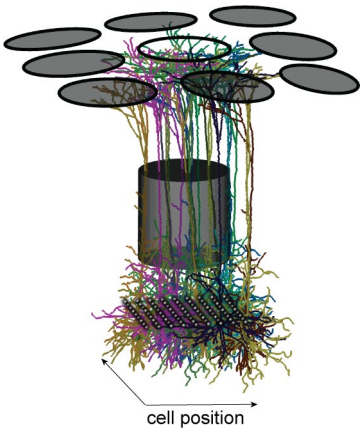

Figure S5. Network embeddings at 81 positions in and around the C2 column of the five *in vivo* reconstructed morphologies.

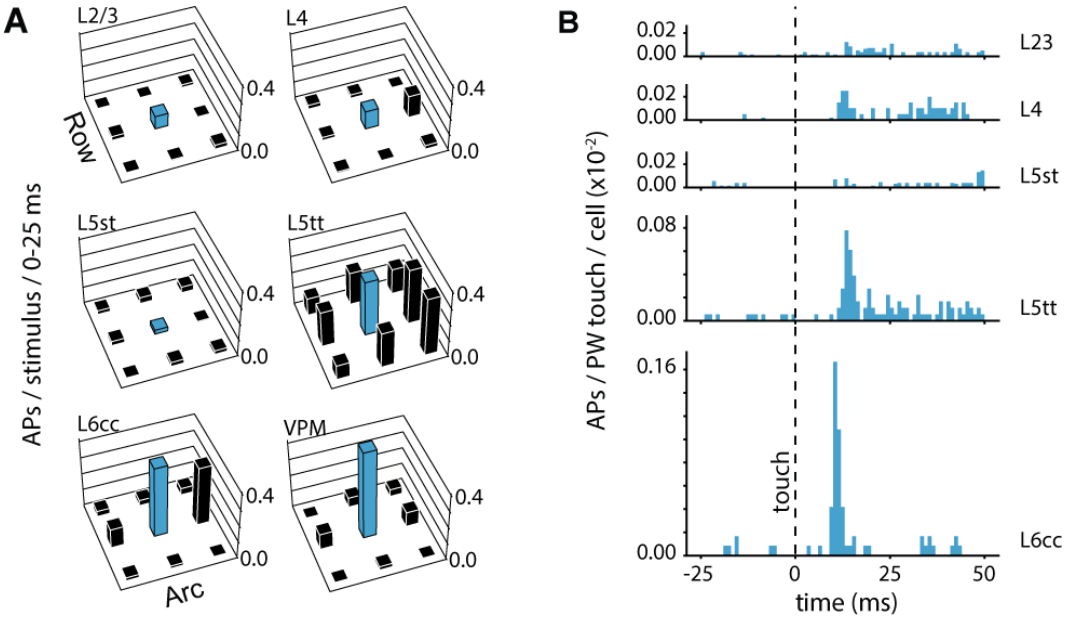

Figure S6. Cell type and stimulus specific activity used to constrain the multi-scale model. A: Average whisker receptive fields of intracortical and thalamic cell types. B: Average post-stimulus time histogram (PSTH) of intracortical cell types.

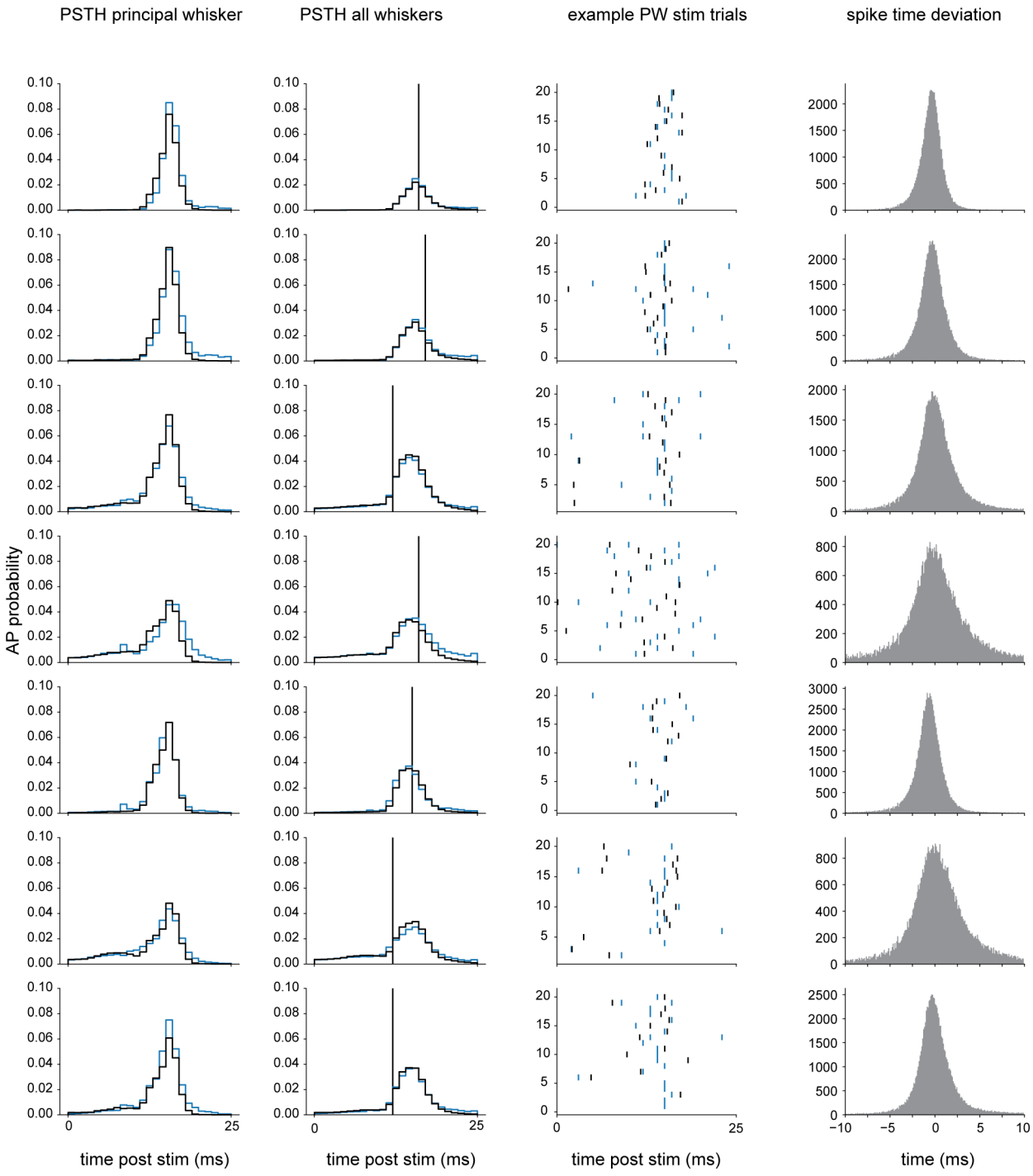

Figure S7. Spike timing in all multi-compartmental (black) versus corresponding reduced models (blue). Rows correspond to each multi-compartmental model. The first row corresponds to example 1 in Fig. 1 and the reduced model in Figure 2. Columns are from left to right: PSTHs of multi-compartmental and reduced models for a PW stimulus, PSTHs of multi-compartmental and reduced models for a PW and the eight SW stimuli (vertical lines

reflect inference time points – i.e., the time point on which the respective GLM was trained), raster plots of 20 example trials with a PW stimulus, deviation in ms between APs predicted by the reduced vs. multi-compartmental model for PW and SW stimuli.

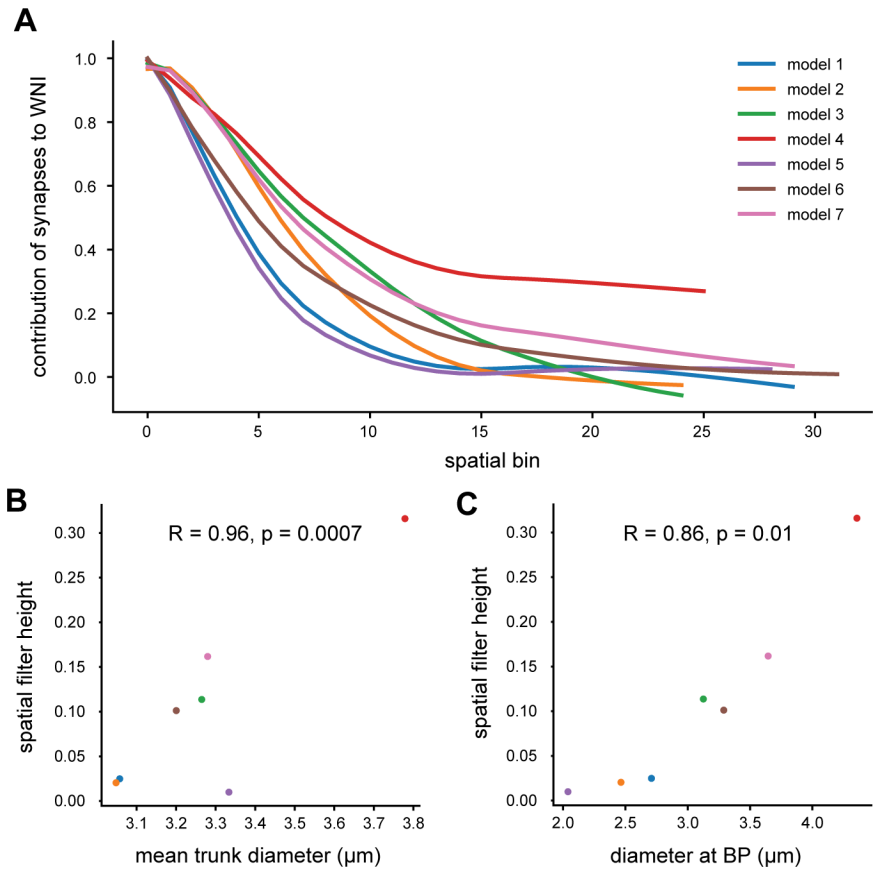

Figure S8. Differences in the decay of the spatial filter can largely be explained by the diameter of the dendritic trunk. A: spatial filters for the seven models, as in Figure 3. B: spatial filter height at the 15th spatial bin (corresponding to a soma distance of 700 to 750 microns) versus the mean trunk diameter. C: as B, but for the diameter at the primary bifurcation point of the neuron.

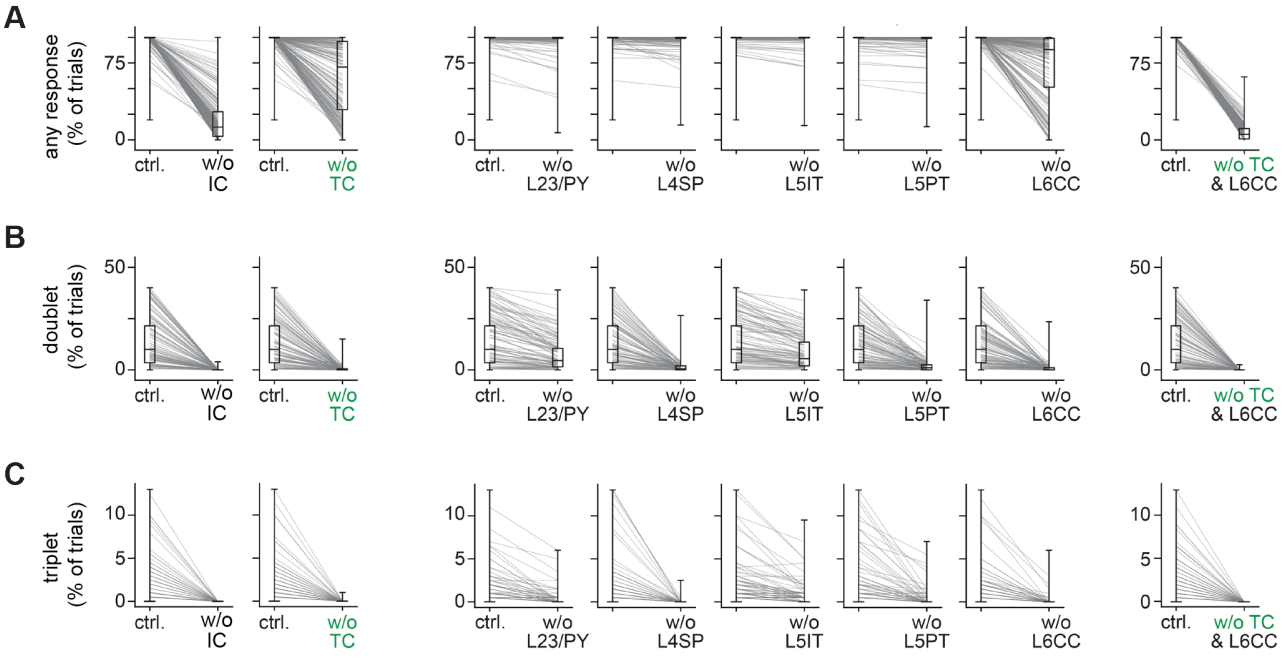

Figure S9. Effect of removing sensory-evoked input by cell type on somatic responses. A: From left to right: effect on response probability across L5PT simulation configurations when removing sensory-evoked input from intracortical excitatory cells (IC), thalamocortical cells in VPM (TC), L2/3 pyramidal cells (L2/3PY), L4 spiny stellates (L4SP), L5 intratelencephalic neurons (L5IT), L5 pyramidal tract neurons (L5PT), L6 corticocortical neurons (L6CC), or thalamocortical and layer 6 corticocortical neurons (TC & L6CC). B: as for panel A, but effect on doublet probability. C: as for A, but effect on triplet probability. *Figure adapted from* (Bast et al., 2021)*.*

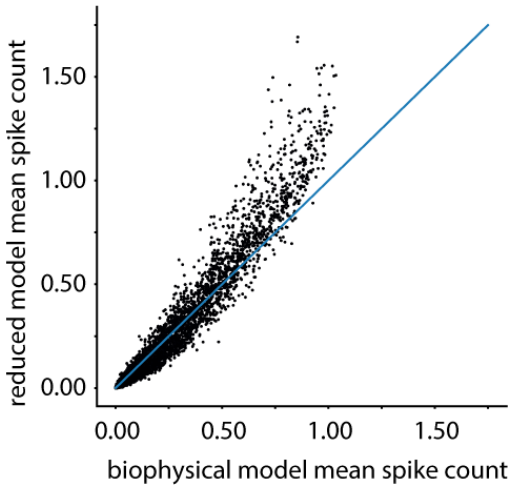

Figure S10. AP counts are not well estimated by the reduced model. Predicted mean number of APs in response to a whisker stimulus for 7 multi-compartmental/reduced models at 81 different positions, for PW and 8 SW stimuli deviates from the multi-compartmental models. In comparison, response probability (See Fig. 13) is very well captured by the reduced models. This indicates that the mechanisms discriminating single AP responses from burst responses are not well captured by the reduced model.

Table S1. Mean test accuracy after one training epoch on the MNIST dataset for recurrent ANNs with different numbers of hidden layer nodes (rows) and different connection probabilities between nodes in the hidden layer (columns).

	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
100	44.6	47.2	50.7	53.5	56.1	55.8	58.4	59.8	61	60	63.9	66.7	68.5	72	72.4	74.1	74.1	76.3	76.9
200	48.6	51.5	58.1	60.2	62.8	64	65.3	65.2	64.8	66	67	73.1	76.5	80.4	81.8	83	83.8	83.1	84.4
300	51.1	58.1	62.2	64.9	66.9	67.6	68.9	69.4	69.3	69.2	75.6	80.1	83.1	85.4	85.1	86	83.6	85.5	84.1
400	51.2	58.2	63.4	67.4	70.5	70.7	71.9	72.3	72	72.2	81.4	85.5	86.8	86.3	87.1	88.2	87.2	87.3	83.8
500	54.4	61.4	67	70.7	71.3	72.7	72.5	74.4	75.8	74.3	84.8	86.7	88.1	88	86	87.9	85.1	88	87
600	55	63.7	68.9	71.9	72.8	72.5	73.7	74.1	76.8	80.3	86.7	88.1	88	89.1	87.4	87.9	86.9	85.9	83.1
700	53.5	63.3	70.4	72.2	73.2	74.1	74.3	76.9	78.8	81.4	87.5	88.8	89.5	88.6	87.4	87.7	86.1	86.6	86.7
800	55.8	64.6	70.4	74.7	76.2	76.4	77.2	79.7	82.1	83.2	88.2	88.7	88.7	88.3	87.3	88.1	85.8	86.6	82.7
900	56.6	66.8	72	74.6	76.6	78.2	80.4	81.9	83.7	84.8	88.9	89.2	90.1	89.4	88.1	86.5	86.2	87.1	85.3
1000	55.7	67.5	73.3	75.6	76.4	78.4	81.5	82.9	83.5	85	88.8	88.8	87.2	89	86	87.7	88.1	85.3	85.6
2000	64.3	75.2	79	83.7	86.3	87.6	88.3	88.4	89.3	88.9	89.9	90.2	89.9	89.1	88.9	87.6	85.6	84.9	84.2
3000	69.1	76.5	81.9	85.9	88.8	89.2	89.3	89.4	89.8	89.7	90.1	89.2	88	85.8	86.1	86.7	84	85.8	83.5
4000	73.9	80.5	84.6	87.2	88.3	89.6	89.4	90.2	90.3	90.6	89.5	89.6	88.1	87.3	87	86.5	84.9	79.9	85.3
5000	76.3	83.8	88.1	88.9	90.4	90.6	90.9	90.7	90.3	91.1	91	89.7	87.6	86.4	86.3	86.5	84.7	83.4	79.2
6000	76.5	84.9	87.9	89.2	89.6	90.4	90.9	90.4	91	90.2	89.9	87.9	89	86.4	85.6	84.2	82.6	79.8	80.8
7000	77.9	85.8	88.8	89.8	90.6	90.5	91.2	91.1	90.8	90	90.4	87.3	88.6	88.6	87.4	86.9	84	82.6	79.7
8000	80.3	87.2	89.3	90	90.7	91.1	90.9	91.1	90.7	90.8	89.3	88.8	85.5	87.9	86	83.2	83.6	83.1	80.9
9000	80.8	87	89.3	91.2	90.6	90.2	89.6	90.6	89.9	89.4	90.2	86.6	84.9	87.8	83.9	79.7	84.4	85.2	78.5
10000	80.9	87.9	89.6	90.1	90.5	91.1	90.3	90.6	89.3	90.8	88.4	89.5	86.5	85.7	85.7	81.4	83.2	82.9	80.3
11000	82	89.2	90.2	90.8	90.4	90.5	91.1	91.1	91.2	91	90	89.8	86.1	83.4	87.6	86.3	84.8	79.3	81.5
12000	84.1	88.9	90	90	90.7	90.7	90.9	90.8	90.6	90.3	90.1	87.9	89	84.9	83	83.6	83.7	78.2	78.6
13000	85.1	88.8	89.8	90.3	90.5	92.1	91.3	91.7	90.4	90.5	89.7	89.1	87.8	84.5	85.4	82	83.7	82.6	80.6
14000	85.3	89.4	90.9	90.5	90.2	91.2	91	91.1	90.3	90.7	88.2	88	88.1	86.8	86.1	82.3	78.6	79.9	73.6
15000	87.3	89.5	90.8	91.1	91.1	91.4	91.2	90.4	91.4	91.1	90.2	88.5	88.4	86	84.4	81.2	77.1	80.2	78.3
16000	86.3	89.3	91.2	90.6	90.6	91.1	90.9	90.4	89.3	90.4	88.9	87	86.3	86.5	82.8	81.4	82.2	76.3	78.9
17000	87.2	90.1	90.3	91.1	91	90.5	91.2	90.4	90.2	91.1	89.4	89.6	85.5	84.5	82.4	80.6	82.5	76.7	77.5
18000	87.5	90.6	90.8	90.6	91	91.2	91.2	90.4	91	90.3	90.3	88.1	88.4	86.3	83	83.6	80.3	76.8	77.8
19000	87.1	89.5	91	90.7	90.3	90.4	91	91.1	90.6	90.6	89.3	87.6	85.5	86.1	84.7	80.6	81.7	77.8	69.8
20000	88.3	89.7	90.5	91.3	91	89.8	91.2	90.9	90.8	90.3	90	87.5	86.1	83.2	82.6	79	75.9	79.7	77.5
21000	88.8	89.9	91	90.5	90.9	90.3	89.7	89.4	90.2	91	90.1	88.8	85.4	88.8	80.5	79.2	79.7	73	79.6
22000	88.5	90.7	90.7	90.6	90.7	91.4	91.3	89.6	90.6	90	89.2	86.5	86.4	82.9	83.2	81.9	82.5	75.7	76.3
23000	88.4	90.3	90.3	90.7	90.6	90.7	91	90.5	89.3	90.1	90.2	86.4	85.2	83.1	81.6	80.5	79.5	79	71.1
24000	88.9	91	90.7	91.6	90.5	91.8	91.5	90.1	90	89.7	89.2	88.2	88.2	87.2	83.2	78.9	77.9	77.7	77.2
26000	89.5	91	91.3	91.2	91.6	91.1	90.6	90.7	89.7	90.7	89.3	87.7	87	86.9	85.9	82.6	83	79.6	77.4
28000	88.9	91.2	91.4	90.7	91.2	90.9	90.5	89.8	90.1	90.6	89.1	88	84.9	85.5	82	78.3	70.3	77.7	76
30000	89.6	90.7	91	90.1	91.3	90.2	90.5	90.1	89.9	90.2	89.2	86.1	85.5	85.7	83.5	81.6	81.2	79.3	75.9
32000	89.5	91	91.3	90.1	91.4	91	90.2	90.5	90.5	89.9	88.3	85.6	84.1	84	81.1	84.9	76.6	77.2	72.8
34000	89.7	91.1	91.8	91.1	90.5	90.9	91.1	90.1	91.1	90.9	88.7	85.5	85	86.1	83.2	80.4	80.1	76.5	66.4
36000	90.2	90.6	91	91.2	90.5	90.3	89.4	90.9	90.2	89.4	87	86.6	88.6	84.3	84.2	84.1	79.4	76.6	74.4
38000	89.7	90.7	90.5	91.2	91.4	90.3	89.2	90.6	90	90.8	90.1	87.7	86.9	83.9	81.6	81.4	77.2	73.7	76.1

Table S2. Mean test accuracy after one training epoch on the CIFAR10 dataset for recurrent ANNs with different numbers of hidden layer nodes (rows) and different connection probabilities between nodes in the hidden layer (columns).

	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
100	23.6	23.9	24	24.2	24.5	24.7	25.1	25.4	25.9	26	28.5	29.9	30.5	30.6	31	30.6	30.7	30.3	30.8
200	24	24.4	25	25.6	26.4	27.1	27.5	27.8	28	28.1	30.3	31.1	31.6	30.9	31.4	31	30.9	31	30.2
300	24.3	24.9	26	26.7	27.4	27.7	28.1	28.5	28.8	29.1	31	32	31.6	31.7	30.6	30.7	30	30.4	30.2
400	24.7	25.8	26.9	27.4	27.8	28.1	28.6	29	29.5	29.8	31.6	31.1	31.1	31	30.7	30	28.8	29.5	29.2
500	25	26.1	27.3	27.8	28.2	28.6	29.1	29.5	29.8	30.4	31.3	31.3	30.7	30.7	29.6	30.3	28.6	28.9	29.7
600	24.8	26.4	27.5	27.8	28.4	28.7	29.4	30	30.4	30.6	31.4	30.8	31.2	30.6	29.7	29.1	29.2	29.2	29.1
700	25.2	26.8	27.7	28.1	28.6	29.3	29.8	29.9	30.6	30.6	31	30.3	30.9	30	29.7	29.9	29.5	28.9	26.7
800	25.3	27.1	27.8	28.2	28.8	29.4	29.7	30	30.5	30.8	31.3	29.9	30.1	30	29.8	29.2	29.5	29.1	28.9
900	25.5	27.4	28.2	28.5	29.1	29.9	30	30.3	30.4	30.4	31	30.3	29.6	29.3	29.6	29	29.1	28.7	28.4
1000	25.4	26.9	27.6	28.5	29.5	30.5	30.5	31.1	30.7	31	30.4	30.6	30.1	29.8	29.6	29.1	29.1	28.9	29
2000	26.6	28.1	28.7	29.5	29.3	30.1	29.8	30.3	30	30.7	29.6	29.6	29.7	29.6	29.3	28.9	28.4	27.8	28.3
3000	27.6	28.7	29.4	29.9	30.3	30.6	30.6	30.4	30.6	30	30.3	29.4	29.3	28.8	28.6	28.4	27.7	28	28.6
4000	27.7	28.8	29.5	29.8	29.5	30	29.5	30.2	30.5	30.2	29.7	28.9	28.8	28.3	27.9	27.4	28.4	28	28.1
5000	28	28.8	29.7	30.3	29.6	29.8	30.4	30	30.5	30.6	29.5	29.2	28.8	28.1	28.1	27.3	27.6	28.5	27.4
6000	27.4	28.1	29.3	29.2	29.2	29.4	29.5	29.6	29.5	30.1	29.6	29.2	28.9	28.8	28.6	29	28.6	27.9	27.7
7000	28.1	29.1	29.5	29.9	30.3	30	30.3	29.5	29.8	28.9	29.9	28.9	28.7	29.3	28.8	30	29.4	27.5	25.3
8000	27.7	28.2	28.6	29.2	29	29.8	29.2	29.3	29.4	29.6	30	29.3	28.1	28.4	29	28.8	26	22.8	22.3
9000	28	28.8	29.3	29.6	29.8	30	29.6	29.9	30.2	30.1	29.2	29.2	28.3	28.3	27.9	27.5	26.3	25	22.6
10000	27.6	29.1	29.1	29.7	29.4	30.2	29.8	29.4	30.1	30.2	29.3	27.8	28.4	29.8	29.2	24.9	22.2	17.4	22.6
11000	27.4	28.3	29.4	29.6	29.9	29.7	30	29.8	29.2	29.3	28.6	28.4	29.2	29.5	27.7	25.4	22.4	20.1	24.1
12000	27.9	29.2	29.4	29.8	29.9	29.9	30.4	28.6	28.9	29.5	29.4	29.5	29.4	27.8	25.7	22.6	20.4	20.6	21.9
13000	27.7	28.7	29.9	29.7	30	29.7	30.1	29.8	29.9	29.4	28.9	29	28.7	29.6	24.1	21.1	22.1	20.9	16.1
14000	28	29.1	29.7	29.1	29.9	29.9	29.6	29.6	29.7	30.1	29.7	29.6	29.4	26	22.8	20.1	20.4	19.8	20.5
15000	28.2	29	29.6	29.9	29.7	30.3	30.1	30.1	29.8	29.4	29.5	29.4	28.7	23.3	21.8	22.3	23.1	22.4	19.4
16000	27.9	28.8	29.3	29	29.9	29.6	29.8	29.2	29.4	30.1	29.1	29.2	29.3	23.5	22.1	24.3	19.9	20.7	21.2
17000	27.7	29.1	29.6	29.9	30.3	30.2	29.4	30.3	29	29.6	28.4	29.3	28.4	24.4	24.8	24.8	23.4	22.7	20.7
18000	28	28.7	29.1	30	29.3	29.4	29.3	29.7	29.4	29.6	28.9	29.4	26.6	23.6	24.2	22.3	20.7	21.6	19.5
19000	27.9	29.6	29.7	29.5	29.9	29.5	29.6	29.7	29.5	29.1	29	29.8	27.6	23.5	24.3	21.3	22.8	22.1	20
20000	27.9	29.4	29.6	29.7	30	30.2	29.8	30.2	29.6	29.7	29.2	29.8	25	25.1	24.4	19.6	19.1	19.9	17.4
21000	28.1	29.4	29.8	30.3	29.7	30.5	30	29.9	29.5	30	29.3	29.7	22.3	23.9	23.9	20.5	21.4	19.6	20.2
22000	27.9	28.9	29.6	29.7	29	30.1	29.7	29.4	29.4	30.3	30	29.5	23.4	25.2	24.1	21	19.8	19.8	19.8
23000	27.8	29	29.3	29.2	29.5	29.2	29.5	29.6	29.4	30	29.9	29.2	23.2	24.1	22.8	21.7	19.9	19	19.8
24000	28.1	28.9	29.6	29	29.3	29.9	29	30.1	28.6	30	29.8	28.8	25.7	25.1	22.5	21.6	20.3	19.7	18.5
26000	27.4	29.2	29.8	29.7	29.8	29	29.8	29.3	29.9	29.5	30.5	26.1	24.6	22	20.7	18.5	20.1	20.1	12.7
28000	28	29.1	28.7	29.2	30	29.6	29.1	30.1	29.7	30.2	29.7	22.7	23	22.4	21.5	19.4	20.1	13.5	13.1
30000	28.1	29.3	29.7	29.9	29.7	30	29.8	29.5	29.9	29.7	30.5	23.9	24.7	23.6	20	21.8	18.4	16	10
32000	27.5	28.7	29.4	29.7	29.2	29.3	28.7	29.3	29.2	28.7	28.7	25.1	23.2	19.7	21.8	18.1	13.9	11	10
34000	27.9	29.1	29	29.6	28.8	29.1	29.3	28.8	29.2	28.8	29.3	24.7	25.4	21.9	19.3	19.4	16.8	10.6	9.8
36000	28.1	28.8	29.2	29.3	29	29.4	29.1	29.9	29.8	29.6	26.2	24.4	21.7	23.3	19.2	16.6	11	11.3	10
38000	27.8	29.2	29.6	29.8	29.6	29.9	29.5	30	29.7	30.1	24	21.9	23.3	21	18.7	13.1	9.9	10	10

Table S3. Mean test accuracy after one training epoch on the Sleep-EDF dataset for recurrent ANNs with different numbers of hidden layer nodes (rows) and different connection probabilities between nodes in the hidden layer (columns).

	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
100	48.5	48.8	49.3	50.3	51.8	53	53.3	52.9	54.2	53.2	53.5	51.7	50.6	51.1	51.1	50.7	50.7	51	50.8
200	49	51	53.3	53.4	54.8	54.5	54.1	53.7	54.1	53.3	51.3	51.1	50.5	51.3	50.5	51.4	50.9	50.4	51.2
300	49.7	54.3	54.9	54.9	54.1	53.4	53.7	53.4	52.9	52.5	50.8	51.9	51.6	51.3	50.6	50.6	50	48.3	50.5
400	51.4	55.1	55.5	54.2	54.3	55.8	53.3	53.1	53.5	52.3	52.2	51.6	50.9	50.8	51.5	51.5	50.3	50.1	50.3
500	52.7	55	55.9	55	54.8	54	52.6	51.8	52.2	52.8	51.1	50.8	51.6	51.4	50.7	51.5	49.6	50.7	50.6
600	53.4	54.5	55.1	53.6	53.6	52.7	53.1	52.2	52.3	48.3	51.6	50.4	51	50.9	51.4	51.3	51	50.2	50.1
700	53.8	55.6	53.8	54.6	53.2	52.7	51.7	51.7	51.3	51.8	51.9	50.3	51.7	51.6	51.3	50.3	51.7	51.1	50.1
800	54.2	55.9	54.8	52.4	53.2	53.1	51.8	51.5	52.8	51.5	50.2	50.5	50.9	50	49.8	49.8	50.9	50.4	50.7
900	55	56.4	54.3	54.3	53.1	52.3	52.5	51.9	52	51.7	51.9	50.5	51.4	50.5	50.6	52	50.5	50.6	49.4
1000	54.8	54.1	54.4	52.8	52	51.7	52.3	51.9	51	53.2	50.8	51.3	51.6	51.1	51.1	50.3	49.8	50	48.8
2000	55.2	54.1	53.1	52.2	51.2	52.4	51.5	52.5	51.3	52.3	52.2	50.8	50.4	50.3	49.9	48.7	50.3	47.3	50
3000	56.2	52.1	52.2	53	51.5	51.7	50.2	52.3	52.9	50.9	50	50.1	51.3	51.2	50.7	50.5	48.2	50.6	49.9
4000	54.9	53.8	52.4	52.3	51.2	51.5	52.5	52.9	51.6	50.3	50.8	51.7	51	51.2	49.5	50.6	49.6	50	48.8
5000	54.8	51.3	52	51.9	52.1	52.1	51.5	51.5	51.9	52.3	52.1	50.7	51.9	50.3	49.8	49.4	45	36.8	32.3
6000	53.6	53	50.7	52.7	52.2	51.3	51.8	52.7	51	52	52.2	50.8	51.3	51.2	50.5	45	40.5	29.1	21.3
7000	53.8	51.4	52.4	51.2	50.7	51.7	51.5	50.5	50.9	51.3	51.4	51.7	51.2	49.4	40.3	18	17.6	17.6	10.3
8000	54.3	53.6	52.1	50.3	50.5	51.6	51.2	50.2	51.6	51.6	51.3	51	49.7	48.5	25.8	14	10.3	10.3	10.3
9000	54.6	52.7	52	52.3	52.2	52.2	51.7	51.1	52.5	51.7	51.5	52	51.6	41.7	28.7	10.3	10.3	10.3	10.3
10000	55.1	53.2	51.8	52.5	51.8	51.5	52.2	51.8	52.2	51.6	51.7	50.6	49.1	28.7	10.3	10.3	10.3	10.3	10.3
11000	55	51.8	52.7	51.6	51.9	50.9	50.8	51.8	50	51.2	50.7	51.7	37	18.1	10.3	10.3	10.3	10.3	10.3
12000	54.9	50	51.4	51.9	52.6	52.6	51	52.1	51.4	51.6	50.8	48.2	25.3	10.3	10.3	10.3	10.3	10.3	10.3
13000	54.3	51.6	52.2	51.6	50.9	51.3	50.1	51.7	51	49.5	51.9	50.2	29.8	14.9	10.3	10.3	10.3	10.3	10.3
14000	54.7	52.8	50.5	51.1	50.1	51.3	52	51.1	50.3	52.1	52.1	45	17.6	10.3	10.3	10.3	10.3	10.3	10.3
15000	54.3	51.6	51.4	52.8	51.6	52.8	51.8	51.1	52.2	52.4	51.2	45.6	17.6	10.3	10.3	10.3	10.3	10.3	10.3
16000	53.5	53.1	52.4	50.1	50.8	50.6	50.5	51.4	51.2	49.7	51.6	28.9	10.3	10.3	10.3	10.3	10.3	10.3	10.3
17000	54.1	51.2	51	51.5	53.3	48.5	52.1	52.4	48.4	51.7	52	37.5	17.6	10.3	10.3	10.3	10.3	10.3	10.3
18000	55	52.4	53.3	51.8	51.9	52	51.1	51	51.3	52.4	52.6	21.8	14	10.3	10.3	10.3	10.3	10.3	10.3
19000	53	51.4	50.7	51.2	50.9	50.8	50.5	51.2	51.8	50.9	50.7	25.3	10.3	10.3	10.3	10.3	10.3	10.3	10.3
20000	54.2	52.4	52.3	50.5	52.2	51.2	49.8	51.4	51.7	50.7	50.8	15.1	10.3	10.3	10.3	10.3	10.3	10.3	10.3
22000	54.7	51.7	52	51.7	50.6	52.2	51.8	51.5	51.5	52.6	49.3	10.3	10.3	10.3	10.3	10.3	10.3	10.3	10.3
24000	54.1	52.7	51.5	51.5	51.3	52.3	51.6	50.6	51	51.7	33.4	10.3	10.3	10.3	10.3	10.3	10.3	10.3	10.3
26000	53.4	51.6	52.2	51.4	51.8	51.6	52.3	51.4	50.1	51.1	29.3	10.3	10.3	10.3	10.3	10.3	10.3	10.3	10.3
28000	55	53.1	50.9	52.3	49.9	52	51.6	51.6	51.3	50.4	17.6	10.3	10.3	10.3	10.3	10.3	10.3	10.3	10.3
30000	53	51.7	52	52	51.4	50.8	51.5	51.7	51.7	53.1	10.3	10.3	10.3	10.3	10.3	10.3	10.3	10.3	10.3
32000	54	51.9	52.4	51.7	52.3	49	51	51.3	51	51.2	10.3	10.3	10.3	10.3	10.3	10.3	10.3	10.3	10.3
34000	53.2	52	52.8	52	51.8	50.4	51.9	51.2	51.8	52.4	10.3	10.3	10.3	10.3	10.3	10.3	10.3	10.3	10.3
36000	53.9	53.1	51.7	50.8	52.6	51	50.2	52.4	51.7	52.3	10.3	10.3	10.3	10.3	10.3	10.3	10.3	10.3	10.3
38000	53.4	51.9	51.9	51.7	51.1	50.1	50.9	49.8	49.3	48.9	10.3	10.3	10.3	10.3	10.3	10.3	10.3	10.3	10.3

Table S4: Biophysical parameters of multi-compartmental models used in this study.

Parameter name/Model ID	***1***	***2***	***3***	***4***	***5***	***6***	***7***
CaDynamics_E2.apic.decay	121.0674	138.7704	188.4791	114.9947	124.7849	94.58709	124.0039
CaDynamics_E2.apic.gamma	0.000813	0.000551	0.000522	0.000579	0.000519	0.000567	0.003474
CaDynamics_E2.axon.decay	119.401	333.7151	346.1883	196.5606	912.4342	973.4132	222.7396
CaDynamics_E2.axon.gamma	0.001766	0.002154	0.001582	0.000642	0.004945	0.000515	0.038612
CaDynamics_E2.soma.decay	301.457	530.8009	327.2006	155.3969	311.9493	213.7342	88.21946
CaDynamics_E2.soma.gamma	0.000517	0.000554	0.010279	0.007425	0.000502	0.03446	0.001622
Ca_HVA.apic.gCa_HVAbar	0.002759	0.004131	0.003894	0.001015	0.004717	0.003176	0.001476
Ca_HVA.axon.gCa_HVAbar	0.000185	0.000488	2.86E-05	0.000819	2.96E-05	2.85E-05	5.13E-05
Ca_HVA.soma.gCa_HVAbar	0.000683	0.000298	3.56E-05	0.00011	0.000564	2.22E-05	0.000502
Ca_LVAst.apic.gCa_LVAstbar	0.001881	0.063474	0.025106	0.002194	0.009637	0.001959	0.024519
Ca_LVAst.axon.gCa_LVAstbar	0.009016	0.006404	0.004188	0.003719	3.58E-05	0.00286	0.00011
Ca_LVAst.soma.gCa_LVAstbar	0.004735	0.002137	0.001545	0.000614	0.006072	5.14E-06	0.008861
Im.apic.gImbar	0.000113	4.52E-05	2.47E-05	2.7E-06	0.000381	4.03E-05	9.71E-07
K_Pst.axon.gK_Pstbar	0.004625	0.296952	0.290188	0.005763	0.104922	0.172764	0.077452
K_Pst.soma.gK_Pstbar	0.005713	0.078117	0.140957	0.003549	0.025351	0.152812	0.009897
K_Tst.axon.gK_Tstbar	0.035837	0.006818	0.096493	0.044702	0.03351	0.027037	0.040092
K_Tst.soma.gK_Tstbar	0.033603	0.003949	0.078375	0.099119	0.081466	0.075429	0.078967
NaTa_t.apic.gNaTa_tbar	0.016966	0.01891	0.017893	0.013856	0.016047	0.020064	0.017836
NaTa_t.axon.gNaTa_tbar	3.863664	3.553574	3.560277	3.899617	3.711704	0.109311	3.348559
NaTa_t.soma.gNaTa_tbar	3.208129	1.833618	3.362703	3.896084	1.325818	3.843608	3.992573
Nap_Et2.axon.gNap_Et2bar	0.009615	0.002898	0.001821	0.009965	0.002535	0.002898	0.009859
Nap_Et2.soma.gNap_Et2bar	0.002371	0.003883	0.009257	0.005123	0.000599	0.000328	0.001849
SK_E2.apic.gSK_E2bar	0.002111	0.002975	0.003351	0.001143	0.003455	0.003749	0.000585
SK_E2.axon.gSK_E2bar	0.009162	0.01546	0.077013	0.014094	0.08137	0.00117	0.020578
SK_E2.soma.gSK_E2bar	0.079714	0.029528	0.024539	0.066667	0.080914	0.059734	0.003856
SKv3_1.apic.gSKv3_1bar	0.002863	0.014618	0.01143	7.1E-05	0.007639	0.003849	8.58E-06
SKv3_1.apic.offset	0.703146	0.679136	0.594989	0.086458		0.845104	0.212368
SKv3_1.apic.slope	-2.73404	-2.04156	-2.79123	-2.98351		-0.87599	-1.09782
SKv3_1.axon.gSKv3_1bar	0.002863	0.014618	0.01143	7.1E-05	0.007639	0.003849	8.58E-06
SKv3_1.soma.gSKv3_1bar	0.002863	0.014618	0.01143	7.1E-05	0.007639	0.003849	8.58E-06
apic.g_pas	7E-05	5.77E-05	3.43E-05	4.54E-05	4.16E-05	3.37E-05	3.54E-05
axon.g_pas	2.42E-05	3.57E-05	2.09E-05	2.07E-05	4.99E-05	2.69E-05	3.49E-05
dend.g_pas	5.3E-05	8.21E-05	7.97E-05	4.15E-05	6.45E-05	6.26E-05	6.19E-05
soma.g_pas	2.31E-05	2.04E-05	4.96E-05	2.25E-05	2.54E-05	3.56E-05	2.01E-05
scale_apical	1	2.181484	2.766078	1.713925	1	2.910863	1.345393

Table S5. Electrophysiological responses of the multi-compartmental models used in this study determined from the stimulus protocols shown in Figure 7H.

	model 1	model 2	model 3	model 4	model 5	model 6	model 7
Somatic AP peak (mV)	36.79	31.89	33.81	30.28	36.70	39.49	32.65
Somatic AP half-width (ms)	1.08	0.96	1.02	1.10	1.16	1.03	1.00
bAP amplitue 180 microns below the BP	24.12	31.17	37.65	41.35	27.64	26.82	40.78
bAP amplitude at the BP	15.04	15.28	19.66	36.83	15.95	10.50	30.78
Somatic AP spike count	1	1	1	1	1	1	1
Mean AP height during 3 AP burst	28.91	22.24	27.25	25.59	26.44	31.99	24.60
Somatic AP ISI	10.29	10.85	10.71	11.46	10.30	12.03	11.87
Somatic AHP depth (mV)	-62.20	-62.79	-63.54	-64.54	-62.45	-65.45	-62.65
Ca2+ AP peak (mV)	6.42	3.14	6.48	5.74	-0.85	3.55	5.11
Ca2+ AP width (ms)	35.48	37.53	36.89	38.60	34.82	38.18	37.99
Somatic AP spike count (BAC stimulus)	3	3	3	3	3	3	3
First spike latency (ms), Step 1	29.96	44.92	32.81	27.14	28.23	33.90	31.53
First spike latency (ms), Step 2	18.23	21.98	19.94	17.85	17.19	20.24	19.69
First spike latency (ms), Step 3	6.61	7.03	7.27	7.15	5.68	6.75	7.32
Initial Burst ISI (ms), Step 1	11.91	67.17	42.92	23.50	15.05	48.00	15.59
Initial Burst ISI (ms), Step 2	8.75	16.49	19.06	14.49	9.26	22.17	10.62
Initial Burst ISI (ms), Step 3	5.78	6.03	6.44	8.43	5.61	6.06	6.04
AP peak (mV), Step 1	22.37	14.71	22.31	17.93	18.05	24.74	19.42
AP peak (mV), Step 2	22.23	14.30	22.20	18.05	17.64	24.80	18.56
AP peak (mV), Step 3	21.07	10.84	21.23	17.75	15.05	24.15	13.98
AP half-width (ms), Step 1	1.06	0.95	0.98	1.11	1.03	0.95	0.98
AP half-width (ms), Step 2	1.06	0.95	0.98	1.10	1.03	0.95	0.98
AP half-width (ms), Step 3	1.06	0.94	0.98	1.09	1.02	0.94	0.98
Fast AHP depth (mV), Step 1	-60.10	-59.82	-61.33	-61.89	-59.58	-62.05	-60.55
Fast AHP depth (mV), Step 2	-60.12	-59.49	-61.26	-62.16	-59.31	-62.01	-60.12
Fast AHP depth (mV), Step 3	-59.64	-56.78	-60.66	-62.36	-57.77	-61.62	-57.46
Slow AHP depth (mV), Step 1	-60.61	-60.77	-62.24	-63.43	-59.48	-62.54	-61.24
Slow AHP depth (mV), Step 2	-60.49	-60.28	-61.98	-63.43	-59.04	-62.51	-60.68
Slow AHP depth (mV), Step 3	-59.81	-57.08	-60.94	-63.28	-57.52	-61.91	-57.38
Slow AHP time, Step 1	0.33	0.16	0.20	0.30	0.29	0.20	0.32
Slow AHP time, Step 2	0.26	0.25	0.25	0.31	0.28	0.23	0.33
Slow AHP time, Step 3	0.22	0.27	0.22	0.22	0.20	0.26	0.15
AP frequency (Hz), Step 1	10.50	7.00	10.00	10.50	11.00	9.50	10.50
AP frequency (Hz), Step 2	15.00	13.50	14.00	13.50	15.50	13.50	14.00
AP frequency (Hz), Step 3	25.00	29.00	23.50	21.00	25.50	21.50	28.00
Adaptation index, Step 1	0.01	0.02	0.01	-0.01	0.02	0.00	0.00
Adaptation index, Step 2	0.02	0.01	0.01	0.00	0.02	0.01	0.01
Adaptation index, Step 3	0.01	0.01	0.01	0.01	0.01	0.01	0.01
ISI-CV, Step 1	0.12	0.09	0.08	0.07	0.12	0.05	0.06
ISI-CV, Step 2	0.16	0.12	0.09	0.06	0.13	0.09	0.12
ISI-CV, Step 3	0.15	0.13	0.10	0.12	0.14	0.13	0.13

Table S6: Single trial accuracy of reduced models versus multi-compartmental models. Tables are confusion matrices representing the number of trials in which at least one AP was elicited in the response window (25ms post whisker stimulus) in the biophysically detailed multi-compartmental models (bio) and reduced models, respectively. Data is provided for all whisker stimuli, i.e., the principal and 8 surround whiskers. The shaded value is the overall accuracy (i.e., the percentage of simulation trials in which the reduced and biophysically detailed model match).

model 1

	AP (bio)	no AP (bio)	%
AP (reduced)	61348	13092	82.41
no AP (reduced)	10660	643900	98.37
%	85.20	98.01	96.74

model 5

	AP (bio)	no AP (bio)	%
AP (reduced)	108704	23461	82.25
no AP (reduced)	23444	573391	96.07
%	82.26	96.07	93.57

model 2

	AP (bio)	no AP (bio)	%
AP (reduced)	91659	23253	79.76
no AP (reduced)	21801	592287	96.45
%	80.79	96.22	93.82

model 6

	AP (bio)	no AP (bio)	%
AP (reduced)	124309	38886	76.17
no AP (reduced)	57407	508398	89.85
%	68.41	92.89	86.79

model 3

	AP (bio)	no AP (bio)	%
AP (reduced)	163240	30828	84.11
no AP (reduced)	52914	482018	90.11
%	75.52	93.99	88.51

model 7

	AP (bio)	no AP (bio)	%
AP (reduced)	127239	23081	84.65
no AP (reduced)	37475	541205	93.52
%	77.25	95.91	91.69

model 4

	AP (bio)	no AP (bio)	%
AP (reduced)	127538	57207	69.03
no AP (reduced)	52427	491828	90.37
%	70.87	89.58	84.96

10 References

Agnati, L. F., Guidolin, D., Carone, C., Dam, M., Genedani, S., and Fuxe, K. (2008). Understanding neuronal molecular networks builds on neuronal cellular network architecture. *Brain Research Reviews* 58, 379–399. doi: 10.1016/j.brainresrev.2007.11.002

Amsalem, O., Eyal, G., Rogozinski, N., Gevaert, M., Kumbhar, P., Schürmann, F., et al. (2020). An efficient analytical reduction of detailed nonlinear neuron models. *Nat Commun* 11, 288. doi: 10.1038/s41467-019-13932-6

Averbeck, B. B., Latham, P. E., and Pouget, A. (2006). Neural correlations, population coding and computation. *Nat Rev Neurosci* 7, 358–366. doi: 10.1038/nrn1888

Bast, A., Guest, J. M., Fruengel, R., Narayanan, R. T., De Kock, C. P. J., and Oberlaender, M. (2021). Thalamus drives active dendritic computations in cortex. doi: 10.1101/2021.10.21.465325

Bast, A., and Oberlaender, M. (2021). Ion channel distributions in cortical neurons are optimized for energy-efficient active dendritic computations. doi: 10.1101/2021.12.11.472235

Beniaguev, D., Segev, I., and London, M. (2021). Single cortical neurons as deep artificial neural networks. *Neuron* 109, 2727–27393. doi: 10.1016/j.neuron.2021.07.002

Bi, G., and Poo, M. (2001). Synaptic Modification by Correlated Activity: Hebb's Postulate Revisited. *Annu. Rev. Neurosci.* 24, 139–166. doi: 10.1146/annurev.neuro.24.1.139

Caron, S. J. C., Ruta, V., Abbott, L. F., and Axel, R. (2013). Random convergence of olfactory inputs in the Drosophila mushroom body. *Nature* 497, 113–117. doi: 10.1038/nature12063

Chen, X., Leischner, U., Rochefort, N. L., Nelken, I., and Konnerth, A. (2011). Functional mapping of single spines in cortical neurons in vivo. *Nature* 475, 501–505. doi: 10.1038/nature10193

Cornford, J., Kalajdzievski, D., Leite, M., Lamarquette, A., Kullmann, D. M., and Richards, B. (2020). Learning to live with Dale's principle: ANNs with separate excitatory and inhibitory units. doi: 10.1101/2020.11.02.364968

Dapello, J., Feather, J., Le, H., Marques, T., Cox, D., McDermott, J., et al. (2021). Neural Population Geometry Reveals the Role of Stochasticity in Robust Perception., in *Advances in Neural Information Processing Systems*, eds. M. Ranzato, A. Beygelzimer, Y. Dauphin, P. S. Liang, and J. W. Vaughan (Curran Associates, Inc.), 15595–15607. Available at:

https://proceedings.neurips.cc/paper_files/paper/2021/file/8383f931b0cefcc631f070480ef340e1-Paper.pdf

Dask Development Team (2016). Dask: Library for dynamic task scheduling. Available at: https://dask.org

de Kock, C. P. J., Bruno, R. M., Spors, H., and Sakmann, B. (2007). Layer- and cell-type-specific suprathreshold stimulus representation in rat primary somatosensory cortex. *The Journal of physiology* 581, 139–154. doi: 10.1113/jphysiol.2006.124321

De Kock, C. P. J., Pie, J., Pieneman, A. W., Mease, R. A., Bast, A., Guest, J. M., et al. (2021). High-frequency burst spiking in layer 5 thick-tufted pyramids of rat primary somatosensory cortex encodes exploratory touch. *Commun Biol* 4, 709. doi: 10.1038/s42003-021-02241-8

De Kock, C. P. J., and Sakmann, B. (2009). Spiking in primary somatosensory cortex during natural whisking in awake head-restrained rats is cell-type specific. *Proc. Natl. Acad. Sci. U.S.A.* 106, 16446–16450. doi: 10.1073/pnas.0904143106

Dehaene, S., and Changeux, J.-P. (2011). Experimental and Theoretical Approaches to Conscious Processing. *Neuron* 70, 200–227. doi: 10.1016/j.neuron.2011.03.018

Dettmers, T., and Zettlemoyer, L. (2019). Sparse Networks from Scratch: Faster Training without Losing Performance. doi: 10.48550/ARXIV.1907.04840

Eccles, J. C. (1976). From electrical to chemical transmission in the central nervous system: The closing address of the Sir Henry Dale Centennial Symposium Cambridge, 19 September 1975. *Notes Rec. R. Soc. Lond.* 30, 219–230. doi: 10.1098/rsnr.1976.0015

Egger, R., Dercksen, V. J., Udvary, D., Hege, H.-C., and Oberlaender, M. (2014). Generation of dense statistical connectomes from sparse morphological data. *Front Neuroanat* 8, 129. doi: 10.3389/fnana.2014.00129

Egger, R., Narayanan, R. T., Guest, J. M., Bast, A., Udvary, D., Messore, L. F., et al. (2020). Cortical Output Is Gated by Horizontally Projecting Neurons in the Deep Layers. *Neuron* 105, 122–1378. doi: 10.1016/j.neuron.2019.10.011

Evci, U., Pedregosa, F., Gomez, A., and Elsen, E. (2019). The Difficulty of Training Sparse Neural Networks. Available at: http://arxiv.org/pdf/1906.10732v3

Feldmeyer, D., Brecht, M., Helmchen, F., Petersen, C. C. H., Poulet, J. F. A., Staiger, J. F., et al. (2013). Barrel cortex function. *Progress in neurobiology* 103, 3–27. doi: 10.1016/j.pneurobio.2012.11.002

Frankle, J., and Carbin, M. (2019). The Lottery Ticket Hypothesis: Finding Sparse, Trainable Neural Networks. Available at: http://arxiv.org/pdf/1803.03635v5

Gomez-Marin, A. (2021). Promisomics and the Short-Circuiting of Mind. *eNeuro* 8. doi: 10.1523/ENEURO.0521-20.2021

Groh, A., De Kock, C. P. J., Wimmer, V. C., Sakmann, B., and Kuner, T. (2008). Driver or Coincidence Detector: Modal Switch of a Corticothalamic Giant Synapse Controlled by Spontaneous Activity and Short-Term Depression. *J. Neurosci.* 28, 9652–9663. doi: 10.1523/JNEUROSCI.1554-08.2008

Han, S., Pool, J., Tran, J., and Dally, W. (2015). Learning both Weights and Connections for Efficient Neural Network., in *Advances in Neural Information Processing Systems*, eds. C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett (Curran Associates, Inc.). Available at: https://proceedings.neurips.cc/paper_files/paper/2015/file/ae0eb3eed39d2bcef4622b2499a05fe6-Paper.pdf

Harris, K. D., and Shepherd, G. M. G. (2015). The neocortical circuit: themes and variations. *Nat Neurosci* 18, 170–181. doi: 10.1038/nn.3917

Hay, E., Hill, S., Schürmann, F., Markram, H., and Segev, I. (2011). Models of neocortical layer 5b pyramidal cells capturing a wide range of dendritic and perisomatic active properties. *PLoS computational biology* 7, e1002107. doi: 10.1371/journal.pcbi.1002107

Hay, E., Schürmann, F., Markram, H., and Segev, I. (2013). Preserving axosomatic spiking features despite diverse dendritic morphology. *Journal of Neurophysiology* 109, 2972–2981. doi: 10.1152/jn.00048.2013

Hay, E., and Segev, I. (2015). Dendritic Excitability and Gain Control in Recurrent Cortical Microcircuits. *Cerebral cortex (New York, N.Y.: 1991)* 25, 3561–3571. doi: 10.1093/cercor/bhu200

Helmchen, F., Svoboda, K., Denk, W., and Tank, D. W. (1999). In vivo dendritic calcium dynamics in deep-layer cortical pyramidal neurons. *Nat Neurosci* 2, 989–996. doi: 10.1038/14788

Herculano-Houzel, S., Mota, B., Wong, P., and Kaas, J. H. (2010). Connectivity-driven white matter scaling and folding in primate cerebral cortex. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19008–19013. doi: 10.1073/pnas.1012590107

Hines, M. L., and Carnevale, N. T. (1997). The NEURON simulation environment. *Neural computation* 9, 1179–1209. doi: 10.1162/neco.1997.9.6.1179

Ilyas, A., Santurkar, S., Tsipras, D., Engstrom, L., Tran, B., and Madry, A. (2019). Adversarial Examples Are Not Bugs, They Are Features. Available at: http://arxiv.org/abs/1905.02175 (Accessed March 8, 2022).

Jia, H., Varga, Z., Sakmann, B., and Konnerth, A. (2014). Linear integration of spine Ca $^{2+}$ signals in layer 4 cortical neurons in vivo. *Proc. Natl. Acad. Sci. U.S.A.* 111, 9277–9282. doi: 10.1073/pnas.1408525111

Jonas, E., and Kording, K. P. (2017). Could a Neuroscientist Understand a Microprocessor? *PLoS Comput Biol* 13, 1005268. doi: 10.1371/journal.pcbi.1005268

Kemp, B., Zwinderman, A. H., Tuk, B., Kamphuisen, H. A. C., and Oberye, J. J. L. (2000). Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. *IEEE Transactions on Biomedical Engineering* 47, 1185–1194. doi: 10.1109/10.867928

Kovács, I. A., Barabási, D. L., and Barabási, A.-L. (2020). Uncovering the genetic blueprint of the C. elegans nervous system. *Proc Natl Acad Sci U S A* 117, 33570–33577. doi: 10.1073/pnas.2009093117

Krizhevsky, A. (2012). Learning Multiple Layers of Features from Tiny Images. *University of Toronto*.

Landau, I. D., Egger, R., Dercksen, V. J., Oberlaender, M., and Sompolinsky, H. (2016). The Impact of Structural Heterogeneity on Excitation-Inhibition Balance in Cortical Networks. *Neuron* 92, 1106–1121. doi: 10.1016/j.neuron.2016.10.027

Larkum, M. E., and Zhu, J. J. (2002). Signaling of Layer 1 and Whisker-Evoked Ca 2+ and Na + Action Potentials in Distal and Terminal Dendrites of Rat Neocortical Pyramidal Neurons In Vitro and In Vivo. *J. Neurosci.* 22, 6991–7005. doi: 10.1523/JNEUROSCI.22-16-06991.2002

Larkum, M. E., Zhu, J. J., and Sakmann, B. (1999). A new cellular mechanism for coupling inputs arriving at different cortical layers. *Nature* 398, 338–41. doi: 10.1038/18686

Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proc. IEEE* 86, 2278–2324. doi: 10.1109/5.726791

Lefort, S., Tomm, C., Floyd Sarria, J.-C., and Petersen, C. C. H. (2009). The excitatory neuronal network of the C2 barrel column in mouse primary somatosensory cortex. *Neuron* 61, 301–316. doi: 10.1016/j.neuron.2008.12.020

Liu, S., Ni'mah, I., Menkovski, V., Mocanu, D. C., and Pechenizkiy, M. (2021). Efficient and effective training of sparse recurrent neural networks. *Neural Comput & Applic* 33, 9625–9636. doi: 10.1007/s00521-021-05727-y

Maass, W. (2011). "Liquid State Machines: Motivation, Theory, and Applications," in *Computability in Context*, (IMPERIAL COLLEGE PRESS), 275–296. doi: 10.1142/9781848162778_0008

Marder, E., and Taylor, A. L. (2011). Multiple models to capture the variability in biological neurons and networks. *Nature neuroscience* 14, 133–138. doi: 10.1038/nn.2735

Mease, R. A., Kuner, T., Fairhall, A. L., and Groh, A. (2017). Multiplexed Spike Coding and Adaptation in the Thalamus. *Cell Reports* 19, 1130–1140. doi: 10.1016/j.celrep.2017.04.050

Mease, R. A., Metz, M., and Groh, A. (2016). Cortical Sensory Responses Are Enhanced by the Higher-Order Thalamus. *Cell Reports* 14, 208–215. doi: 10.1016/j.celrep.2015.12.026

Meyer, H. S., Egger, R., Guest, J. M., Foerster, R., Reissl, S., and Oberlaender, M. (2013). Cellular organization of cortical barrel columns is whisker-specific. *Proc. Natl. Acad. Sci. U.S.A.* 110, 19113–19118. doi: 10.1073/pnas.1312691110

Meyer, H. S., Schwarz, D., Wimmer, V. C., Schmitt, A. C., Kerr, J. N. D., Sakmann, B., et al. (2011). Inhibitory interneurons in a cortical column form hot zones of inhibition in layers 2 and 5A. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16807–16812. doi: 10.1073/pnas.1113648108

Mohajerani, M. H., McVea, D. A., Fingas, M., and Murphy, T. H. (2010). Mirrored Bilateral Slow-Wave Cortical Activity within Local Circuits Revealed by Fast Bihemispheric Voltage-Sensitive Dye Imaging in Anesthetized and Awake Mice. *J. Neurosci.* 30, 3745–3751. doi: 10.1523/JNEUROSCI.6437-09.2010

Murayama, M., and Larkum, M. E. (2009). Enhanced dendritic activity in awake rats. *Proc. Natl. Acad. Sci. U.S.A.* 106, 20482–20486. doi: 10.1073/pnas.0910379106

Murayama, M., Pérez-Garci, E., Lüscher, H.-R., and Larkum, M. E. (2007). Fiberoptic System for Recording Dendritic Calcium Signals in Layer 5 Neocortical Pyramidal Cells in Freely Moving Rats. *Journal of Neurophysiology* 98, 1791–1805. doi: 10.1152/jn.00082.2007

Narang, S., Elsen, E., Diamos, G., and Sengupta, S. (2017). Exploring Sparsity in Recurrent Neural Networks. doi: 10.48550/ARXIV.1704.05119

Nguyen, J. P., Shipley, F. B., Linder, A. N., Plummer, G. S., Liu, M., Setru, S. U., et al. (2016). Whole-brain calcium imaging with cellular resolution in freely behaving *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U.S.A.* 113. doi: 10.1073/pnas.1507110112

Nguyen, T., Raghu, M., and Kornblith, S. (2021). Do Wide and Deep Networks Learn the Same Things? Uncovering How Neural Network Representations Vary with Width and Depth. Available at: http://arxiv.org/abs/2010.15327 (Accessed March 8, 2022).

Palmer, L. M., Schulz, J. M., Murphy, S. C., Ledergerber, D., Murayama, M., and Larkum, M. E. (2012). The Cellular Basis of $GABA_B$ -Mediated Interhemispheric Inhibition. *Science* 335, 989–993. doi: 10.1126/science.1217276

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine Learning in Python. Available at: http://arxiv.org/pdf/1201.0490v4 (Accessed January 2, 2012).

Pernice, V., Staude, B., Cardanobile, S., and Rotter, S. (2011). How Structure Determines Correlations in Neuronal Networks. *PLoS Comput Biol* 7, e1002059. doi: 10.1371/journal.pcbi.1002059

Petersen, C. C. H., Hahn, T. T. G., Mehta, M., Grinvald, A., and Sakmann, B. (2003). Interaction of sensory responses with spontaneous depolarization in layer 2/3 barrel cortex. *Proc. Natl. Acad. Sci. U.S.A.* 100, 13638–13643. doi: 10.1073/pnas.2235811100

Rochefort, N. L., and Konnerth, A. (2012). Dendritic spines: from structure to *in vivo* function. *EMBO Reports* 13, 699–708. doi: 10.1038/embor.2012.102

Rojas-Piloni, G., Guest, J. M., Egger, R., Johnson, A. S., Sakmann, B., and Oberlaender, M. (2017). Relationships between structure, in vivo function and long-range axonal target of cortical pyramidal tract neurons. *Nat Commun* 8, 870. doi: 10.1038/s41467-017-00971-0

Rusakov, D. A., Savtchenko, L. P., and Latham, P. E. (2020). Noisy synaptic conductance: Bug or a feature? *Trends in Neurosciences* 43, 363–372. doi: https://doi.org/10.1016/j.tins.2020.03.009

Scholl, B., Thomas, C. I., Ryan, M. A., Kamasawa, N., and Fitzpatrick, D. (2021). Cortical response selectivity derives from strength in numbers of synapses. *Nature* 590, 111–114. doi: 10.1038/s41586-020-03044-3

Sejnowski, T. J., Churchland, P. S., and Movshon, J. A. (2014). Putting big data to good use in neuroscience. *Nat Neurosci* 17, 1440–1441. doi: 10.1038/nn.3839

Shadlen, M. N., and Newsome, W. T. (1998). The Variable Discharge of Cortical Neurons: Implications for Connectivity, Computation, and Information Coding. *J. Neurosci.* 18, 3870–3896. doi: 10.1523/JNEUROSCI.18-10-03870.1998

Singer, W. (2021). Recurrent dynamics in the cerebral cortex: Integration of sensory evidence with stored knowledge. *Proc. Natl. Acad. Sci. U.S.A.* 118, e2101043118. doi: 10.1073/pnas.2101043118

Song, S., Sjöström, P. J., Reigl, M., Nelson, S., and Chklovskii, D. B. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol* 3, 68. doi: 10.1371/journal.pbio.0030068

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research* 15, 1929–1958.

Suzuki, M., and Larkum, M. E. (2020). General Anesthesia Decouples Cortical Pyramidal Neurons. *Cell* 180, 666-676.e13. doi: 10.1016/j.cell.2020.01.024

Takahashi, N., Ebner, C., Sigl-Glöckner, J., Moberg, S., Nierwetberg, S., and Larkum, M. E. (2020). Active dendritic currents gate descending cortical outputs in perception. *Nat Neurosci* 23, 1277–1285. doi: 10.1038/s41593-020-0677-8

Takahashi, N., Oertner, T. G., Hegemann, P., and Larkum, M. E. (2016). Active cortical dendrites modulate perception. *Science (New York, N.Y.)* 354, 1587–1590. doi: 10.1126/science.aah6066

Udvary, D., Harth, P., Macke, J. H., Hege, H.-C., de Kock, C. P. J., Sakmann, B., et al. (2022). The impact of neuron morphology on cortical network architecture. *Cell Reports* 39, 110677. doi: 10.1016/j.celrep.2022.110677

Ujfalussy, B. B., Makara, J. K., Lengyel, M., and Branco, T. (2018). Global and Multiplexed Dendritic Computations under In Vivo-like Conditions. *Neuron* 100, 579–5925. doi: 10.1016/j.neuron.2018.08.032

van Geit, W., Gevaert, M., Chindemi, G., Rössert, C., Courcol, J.-D., Muller, E. B., et al. (2016). BluePyOpt: Leveraging Open Source Software and Cloud Infrastructure to Optimise Model Parameters in Neuroscience. *Frontiers in neuroinformatics* 10, 17. doi: 10.3389/fninf.2016.00017

Varga, Z., Jia, H., Sakmann, B., and Konnerth, A. (2011). Dendritic coding of multiple sensory inputs in single cortical neurons in vivo. *Proc. Natl. Acad. Sci. U.S.A.* 108, 15420–15425. doi: 10.1073/pnas.1112355108

Ven, L., and Lederer, J. (2021). Regularization and Reparameterization Avoid Vanishing Gradients in Sigmoid-Type Networks. doi: 10.48550/arXiv.2106.02260

Viaene, A. N., Petrof, I., and Sherman, S. M. (2011). Synaptic Properties of Thalamic Input to the Subgranular Layers of Primary Somatosensory and Auditory Cortices in the Mouse. *Journal of Neuroscience* 31, 12738–12747. doi: 10.1523/JNEUROSCI.1565-11.2011

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature methods* 17, 261–272. doi: 10.1038/s41592-019-0686-2

Watson, J. F., Vargas-Barroso, V., Morse-Mora, R. J., Navas-Olive, A., Tavakoli, M. R., Danzl, J. G., et al. (2025). Human hippocampal CA3 uses specific functional connectivity rules for efficient associative memory. *Cell* 188, 501-514.e18. doi: 10.1016/j.cell.2024.11.022

Weber, A. I., and Pillow, J. W. (2016). Capturing the dynamical repertoire of single neurons with generalized linear models. Available at: http://arxiv.org/pdf/1602.07389v3

White, J. G., Southgate, E., Thomson, J. N., and Brenner, S. (1986). The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Phil. Trans. R. Soc. Lond. B* 314, 1–340. doi: 10.1098/rstb.1986.0056

Wybo, W. A., Jordan, J., Ellenberger, B., Marti Mengual, U., Nevian, T., and Senn, W. (2021). Data-driven reduction of dendritic morphologies with preserved dendro-somatic responses. *eLife* 10, e60936. doi: 10.7554/eLife.60936

Wybo, W. A. M., Torben-Nielsen, B., Nevian, T., and Gewaltig, M.-O. (2019). Electrical Compartmentalization in Neurons. *Cell Reports* 26, 1759-1773.e7. doi: 10.1016/j.celrep.2019.01.074

Xu, N., Harnett, M. T., Williams, S. R., Huber, D., O'Connor, D. H., Svoboda, K., et al. (2012). Nonlinear dendritic integration of sensory and motor input during an active sensing task. *Nature* 492, 247–251. doi: 10.1038/nature11601

Xue, M., Atallah, B. V., and Scanziani, M. (2014). Equalizing excitation–inhibition ratios across visual cortical neurons. *Nature* 511, 596–600. doi: 10.1038/nature13321

Zador, A. M. (2019). A critique of pure learning and what artificial neural networks can learn from animal brains. *Nat Commun* 10, 3770. doi: 10.1038/s41467-019-11786-6