Cortical state contributions to response variability in the early visual cortex: A system identification approach

running head: CNN model of A17/18 neurons incorporating cortical states

Jinani Sooriyaarachchi[1], Chang'an Zhan[2], & Curtis L. Baker Jr.[3]*

[1]Department of Physiology, McGill University, Montreal, Quebec, Canada

[2]School of Biomedical Engineering, Southern Medical University, China

[3]Department of Ophthalmology and Visual Sciences, McGill University, Montreal, Quebec, Canada

* Corresponding Author

Email: curtis.baker@mcgill.ca

Acknowledgements: This work was supported by the Canadian Institutes of Health Research Grant MOP-119498 to C.L.B, and a Quebec FRQS doctoral fellowship to J.S.. We thank Guangxing Li for contributions to software and technical support in experiments; and Amol Gharat, Phillipe Nguyen and Guangxing Li for important contributions to data collection.

Abstract

Neurons in the early visual cortex respond selectively to multiple features of visual stimuli, but they respond differently to repeated presentation of the same visual stimulus. Such trial-to-trial response variabilities are often treated as noise and addressed by simple trial-averaging to obtain the stimulus-driven response, though this approach is insufficient to fully remove the response variability. More importantly, there is evidence that response variability may primarily be caused by non-sensory factors, particularly by variations in cortical state.

Here we recorded and analyzed neuronal spiking activity in response to natural images from areas 17 and 18 of cats, along with local population neuronal signals; local field potential (LFP) and multi-unit activity (MUA). We used a variability ratio (VR) measure to quantify the variability of neural responses across trials and two cortical state indicative measures; global fluctuation index (GFI) calculated using MUA and synchrony index (SI) calculated using LFP signals. We propose a compact convolutional neural network model with two parallel pathways, to capture the stimulus-driven activity and the cortical state-driven response variabilities. The stimulus-driven pathway contains a spatiotemporal filter, a parametric rectifier and a Gaussian map and the cortical state-driven pathway contains temporal filters for MUA and LFP. The model parameters are fit to best predict each neuron's spiking activity.

Single neurons showed highly varying degrees of trial-to-trial response variability, even when recorded simultaneously. The fitted model performed with a significantly higher accuracy in predicting neural responses compared to a basic model with a

stimulus-driven pathway alone. The neurons with higher response variability benefited more from the cortical state-driven pathway compared to less variable neurons. These results provide insights to understanding the possibility of trial-to-trial response variabilities emerging as an effect of cortical state dynamics.

Author Summary

Neuronal responses in the early visual cortex to repeated presentation of an identical stimulus can be highly variable across trials, so a neuron's response is not solely predictable from the stimulus alone. In many cases, the variable portion of these neuronal responses can be as large as the stimulus-driven response. The cortical state of animals can vary continuously during a data recording session giving rise to the observed response variabilities, and these dynamics are linked to simultaneously recorded cortical signals including local field potentials and multi-unit activity. Here we demonstrate that a model combining these cortical signals along with a visual stimulus processing pathway can predict neural responses significantly better than a model containing a stimulus-driven pathway alone. This work provides insights to understanding the role of visual cortical neurons not only in processing visual stimuli, but also processing visual inputs along with non-sensory cognitive factors like cortical states fluctuations.

Introduction

A visual cortex neuron's response depends on the characteristics of a visual stimulus in relation to its receptive field (RF). However, the neuronal responses elicited by repeated presentations of the same stimulus can be highly variable (Carandini, 2004; Tolhurst et al., 1983; Vogels et al., 1989). This trial-to-trial response variability is conventionally treated as noise and mitigated by simple averaging across trials. However, simultaneous recordings from multiple single neurons have demonstrated that such trial-wise response variability are correlated across neural populations (Averbeck et al., 2006; Zohary et al., 1994), suggesting that simple averaging places critical constraints on understanding biologically plausible visual information processing. In fact, cortical neurons can produce highly reliable spike trains for intracellularly injected current (Mainen & Sejnowski, 1995). Thus, the observed response variability may not be solely caused by intrinsic noise of the neuron itself, but rather the network activity (e.g., cortical state, Harris & Thiele 2011) within which the neuron is embedded.

Recent attempts at developing models to explain this trial-to-trial variability in the visual cortex used population neural signals, e.g. local field potentials (LFP) and multi-unit activities (MUA), to infer cortical variability (Cui et al., 2016; Kelly et al., 2010). However, these studies did not address the extent to which such response variability were heterogeneous across neurons, and effects of cortical states in simultaneously recorded neurons. Understanding the nature of trial-to-trial response variability at the single neuron level could guide development of biologically plausible models to explain visual cortical computations and circuits. Furthermore, separating neurons' responses to

visual stimuli from brain state effects could facilitate improved characterization of neuronal sensory RFs.

The trial-to-trial variability of neural responses has been thought to be caused by the variability of synaptic inputs, reflecting the coordinated activities in the cortical network (Arieli et al., 1996; Carandini, 2004). These ongoing activity patterns were historically studied in terms of EEG, e.g. Berger's observation of an 8-12Hz rhythm while his son sat passively with his eyes closed and a 14-30Hz rhythm when he opened his eyes (Berger, 1929). Such fluctuations of brain oscillations are often interpreted as reflecting “cortical states”, linked to multiple behavioral dynamics (Harris & Thiele, 2011). Although cortical states have been most studied as categorically distinct in relation to sleep stages (Brown et al., 2012) and depth of anesthesia (Liu et al., 2021; Ward-Flanagan et al., 2022), recent studies have indicated a more complex picture in which cortical state varies in a continuum during wakefulness (Reimer et al., 2014) due to factors including attention (Saalmann et al., 2007), working memory (Mendoza-Halliday et al., 2014), or general arousal (Constantinople & Bruno, 2011; Vinck et al., 2015). Recent studies on rodents demonstrated a relationship between brain-wide neural response variability and behavioral patterns including locomotor activity, pupil size and whisker movements (Musall et al., 2019; Stringer et al., 2019). Thus it seems likely that trial-to-trial variability may reflect the effects of cortical states, and their influences in sensory cortical function, rather than being just “noise”.

Disentangling the effects of cortical states on sensory information processing is challenging as the effects are confounded with the responses to the sensory stimuli,

and so it is critical to accurately assess cortical state fluctuations over time. Moreover, the brain state effects may not be uniform across a population of neurons. Some studies have demonstrated varying coupling of neuronal firing to the population activity, with the strongly coupled neurons showing strong activation during population-wide events including top-down modulation (e.g. motor intention in saccade preparation) (Okun et al., 2015). Therefore, it's important to study the cortical state effects at single neuronal scale to fully understand the differential effects of cortical states.

Broadband multielectrode recordings have been observed to co-vary with cortical states. These signals contain lower frequency components (0.5-100 Hz), commonly termed local field potentials (LFPs), which reflect synaptic inputs to a given cortical region (Buzsáki et al., 2012; Buzsáki & Draguhn, 2004), and more global EEG, which are thought to be indicative of brain states (Roohi-Azizi et al., 2017). The high frequency (300-5000 Hz) components (usually referred as to multi-unit activity, MUA) in nearby channels have also been found to predict single neurons' spontaneous spiking activity (Okun et al., 2015; Schölvinck et al., 2015), suggesting that MUA may also be indicative of brain states. The LFP and MUA signals along with the visual stimulus information may be quantitatively analyzed jointly to uncover possible interactions between them, rather than studying them in isolation.

Most early modelling efforts have assumed that cortical state acts either additively or multiplicatively on sensory neuronal responses. Multiplicative effects could conceivably result from top-down feedback related to attentional modulation (Armstrong & Moore, 2007), and some modeling works have assumed multiplicative (gain)

modulation (Ecker et al., 2014). However, several studies modelled the cortical state modulatory effects on sensory processing as additive (Cui et al., 2016; Kelly et al., 2010; Murray, 2008; Thiele et al., 2009). In addition, one research revealed that a combination of both additive and multiplicative modulatory effects could better explain the trial-to-trial variability compared to either effect alone (Lin et al., 2015). These attempts were mainly aimed at improving accuracy in predicting neural responses but have seldom explored the effects of cortical states in estimating the RFs of the studied neurons. Furthermore, most of these studies used pre-defined LFP frequency bands like those conventionally used for EEG. However, the implied categorical nature of these frequency bands, and the assumed boundaries between them, might result in missing important relationships and information provided in the raw signals (Cohen, 2020).

Here, we extended a system identification model of visual responses which estimates the receptive fields of early visual cortex neurons based on a simple convolution neural network (Nguyen et al., 2024). We incorporated the fluctuations of cortical states inferred from LFP and MUA signals recorded simultaneously from nearby electrode sites to explain the response variability of visual cortex neurons to repeated stimuli. The individual brain state effects on each neuron were captured using separate temporal filters applied to raw LFP and MUA signals in two different parallel pathways. We find that visual cortex neurons differ greatly from one another in their response variability, even when they are recorded simultaneously. Our complete system identification model including parallel pathways for visual stimuli, LFPs and MUA, and

possible interactions between them, can substantially improve the predictive accuracy of neuronal responses by better accounting for the response variability. Moreover, the improvement is greater for neurons with higher trial-to-trial response variability. Furthermore, we find that the quantitative improvement in predictive accuracy is also accompanied by improvements in the estimated RFs. The estimated LFP temporal filters extracted different important LFP frequencies uniquely for each neuron, showing differential effects of cortical states on different neurons including the ones that were recorded simultaneously. These results thus demonstrate the significance of accounting for cortical state effects in characterizing the sensory responses at the individual neuronal level.

Materials and Methods

Animal Preparation

Detailed descriptions of the procedures for animal surgery, recording, and maintenance have been published previously (St-Amand & Baker, 2023) and will only be summarized here. Adult cats of either sex were anesthetized with isofluorane/oxygen (3–5%), followed by intravenous (i.v.) cannulation. A bolus i.v. injection (5 $mg \cdot kg^{-1}$) followed by a continuous infusion (5.3 $mg \cdot kg^{-1} \cdot h^{-1}$) of propofol was provided for surgical anesthesia. The animal was secured in a stereotaxic apparatus, and a craniotomy and small duratomy were performed over the cortical region of interest (Areas 17/18). The cortical surface was protected by 2% agarose and petroleum jelly. Body temperature was maintained using a thermister-controlled heating pad. The corneas were protected

during the initial surgery using topical ophthalmic carboxymethylcellulose (1%), and subsequently, neutral contact lenses were inserted for long-term corneal protection. The eyes were brought into focus at the distance at which visual stimuli would be presented (57cm) using spectacle lenses, with artificial pupils (2.5mm) to further improve optical quality. Nictitating membranes were retracted, and pupils were dilated using topical phenylephrine hydrochloride (2.5%) and a mydiatic (atropine sulfate, 1%, or cyclopentolate, 1.0 %), which were re-applied daily.

The animal was connected to a ventilator supplying oxygen/nitrous oxide (70:30) after the completion of all surgical procedures. Gallamine triethiodide (i.v. bolus, to effect, followed by infusion, 10mg·kg^{-1}h^{-1}) was used to induce and maintain paralysis. The propofol infusion rate was reduced (5.3 mg·kg^{-1}h^{-1}) and additional anaesthesia was provided with i.v. remifentanil (bolus injection, 1.25 µg·kg^{-1}, followed by infusion, 3.7 µg·kg^{-1}h^{-1}). Vital signals (temperature, heart rate, expired CO2 and EEG activity) were monitored and kept at normal levels during recording experiments conducted over a period of 3-4 days.

All animal procedures were approved by the McGill University Animal Care Committee and are in accordance with the guidelines of the Canadian Council on Animal Care.

Electrophysiology and Visual Stimuli

Extracellular recordings were obtained from 32-channel multielectrode probes (NeuroNexus, linear array: A1x32-6mm-100-177, linear edge: A1x32-Edge-5mm-100-177, or polytrode: A1x32-Poly2-5mm-50-177), using a Plexon Recorder (f_s=31.25kHz)

or OpenEphys (f_s=30kHz) data acquisition system. Spiking responses (300Hz-5kHz) from a channel of interest were monitored “on-line”. All 32 channels of broadband neural data, as well as CRT photodiode signals (see below), were streamed to a hard disk for later off-line spike sorting using Kilosort software (Pachitariu et al., 2016), MUA and LFP signal extraction, and temporal registration between stimuli and neural signals. The sorted neural spikes, MUA and LFP signals were then analyzed in detail as described below.

Probes were inserted in approximately vertical, near-columnar penetrations, so that the recorded neurons generally had similar receptive field location, preferred spatial frequency, and preferred orientation (DeAngelis et al., 1999; Hubel & Wiesel, 1962; Tootell et al., 1981). An initial estimate of the neurons’ receptive field location and dominant eye were identified using an interactively controlled bar stimulus. The non-dominant eye was occluded, and the CRT monitor centered approximately on the receptive field.

Visual stimuli were generated on a CRT monitor (NEC FP1350, 20”, 640 x 480 pixels, 150 Hz, 36 cd/m^2) using custom software written in MATLAB (MathWorks, USA) with Psychophysics Toolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) running on a Macintosh computer (MacPro, 2.66 GHz Quad Core Intel Xeon, 6 GB, NVIDIA GeForce GT 120). The monitor’s gamma nonlinearity was measured with a photometer (United Detector Technology) and corrected using inverse lookup tables. Stimulus onset/offset timing and electrophysiological recordings were temporally registered using

a photodiode (TAOS, TSL12S) in a corner of the monitor screen. Photodiode signals were also used to verify the absence of dropped frames.

Detailed descriptions of the production of visual stimuli have been published (Talebi & Baker, 2012) and will only be summarized here. Natural image stimuli were constructed from monochrome 480x480 pixel subsets of photographs from the McGill Calibrated Color Image Database (Olmos & Kingdom, 2004). Images having low pixel standard deviations were discarded, since these images were nearly blank. We subtracted the mean luminance in remaining images and clipped to 8 bits. These natural images were the same as used previously (Nguyen et al., 2024; Talebi & Baker, 2012), except that in some experiments the image stimuli had higher RMS contrast, with the RMS contrast normalized or not. The observed final results were independent of these variations in the stimuli (S1 Fig) and will be considered jointly. Ensembles of 375 such images were presented in 5 sec trials. The CRT refresh rate was 150Hz, with each stimulus frame presented twice, for an effective rate of 75Hz. Stimulus image ensembles were divided into 3 sets, designed to be used for training, regularization, and testing (see below). Data collection was carried out in 5 trial blocks (henceforth referred to as "sweeps") each consisting of 20 training ensembles, and 5 ensembles each for regularization and test sets. Image ensemble sets were presented in pseudorandom order over a period of ca. 60 minutes.

Neural Data Processing

Neural data was processed in MATLAB (MathWorks, USA). The recorded broadband neural signals (sampling rate f_s=31.25kHz with Plexon or f_s=30kHz with OpenEphys)

were filtered using Matlab's *designfilt* and *filtfilt* functions. Neural signals were first low-pass filtered (<100Hz: passband frequency: 100, stopband frequency: 120, design method: '*kaiserwin*', impulse response: '*fir*', passband ripple: 0.1, stop band attenuation: 60), and also notch filtered to remove mains noise (60 Hz: half power frequencies: 59 and 61, design method: '*butter*', impulse response: '*iir*', filter order: 2) to obtain local field potentials (LFPs), and high-pass filtered (>600Hz: passband frequency: 600, stopband frequency: 550, design method: '*kaiserwin*', impulse response: '*fir*', passband ripple: 0.1, stop band attenuation: 60) to obtain multiunit spiking activity.

To identify the recording channels in grey matter, current source density (CSD) analysis based on the LFPs was employed. LFPs were extracted from the responses to drifting sinewave gratings at varying spatial or orientation frequencies collected for conventional tuning curve measurements. These LFPs were analysed across the channels (see also S2 Fig) using the *CSD plotter* software (Pettersen et al., 2006) to identify the grey matter boundaries (S2 Fig). Further analyses were performed only for grey matter channels.

Spike waveforms were extracted and classified using Kilosort software in Python (Pachitariu et al., 2016) to obtain single neuron spike time series, and to associate each neuron to a principal recording channel (Kilosort parameters include *Th*: [10,4], *minFR*: 1, *spkTh*: -5, *AUCsplit*: 0.9). Spike waveform clusters identified from Kilosort were manually accepted if the spike time correlograms showed a clear minimum near zero, spike amplitudes showed a normal distribution unique from the multi-unit spike amplitude distribution, and the shape of the spike waveform showed a clear distinction

from noise. The acceptance, rejection and merging of spike clusters were carried out manually in Phy-GUI which is integrated with Kilosort. Spike-sorted neurons having an average frequency <1.0 spikes/sec in response to natural image ensembles were considered to be unresponsive and discarded from further analysis.

Trial-to-trial response variability

As illustrated in Fig 1, a neuron's response to repeated presentation of a natural image movie typically varied across trials. For each neuron we estimated a measure of its trial-by-trial variability as follows. For each of the n=20 repetitions of the regularization set, the Pearson's correlation (*R*) between the neuron's response to that repetition and the average response for the (n-1) other repetitions of the same stimulus were calculated. R_i^2 provided a reliability measure of each neuron's response for a given trial *i* (*i=1, …, n*):

$$R_i^2 = \left[corr\left(r_i\,, \frac{\sum_{j=1, j\neq i}^{n} r_j}{n-1}\right)\right]^2 = \left[corr\left(r_i\,, \frac{\sum_{j=1, j\neq i}^{20} r_j}{19}\right)\right]^2 \qquad (1)$$

where r_i refers to the neuron's spiking response for the *i*-th trial, and $corr$ denotes Pearson's correlation. The average of these correlation coefficients was calculated:

$$\overline{R^2} = \frac{\sum_{i=1}^{n}(R_i^2)}{n} = \frac{\sum_{i=1}^{20}(R_i^2)}{20} \qquad (2)$$

This $\overline{R^2}$ value is similar in principle to the "reliability ratio" of a neuron (Borst & Theunissen, 1999; Lesica et al., 2007). The variability ratio for each neuron was then quantified as:

$$VR = 1 - \overline{R^2} \qquad (3)$$

VR values can range from zero (perfect reproducibility) to unity (maximal trial-wise variability).

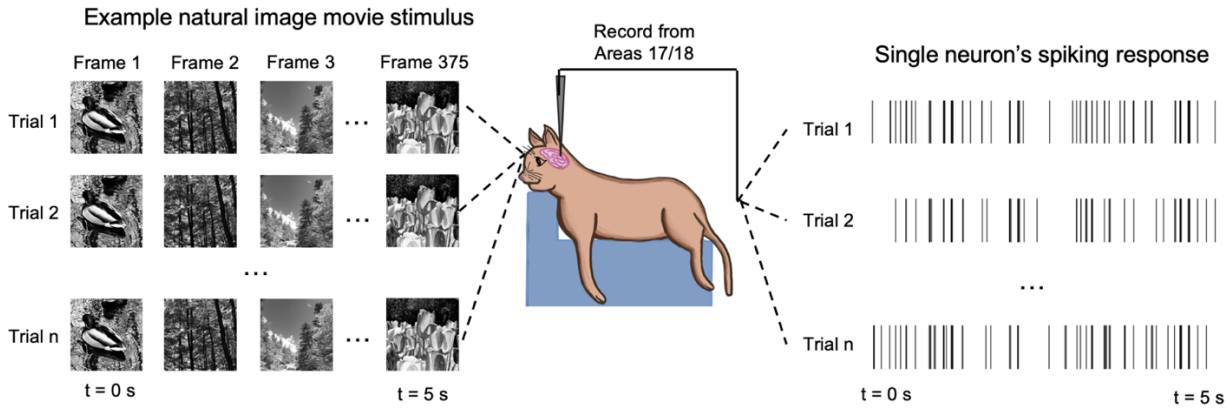

Fig 1: Trial-wise variability. Each natural image movie is presented across multiple (n) trials and the corresponding single neuron responses are compared to calculate a variability ratio, *VR* (Eq 1-3).

Spike-LFP phase locking

We analysed the relationship of each neuron's spiking activity to different LFP frequency bands, i.e. δ (0.5-4Hz), θ (4-7.5 Hz), α (7.5-12.5Hz), β (12.5-30Hz) and γ (30-100Hz) (Fig 2A and B). These LFP frequency bands were extracted in MATLAB by designing finite impulse response (FIR) bandpass filters using a Kaiser window applied to downsampled (x20) raw signals (i.e. for Plexon recorded data: from 31.25kHz downsampled to 1562.5Hz, OpenEphys recorded data: from 30kHz downsampled to 1500Hz). We then calculated a phase locking strength (PLS_f) of each neuron's firing to each filtered LFP band (Cui et al., 2016; Lachaux et al., 1999):

$$PLS_f = \frac{1}{N}\left|\sum_{k=1}^{N} e^{j\emptyset_f(t_k)}\right| \quad (4)$$

where *f* indexes the LFP frequency band, *N* is the total number of spikes for the neuron, and $\emptyset_f(t_k)$ is the phase of the LFP frequency band *f* at a given time t_k of the *k*-th spike. $\emptyset_f(\mathrm{t})$ was extracted using MATLAB's *hilbert* function.

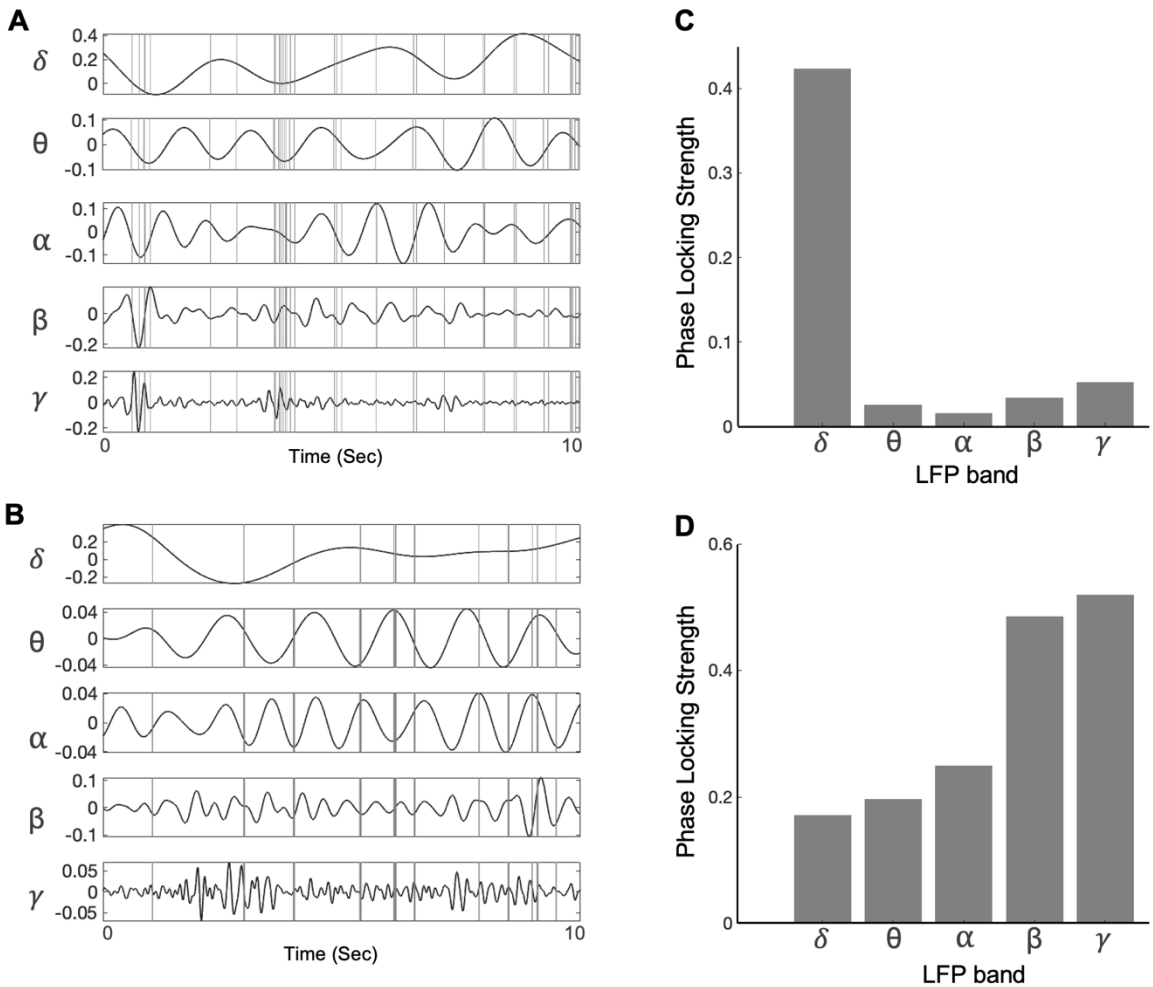

Fig 2: Spiking activity vs LFP frequency bands. A) Spiking activity of an example neuron 1 along with bandpass-filtered LFP signals recorded simultaneously. **B)** Similar to A, for example neuron 2. **C)** Phase locking strength (*PLS*) values (comparing spiking activity in relation to the phase of LFP bands) calculated for neuron 1. **D)** Similar to C, for example neuron 2.

Cortical state dynamics

Local field potentials (LFPs) and multi-unit activities (MUAs) show strong correlations with cortical state dynamics (Harris & Thiele, 2011). To identify the fluctuations of cortical state reflected by the population spiking activity in discrete time points, we extracted MUAs (level crossings at $\pm$ 5 SD of the >600 Hz high-passed signal) from each of the channels in grey matter. To avoid confusion with responses to the visual stimuli, MUA-based cortical state analysis was performed during the 4 blank screen periods between the 5 sweeps. These MUA spikes were binned (100 ms) to examine the up/down phases in network firing using a global fluctuation index (*GFI*, Eq 5) (Schölvinck et al., 2015). For each blank screen period, the binned firing rates were averaged across the channels to obtain channel-averaged firing rates (FR_{ch_avg}). The *GFI* was taken as the ratio between the standard deviation and the mean of FR_{ch_avg}:

$$GFI = \frac{SD\ (FR_{ch_avg})}{mean\ (FR_{ch_avg})} \tag{5}$$

To analyse cortical state fluctuations in a continuous time scale throughout the data recording session, we calculated a synchrony index (Saleem et al., 2010) based on the LFP signals from the middle channel in layer IV identified from CSD analysis. The power spectral density (PSD) was first computed using Matlab's *pspectrum* function (*type*: *'spectrogram'*, *FrequencyLimits*: [0 100], '*TimeResolution*': 30 and '*OverlapPercent*': 83.33) spanning throughout the data recording session. Then the Synchrony Index (*SI*) (Saleem et al., 2010) was calculated:

$$SI = \frac{P_{low}}{(P_{low} + P_{high})} \quad (6)$$

where P_{low} refers to power in the frequency range of [0.5-15Hz] and P_{high} refers to power in the frequency range of [15-100Hz] obtained from the PSD analysis. *SI* could vary from zero to unity, with higher *SI* values indicating a synchronized brain state and lower *SI* values indicating a desynchronized brain state.

Model architecture and parameter estimation

We introduce a convolutional neural network model including three parallel pathways: a stimulus-driven pathway to capture the stimulus-dependent neuron's response (R_{stim}), and two cortical state modulatory pathways with inputs from LFP and MUA signals to capture the variability of a neuron's response (Fig 3).

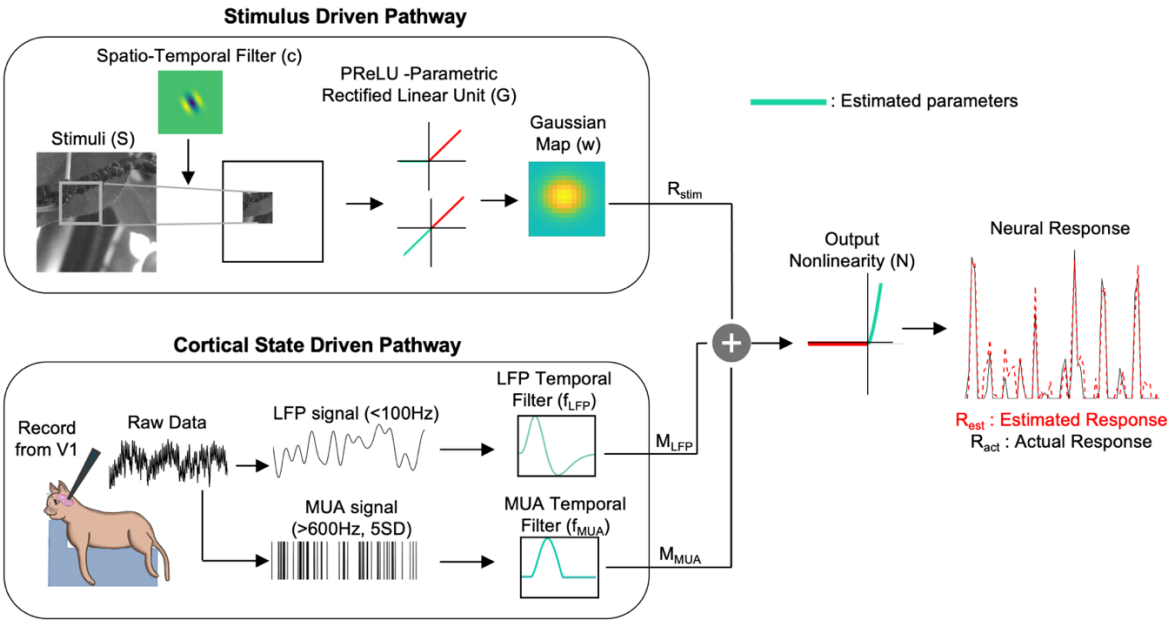

Fig 3: Model architecture for system identification. Visual stimuli (*S*) are convolved with a spatiotemporal filter (c), PReLU (*G*) and a gaussian map (*w*) to provide the stimulus-driven response (R_{stim}). Concurrent LFP and MUA signals provide inputs to the cortical state pathways, passing through LFP (f_{LFP}) and MUA (f_{MUA}) temporal filters to produce cortical state modulatory responses (M_{LFP} and M_{MUA}). These signals summate with R_{stim} and then are subjected to a final output nonlinearity (*N*) to produce the estimated neural response $R_{est.}$

Stimulus-driven pathway

Details of the model architecture for the stimulus-driven pathway (Eq 7) have been described previously (Nguyen et al., 2024), so here will only be briefly summarized.

$$R_{stim} = \sum_{x,y} w[x,y] G\left(\sum_{j,k,\tau} c[j,k,\tau] s[x-j, y-k, t-\tau]\right) \tag{7}$$

A cropped and downsampled stimulus *S* (see below) was first spatiotemporally convolved with the filter *c*. This was similar to the initial linear component of a linear-nonlinear model for a small region of the stimulus. The two-dimensional image output of this convolution, at a given time *t* was then passed through a static nonlinearity *G*, modelled using a Parametric Rectified Linear Unit (PReLU) with a single parameter $PReLU_{\alpha}$. The PReLU parameter, α, can take a value in a continuum from -1 to +1 indicating simple or complex type neural responses as well as intermediate types. The "map" layer *w* was modelled as a parameterized two-dimensional spatial Gaussian to identify the regions in visual space where the subunit filter *c* is active.

LFP-driven pathway

Phase locking strength (PLS; Eq 4) values measured on a subset of neurons indicated that the spiking responses of different neurons were phase-coupled at different strengths with the LFP components of different frequency bands (Fig 2B and 2C). Therefore, we included LFP signals containing all frequencies up to 100Hz and let the model parameter-fitting procedure (below) estimate unique temporal filters applied to the LFP, for each neuron.

The LFP signals were first downsampled 10-fold (fs_{down}= 3 kHz for OpenEphys and fs_{down}= 3.125 kHz for Plexon-recorded datasets). These downsampled signals were then low pass filtered (<100 Hz), line noise was removed (notch filter at 60Hz) and they were segmented into windows of length 150 time points (corresponding to T_{window} = 50 ms for OpenEphys and 48 ms for Plexon-recorded datasets). These LFP segments were computed for each frame of a natural image movie starting from the frame onset to a window length of 150 time points. Preliminary analysis with a subset of neurons indicated the sufficiency of using only 150 time points - longer window lengths did not improve the model prediction ability.

To ensure that the ability of the model to predict spiking activity from LFPs was not affected by any contamination from the spike waveforms, we avoided using LFPs from the same channel that the neuron was primarily recorded from (see Discussion). We evaluated the contribution of LFPs, gradually starting by incorporating LFPs from a nearby (located 100 μm away) channel to incorporating LFP from all the channels (except the same channel) in grey matter. For recordings with polytrodes (having 2

columns of 16 recording sites), we used LFPs from the channels which are located along the same column as the recorded neuron, and we used every other channel (located 100μm apart) to be consistent with the linear array and edge recordings.

MUA-driven pathway

A neuron's firing can be correlated to different degrees with nearby population firing, independent of stimulus conditions, which might occur due to non-sensory global variables including cortical state fluctuations (Okun et al., 2015). Therefore, we incorporated MUA from two nearby channels (100μm above and below) and included a MUA-driven cortical state pathway in our overall model. Unlike the LFP-driven pathway, here we did not use MUA signals from all the channels, because a given neuron's firing correlates primarily with population firing that originates nearby (Cui et al, 2016). This idea was corroborated here by preliminary analyzes in a subset of neurons – the prediction accuracy of the system identification model did not further improve by including more MUA channels, and including all the channels resulted in estimated MUA temporal filter weights close to zero in channels further away from the primary channel (S3 Fig). The MUA from the same channel as the recorded neuron was always excluded to avoid potential signal contamination.

Signals from the two nearby channels were high pass filtered (>600Hz), and spikes exceeding a threshold of 5 standard deviations away from the mean were detected to generate the MUA signal. This MUA signal was segmented into windows of 1500 time points (corresponding to T_{window} = 500 ms for OpenEphys and 480 ms for Plexon-recorded datasets) starting at the onset of each natural image frame. These

segments were downsampled 10-fold to yield segments of 150 data points in length (corresponding to T_{window} = 50 ms for OpenEphys and 48 ms for Plexon-recorded datasets), similar to those from the LFP processing. An important issue is that in some cases, the spike waveform for one (sorted) neuron can appear in multiple channels. To avoid this kind of potential contamination, such "spillover" spikes were removed from the nearby channels before computing the MUA segments (S3 Fig).

Model architecture

LFPs and MUA, recorded simultaneously with spiking activity, both showed variations in *SI* and *GFI* (explained above) in our recorded data. We used these brain signals as indicators of cortical states to capture the variability in neural responses and boost the predictive accuracy in estimating neural responses. To implement this, frames of natural image movies and the corresponding LFP and MUA segments were fed in parallel to the stimulus-driven pathway and LFP- and MUA-driven pathways of the model, respectively.

Overall activity from the cortical state-related signals was combined additively with responses from the stimulus-driven pathway, and subjected to a final output nonlinearity (*N):*

$$R_{est} = N\,(R_{stim} + M_{LFP} + M_{MUA}) \tag{8}$$

The output nonlinearity was modelled as a rectified power law, with gain (*g*) and exponent (*exp*):

$$N(x) = \begin{cases} g.x^{exp}, & if\ x > 0 \\ 0 & if\ x \leq 0 \end{cases} \quad (9)$$

We also evaluated models in which the cortical state signals acted multiplicatively (Ecker et al., 2014; Goris et al., 2014), or with divisive normalization (Carandini & Heeger, 2011), but additive modulation of stimulus-driven neural activity by the cortical state-driven activity provided higher predictive accuracy (S4 Fig).

Model parameter estimation

Optimization of the model parameters was implemented in Python using tools from the TensorFlow 2.0.0 (Abadi et al., 2016) and Keras 1.1.2 (Chollet & others, 2018) packages. To avoid overfitting, we applied L2 regularization to the model coefficients from both LFP and MUA temporal filters and incorporated early stopping in the complete model with a patience value of 50 epochs. Preliminary analysis showed that the use of L2 regularization in the stimulus-driven pathway did not bring additional improvements in prediction accuracy. Model parameters were optimized as explained below.

First, a three-pass cropping procedure (Nguyen et al., 2024) was employed to identify regions of the stimulus that encompass the receptive field. We initially estimated the parameters of the model containing only the stimulus-driven pathway using full 480x480 stimulus images downsampled to 30x30 in the first pass. During the second pass, a square-shaped cropping window enclosing an area slightly larger than the apparent receptive field was identified and downsampled to 30x30 to re-train the model. In the final pass, we adjusted the cropping window based on the model estimated from the second pass to identify more accurate boundaries of the receptive field. We imposed a minimum cropping window size of 120 pixels to handle very small receptive

fields. The final cropped stimulus from the third pass was downsampled to 30x30 to be used in further analyses as input to the stimulus-driven pathway.

The two parameters of the final output nonlinearity (Eq 9) were estimated in two stages, as previously described (Nguyen et al., 2024). In the first stage we set *N* to be a simple ReLU activation function and obtained a preliminary estimate of the receptive field and filters for the LFP and MUA signals. For this optimization, we initialized the spatiotemporal receptive field with Glorot normal initializer, PReLU α at 0.5 and the Gaussian map layer at the center of the image with standard deviation equal to the length of the image. For the second stage, we used the pre-trained receptive field weights from the first stage as initial conditions of the spatiotemporal receptive field (with other model parameters initialized similar to the first stage), and re-trained the model with *N* set to be an adaptive exponent output nonlinearity defined in a custom Keras layer with two trainable parameters, gain (*g*) and power law exponent (*exp*) as in Eq 9.

The models' predictive accuracy for the hold-back test dataset was measured by the variance accounted for (*VAF*), computed as the square of the correlation between the neuron's measured response (R_{act}) and the response of the estimated model (R_{est}) expressed as a percentage:

$$VAF = [(corr(R_{act}, R_{est}))^2 \; x \; 100 \; \%] \qquad (10)$$

To quantify the trial-wise response variability explained by our cortical state-driven pathway we calculated a "$VAF_{ongoing}$". For this, we calculated the variable portion of the trial-wise response which is considered as the "ongoing activity" (Schölvinck et

al., 2015). The recorded ongoing activity ($Ongoing_{act}$) for the test dataset was derived by taking the difference between each trial's response and the trial-averaged response. The trial-averaged response is considered as the visual evoked response.

$$Ongoing_{act} = Trial\ Response - Trial\ averaged\ response \quad (11)$$

From the system identification model, we calculated the difference between the predicted responses from the full model containing the stimulus and cortical state-driven pathway and from the base model containing only the stimulus-driven pathway, which we term the estimated ongoing activity by the model ($Ongoing_{est}$).

$$Ongoing_{est} = Estimated\ response_{full\ model} - Estimated\ response_{base\ model} \quad (12)$$

By comparing these two, we obtained the $VAF_{ongoing}$:

$$VAF_{ongoing} = [(corr(Ongoing_{act}, Ongoing_{est}))^2\ x\ 100\ \%] \quad (13)$$

The $VAF_{ongoing}$ can vary between 0% where the cortical state-driven pathway doesn't account for the observed ongoing activity to 100% where the cortical state-driven pathway completely account for the observed ongoing activity.

Quality of receptive field estimates

We obtained receptive field restorations (henceforth termed "receptive fields") by convolving the filter function (c) with the map function (w) (Nguyen et al., 2024). To quantify the statistical quality of estimated receptive fields, with or without incorporating brain state fluctuations, we calculated the z-score at the peak time lag of the receptive field using:

$$z_{i,j} = \frac{x_{i,j} - \mu}{\sigma} \tag{14}$$

$$Z = var(z) \tag{15}$$

Here, $x_{i,j}$ is the *i,j-th* filter weight for the peak time lag of the receptive field, μ is the mean value of all the filter weights in the peak time lag, and σ is the standard deviation of the filter weights in the zeroth time lag of the spatiotemporal receptive field. Here z is a 2D array with values $z_{i,j}$ corresponding to each filter weight in the peak time lag. The variance of these values (Eq 15) was utilized in comparisons between models.

In the calculation above we used σ from the zeroth time lag as a measure of background noise of the estimated receptive fields (Eq 14). Since this lag is only nominally zero, but actually represents a time bin of 13.3 msec, it seemed prudent to check if there might be significant filter weights present at this lag depending on whether we incorporate brain state effects. To do this we confirmed the similarity of σ in the zeroth time lag of the estimated receptive fields before and after incorporating cortical state effects by using the Wilcoxon rank-sum test (S5 Fig).

Dimensionality reduction of the filtered LFP power spectra

To examine the LFP frequencies of importance for different neurons, we further analysed the power spectra of LFPs filtered by the estimated temporal filters for neurons which showed a VAF improvement >5% (S6 Fig). We used the UMAP python package (McInnes et al., 2018) for nonlinear dimensionality reduction of the normalized power spectra (magnitude varying from 0 to 1), with parameters *n_neighbors=20, min_dist=0.1,* to obtain a high-dimensional graph. The graph provided by UMAP was

then used for the Louvain community detection algorithm (Blondel et al., 2008) to identify distinct clusters (*resolution = 1.0*). The *resolution* parameter was selected to maximize the modularity score which is a measure of the quality of clusters (S7 Fig). For visualization purposes, we projected the high-dimensional graph into two dimensions using UMAP. We compared the variability ratios and the VAF improvements of neurons in different clusters provided by Louvain clustering, to assess differential effects of different LFP frequency power spectra.

Results

We recorded broadband neural data from 31 penetrations in Areas 17/18 of 14 anesthetized and paralyzed cats. The recorded data was used in multiple previous studies in the same lab (Gharat & Baker, 2017; Nguyen et al., 2024; St-Amand & Baker, 2023). The present study was based on analysis of 154 spike-sorted neurons (131 from OpenEphys and 23 from Plexon data acquisition systems) meeting the inclusion criteria (spike frequency >1.0 spikes/sec in response to natural image movies), along with simultaneously recorded LFP and MUA signals.

Trial-to-trial response variability

We often observed single neuron spiking responses which differed across repeated presentations of the same stimulus. We quantified the extent of these trial-wise differences by a “variability ratio (*VR*)” (see Methods) which compares the similarity of individual trial trial-wise responses to the trial-averaged response. Fig 4A shows spikes from an example neuron with a relatively low variability in response to repeated

presentation of a 5 sec natural image movie. For this neuron the trial-averaged response and individual trial-wise responses were comparatively similar, yielding a low *VR* of 0.31. However, another example neuron as shown in Fig 4B exhibited much greater variability (*VR*=0.87).

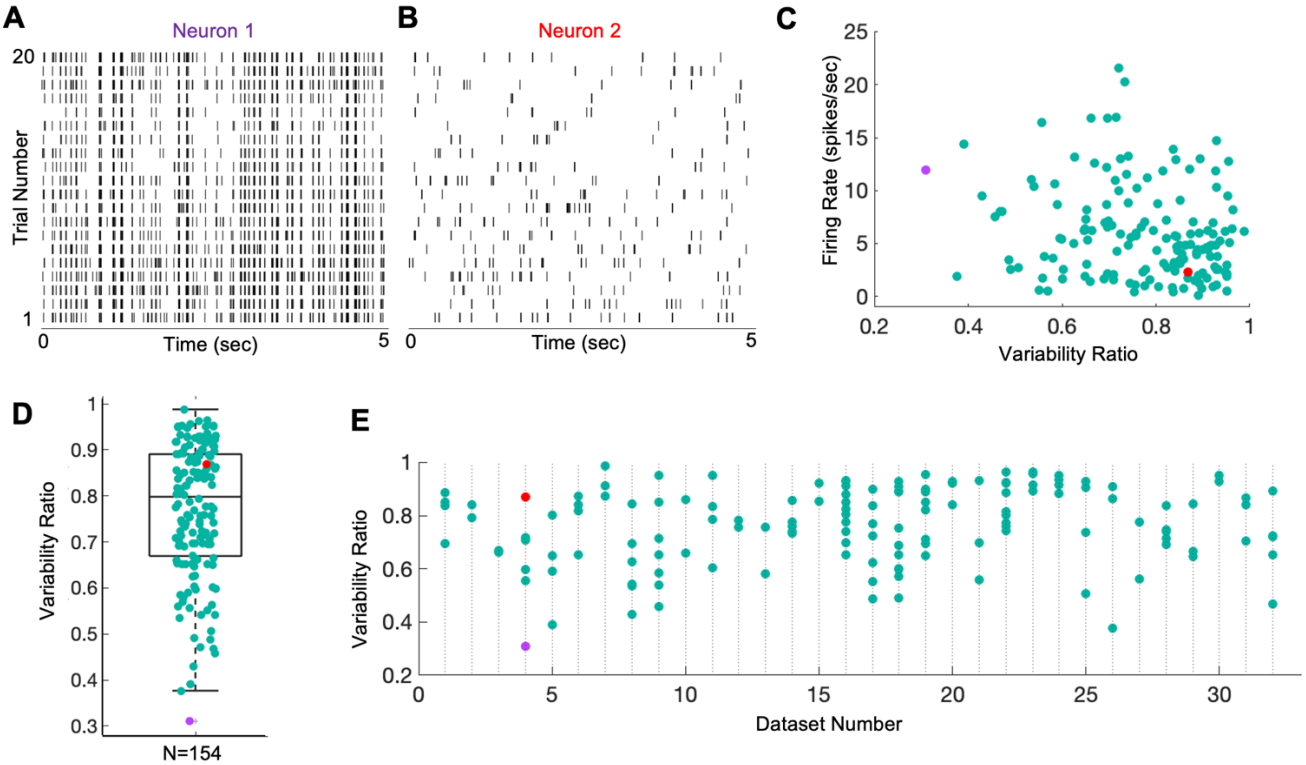

Fig 4: Trial-wise response variability of neurons. A) Spiking activity of a neuron with a low trial-to-trial variability (VR=0.31, the same neuron as in Fig 2A). **B)** Spiking activity of a neuron with a higher trial-to-trial variability (VR=0.87, same neuron as in Fig 2B) recorded simultaneously with the neuron in panel A. **C)** Scatter plot for the population of neurons (n=154) showing firing rate (FR) plotted against VR. Correlation coefficient between FR and VR is -0.21 and p=9.9e-03. **D)** Distribution of variability ratio across the population of neurons (n=154). **E)** Distribution of variability ratios of neurons across different datasets containing at least 2 simultaneously recorded neurons. In C, D, and E, each data point on a column shows *VR* of a simultaneously recorded neuron in that dataset. Example neuron 1 is shown in purple and neuron 2 in red.

Comparing the two example neurons in Fig 4A and B, one might suppose that higher VRs and found in neurons with lower firing rates. To examine whether VRs are related to the mean FRs across the sample population, a scatter plot is presented in Fig 4C for the 154 neurons. While there is little evident relationship, statistical analysis shows a slight though significant negative correlation between the mean FRs and VRs (n=154, r= -0.18; p=0.03), suggesting that neurons with a low mean FR tend to have a higher VR and vice versa.

The *VR*s across the sampled population of neurons (n=154) are displayed in Fig 4D, showing *VR*s ranging from 0.31 to 0.98 with a mean of 0.73 (median 0.79) and standard deviation of 0.14 (inter-quartile 0.21). The high mean value of VRs suggest that the neural spiking activities are poorly predictable or explainable from the visual stimulus inputs alone. Traditional system identification models usually ignored such variabilities and sought only to predict the trial-averaged firing rates (Nguyen et al., 2024).

It would seem conceivable that neurons with diverse VR values were those recorded at different times of higher or lower variability - in that case, one might expect that simultaneously recorded neurons would have similar VRs. Fig 4E shows a scatter bar plot of neuronal *VR*s for 33 datasets containing at least 2 simultaneously recorded neurons. In many instances, even the simultaneously recorded neurons showed very different amounts of trial-to-trial response variability, ranging from a maximum difference of 0.56 (dataset number 4 of Fig 4E) to a minimum of 0.01 (dataset number 3 of Fig 4E). This result shows that the sources causing the response variabilities can affect

simultaneously recorded neurons differentially, and therefore neurons should be treated at an individual level to account for the observed response variabilities.

Cortical state dynamics

The fluctuations of cortical network state might give rise to trial-to-trial response variability (Harris & Thiele, 2011). We thus analysed MUA and deep layer LFPs, as indicators of cortical network state fluctuations, using indices based on the firing patterns of simultaneously recorded populations of neurons and the LFP power in different frequency bands. During a synchronized brain state, MUA shows prominent "up phases" where a majority of neurons are active and "down phases" where they are relatively silent (Destexhe et al., 2007; Harris & Thiele, 2011; Wilson, 2008). Therefore, we analysed MUA signals during periods of spontaneous activity when the neuronal firing patterns are not affected by visual stimuli.

LFP signals show high power in low frequency ranges and low power in high frequency rages during synchronized states and vice versa during desynchronized states (Harris & Thiele, 2011). Since LFP signals (<100Hz) are less affected by high frequency firing due to visual stimuli, we analysed LFPs to continuously measure brain states throughout a recording session, complementary to MUA that provided a discrete measure of cortical states.

We computed MUA firing rate (100ms bins) for each of the channels in the grey matter during the four inter-sweep intervals, when no visual stimuli were presented. In some intervals (e.g. 1, 3, and 4), firing activities showed clear synchronization across channels, firing during "up phases" and being completely silent during "down phases".

However, firing patterns in interval 2 were much less synchronized to the repeated stimuli, indicative of a relatively desynchronized state. The average firing rates across the channels (Fig 5B) confirmed the above impression that intervals 1, 3 and 4 show pronounced fluctuations, and higher peaks for the “up phases”, than interval 2. The synchronized vs. desynchronized phases are further confirmed by quantifying a global fluctuation index, *GFI*, (Schölvinck et al., 2015), defined as the ratio of the standard deviation (SD) to the mean of channel-averaged firing rate (see Methods, Eq 5), in Fig 5C (see also S8 Fig B, D, and F). Interval 2 showed a much lower GFI value than intervals 1, 3, and 4. Although these GFI values are calculated from responses when no visual stimuli were presented, previous studies show that the ongoing cortical state fluctuations can outlast visual stimulus presentations (Schölvinck et al., 2015).

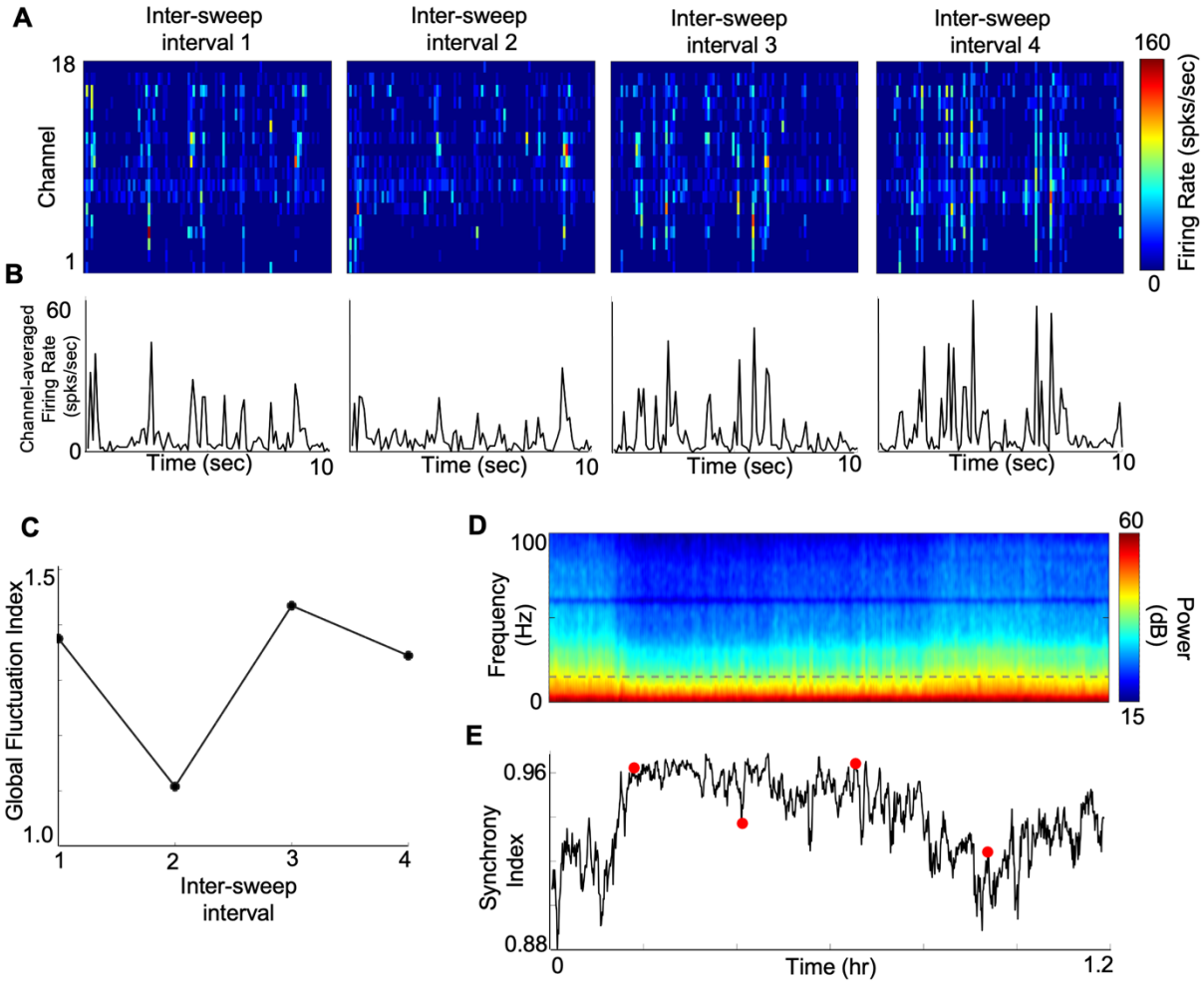

Fig 5. Cortical state fluctuations within data recordings. A) Binned firing rate of MUA across 18 channels in the grey matter (100 ms bins) during inter-sweep intervals while presenting a uniform grey screen. **B)** MUA firing rate averaged across channels in panel A. **C)** Global fluctuation index (GFI, Eq 5) using channel-averaged MUA firing rates in panel B. **D)** Power spectral density (PSD) of a deep layer LFP (middle channel of layer IV) across a data recording session (from same dataset as A-D). **E)** Synchrony index (SI, Eq 6) using PSD of panel D. Red data points indicate inter-sweep intervals.

Cortical state dynamics observed in terms of GFI could be obtained only at discrete times (i.e. during inter-sweep periods without any visual stimulation) during a data recording session. To identify network fluctuations in a more continuous manner,

we analysed LFP signals spanning throughout a data recording session. Time frequency analysis of a deep layer LFP showed that its power spectral density varied greatly at different frequency ranges and over time (Fig 5D, results from same recording session as for Fig 5A-C). We defined a synchrony index (*SI*) as the ratio of LFP power in low frequency ranges (0.5-15Hz) divided by the total power in low and high (0.5-100Hz) frequency ranges (Saleem et al., 2010). Overall, SI values in this example varied in a range approaching unity (range of 0.88 and 0.96) which could be due to the global cortical state under anesthesia. However, we observed continuous fluctuations of SI across time during the recording session (Fig 5E, also see S8 Fig for other examples) showing small rapid variations of network state dynamics which might result in the observed trial-to-trial spiking response variability. We observed a similar pattern of changes in the cortical state fluctuations from MUA and LFP when we overlaid them (and also in the examples in S8 Fig), however, the correlation between the two measures is not statistically significant (Fig 5C, E; $r=0.58$, $p=0.43$, $n=4$).

Consequently, we analysed how much trial-wise spiking response variability might be predicted given sufficient knowledge of the ongoing cortical dynamics, as inferred from LFP and MUA recordings using system identification (see below).

System Identification Model

To examine how cortical state variations might influence neuronal activity, we incorporated candidate brain state indicators in a system identification model optimized to give best predictive performance of the neuronal spiking responses to visual stimuli.

We developed the model (Fig 3) in stages incorporating successively more indicators of brain state.

Stimulus-driven (base) model

First, we consider the performance of the model for the stimulus-driven pathway alone (upper part of Fig. 3), without incorporating brain state effects, which we term the "base model". With the recorded neural data and the corresponding stimulus input, the model parameters estimated for the example neuron 1 are presented in Fig 6A. The estimated model explained well the responses of this neuron (VAF 48.9%). The estimated filter weights contained a central inhibitory region surrounded by an excitatory region peaking at short time lags, which is reversed at larger time lags. The map layer, parameterized as a Gaussian, localized the activity of the filter. This example neuron had an PReLU alpha value of 0.62, corresponding to a nearly linear relationship indicative of a simple-type cell. The receptive field (i.e., the result of convolving the spatiotemporal filter weights with the Gaussian map) appeared less noisy and had a more clearly delineated excitatory surround. The convolutional output from the map layer passed through the final nonlinearity (exponent 0.96) to generate the final neural response.

The base model estimates for the second example neuron provided a VAF of 5.23% which was much lower than that for neuron 1. Fig 6B shows the spatiotemporal filter estimate (top row), PReLU alpha of 0.05 (indicative of a complex-like cell), estimated gaussian map, final output nonlinearity with exponent 1.50, and the receptive field (bottom row). The estimated receptive field for this neuron was spatially Gabor-like,

i.e. with elongated, alternating excitatory and inhibitory regions, which reversed polarity at larger time lags.

Neurons with VAF<1% using the model consisting of stimulus-driven pathway only, were discarded from further analysis to help ensure reliability of the estimated model parameters.

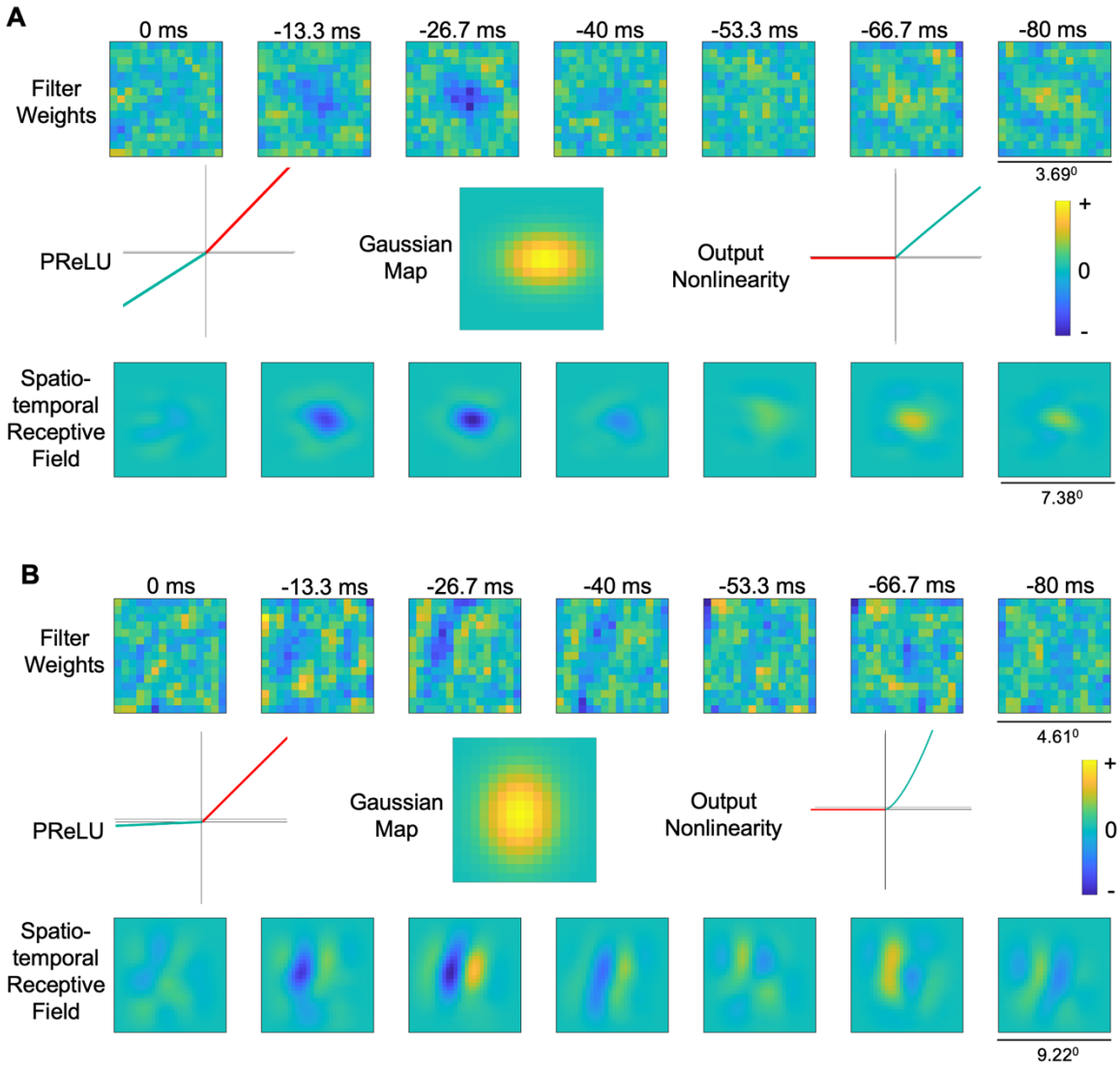

Fig 6: Estimated parameters for the stimulus-driven pathway for two example neurons. A) Estimated parameters from the stimulus-driven pathway for example neuron 1 (same neuron as in Fig 2B); spatiotemporal filter weights (top row), PReLU

alpha, Gaussian map, Output nonlinearity and the final restoration of the receptive field (bottom row). **B)** Same as A, for example neuron 2 (same neuron as in Fig 2C)

We reasoned that the lower VAF for the second neuron compared to the first neuron could be a result of its higher trial-to-trial response variability (Fig 4A, B). Thus, we next incorporated the influence of cortical state indicators to see whether there are improvements in the model performance, i.e. predictive VAF accuracy.

Performance improvement with cortical states estimated using LFPs

Since cortical state fluctuation patterns can be detected with LFPs recorded simultaneously with spiking activity (Banerjee et al., 2012; Kelly et al., 2010), we incorporated LFP signals in a cortical state-driven pathway, parallel to the stimulus-driven pathway in our model architecture (Fig 3) to predict trial-to-trial response variability. For each of the neurons we re-trained the base model with the cortical state-driven pathway, and the optimization algorithm estimated a temporal filter to enhance LFP frequencies from a nearby channel related to the trial-wise variable response.

The estimated temporal filter for the first example neuron (Fig 7A) had a relatively low gain ($\pm$ 0.001) with little evident temporal structure. However, the LFP temporal filter for the second neuron (Fig 7B) was ~20 times larger in gain ($\pm$ 0.02), with clear positive weighting of shorter time lags, and a negative peak around 25ms. The filter weights for the first example neuron oscillated around zero for all the time lags and therefore, it would dampen all the LFP signals passing through it. For the second example neuron, the measured temporal dynamics could retain relevant frequencies for the neuron and filter out less important frequencies.

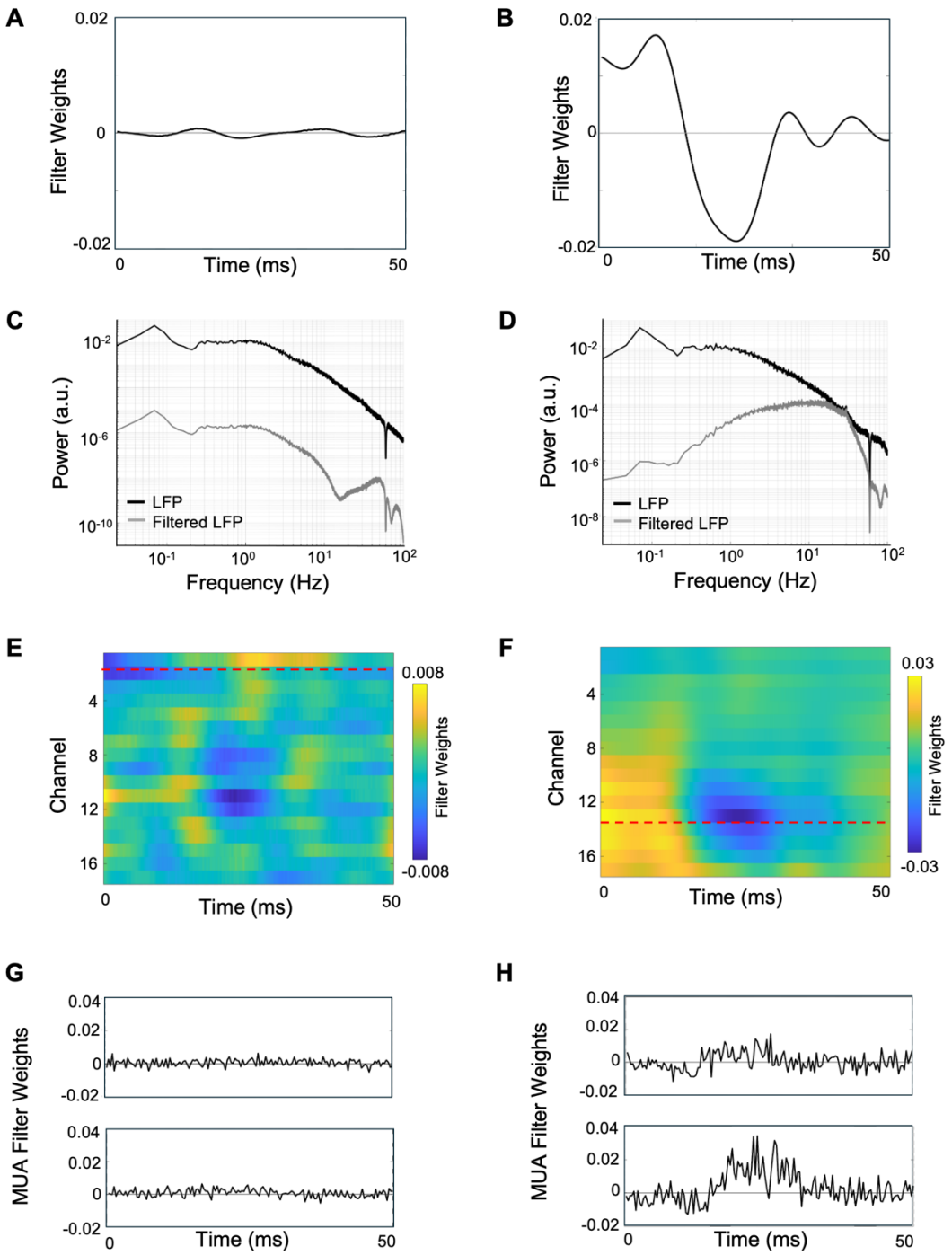

Fig 7: Estimated parameters for the cortical state-driven pathway. **A)** Estimated temporal filter from LFP-driven pathway for example neuron 1 when using LFP only from the channel above. **B)** Similar to A, for example neuron 2. Note ordinate scales are same for A and B, to facilitate comparison. **C)** Frequency spectrum of LFP before

(black) and after (grey) applying the temporal filter shown in panel A. **D)** Similar to C, for neuron 2. **E)** Estimated temporal filters from LFP-driven pathway for neuron 1 when using LFPs from all the grey matter channels. Dashed red lines indicate recording channel of the neuron. **F)** Similar to E for neuron 2. **G)** Estimated temporal filters from MUA-driven pathway for neuron 1 when using MUA from the above and below channels. **H)** Similar to G for neuron 2. Note ordinate scales are same for G and H. (See also S6, S9 Fig.s for additional examples).

To identify the relevant frequency bands of importance for each neuron we examined the frequency spectrum of the LFP filtered through the estimated temporal filter. For the first example neuron, the filtered LFP had a similar power spectrum as the raw LFP, but with a broadband reduction in power (Fig 7C). The power spectrum for this filtered LFP had higher power in the delta frequency range (0.5-4Hz), comparable with the LFP-spike phase locking values for the neuron, showing an abundance of spikes phase-locked to the delta band (0.5-4Hz) (Fig 2C). For the second example neuron, the filtered LFP spectrum's shape was markedly different compared to the raw LFP power spectrum, with greater power in the higher frequency ranges including beta (15-30Hz) and gamma (30-100Hz) compared to the lower frequencies (Fig 7D). This behavior was comparable with the LFP-spike phase-locking values for the neuron, showing spikes more phase-locked to the higher frequencies (Fig 2D). Overall, phase-locking strengths for the second neuron had relatively larger values (0-0.6) compared to the first neuron (0-0.4), showing higher spike-LFP coherence for the second neuron (Fig 2B, C).

Consistent with these observations, the predictive VAF accuracy for the second neuron improved from 5.23% to 8.14% when the LFP pathway was included in the

model. This improvement was higher compared to the first neuron, which only improved from 48.9% to 49.1%. These observations suggested that the second neuron shows higher cortical state effects compared to the first neuron, as evident from its higher trial-to-trial variability ratio and higher VAF improvement with the incorporation of LFP in the system identification model.

To further examine how neurons varied in terms of the LFP frequencies filtered by the estimated temporal filters, and whether this was related to the VAF improvements, we examined the power spectra of filtered LFPs of all the sampled neurons which showed VAF improvements >5%. We excluded neurons with smaller VAF improvements since their temporal filter values were close to zero (e.g. Fig 7A) - the filtered LFP spectra for a representative subset of the remaining n=61 neurons are shown in S6 Fig. Non-linear dimensionality reduction of the filtered LFP power spectra (see Methods) identified three distinct clusters (Fig 8A). One cluster showed a low-pass filtered LFP power spectrum ("class 1": n=22, Fig 8A-B, purple), while a second group showed a broader band-pass filtered spectrum ("class 2": n=15, Fig 8A-B, yellow) and a third cluster showed a narrow band-pass filtered spectrum ("class 3": n=24, Fig 8A-B, green). However, the VAF improvements and the variability ratios of the neurons in these three classes of filtered LFP power spectra were not significantly different (Fig 8C-D, $p>0.05$). These observations suggest that different neurons can show a range of response variabilities and VAF improvements independent from the LFP frequency ranges that influence them.

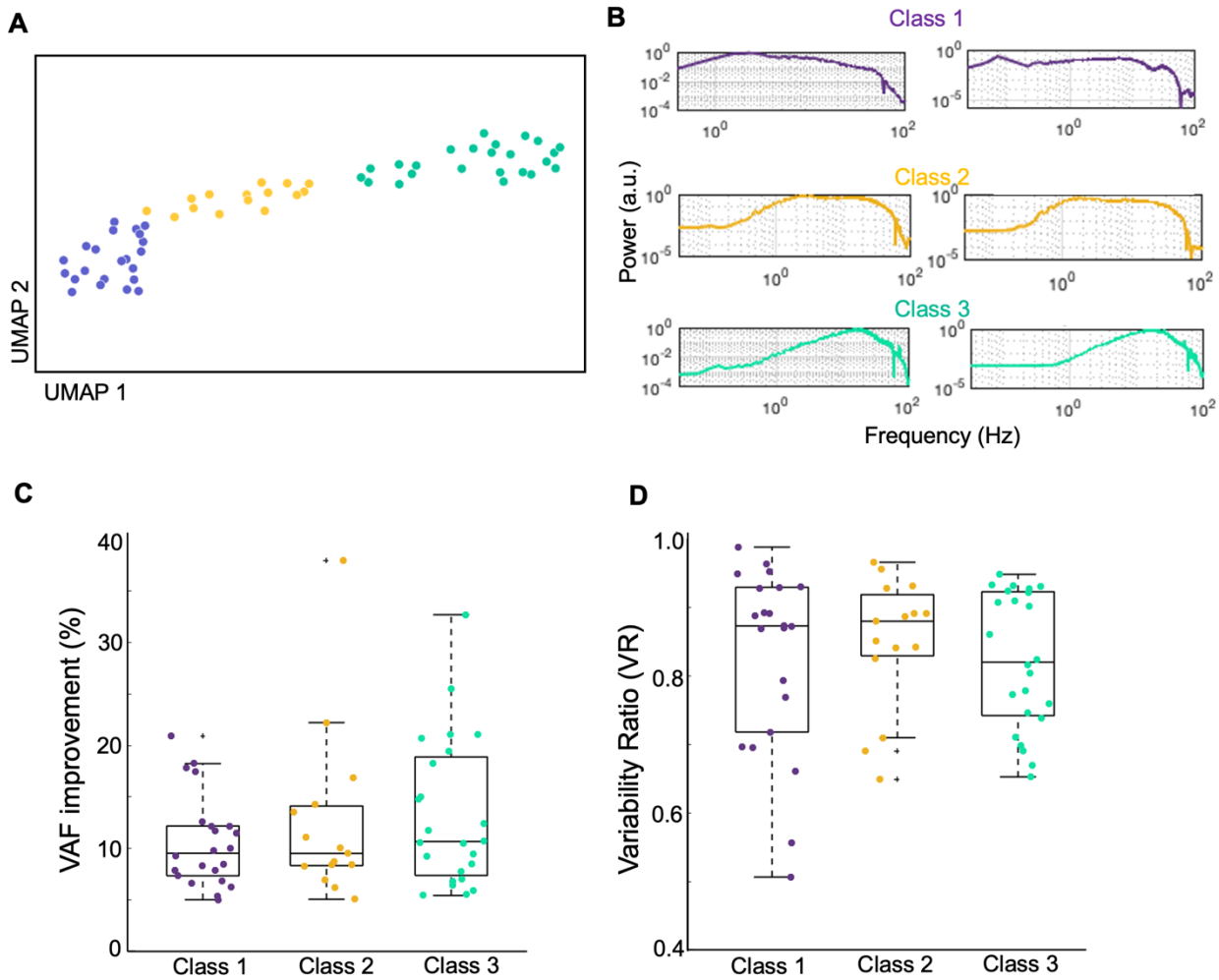

Fig 8: Diversity of filtered LFP power spectra. **A)** Scatter plot of filtered LFP power spectra in UMAP space colored by Louvain cluster membership. Neurons with a VAF improvement > 5% (n=61) were used in this analysis. **B)** Two examples for each class of power spectra clusters in panel A. **C)** VAF improvements of neurons which belong to the 3 classes. **D)** Similar to C, for the variability ratios of neurons.

In the above analysis, only the LFP from one nearby channel was used in the model. We evaluated further improvements in VAF by including LFPs from all the channels in grey matter (except the primary channel for the recorded neuron). Re-training of this model including the cortical state-driven pathway with all LFP channels

provided separate temporal filters for each channel. Estimated filter weights for the first example neuron (Fig 7E) had lower magnitudes ($\pm$ 0.008) for the first neuron compared to the second neuron (Fig 7F, $\pm$ 0.03). The filter weights showed clear spatial distribution for the second neuron, i.e., having higher magnitude weights at recording sites distributed around the primary channel for the neuron (red dashed line in Fig 7F). This spatial distribution of estimated weights was not observed for the first neuron (Fig 7E). The model prediction accuracy (VAF) improved by 2-fold for the second neuron (VAF 5.23% to 11.55%) while negligible improvement was observed for the first neuron (VAF 48.9% to 49.9%). Thus, these results suggest that the variable portion of a neuron's response might be caused by cortical state dynamics which are reflected in the LFP. Further, the second neuron which had a higher *VR* benefited more from the LFP-driven pathway, while showing strong coupling of the spiking response to ongoing network activity across nearby channels.

Contribution of MUA-based cortical state to model performance

Recent work has shown that average firing rate of a population of simultaneously recorded nearby neurons can be predictive of the observed trial-to-trial response variability of individual neurons (Okun et al., 2015). We evaluated whether the predictive performance for a single neuron's spiking activity improves by incorporating MUA-based cortical state dynamics measured from two nearby channels (i.e., 100μm above and below) in an additional parallel pathway in our model. MUA from the same channel was not used, to avoid potential contamination from the recorded neuron itself. Further signal "cleaning" was carried out in nearby MUA signals to avoid the effect of spikes

from the same neuron appearing in multiple nearby channels (see Methods; S3 Fig). Preliminary analysis showed the sufficiency of using only two MUA channels - the predictive accuracy with just two channels was higher (or similar) compared to that using additional channels. The MUA filter weights located further from two channels were close to zero in magnitude (S3 Fig), so including them in the model can cause the predictive accuracy to drop due to the increase of the number model parameters. Therefore, we used MUA signals from only two nearby channels separated by 100μm from the neuron's recorded channel and re-trained the model with the stimulus-driven pathway along with the pathway containing LFP signals from all the channels.

The pair of MUA temporal filters from two channels for the first example neuron (Fig 7G) were very small (+0.006 / -0.004) compared to those for the second neuron (Fig 7H) (+0.04 / -0.02). The temporal filters showed prominent peaks at latencies around 25-30ms for the second neuron, but this effect was not observed for the first neuron. The VAF accuracy further improved for the second neuron to 17.4%, while negligible improvement was observed for the first neuron (VAF = 49.5%). Thus, the second example neuron benefited comparatively more from inclusion of the MUA-driven pathway, similar to the effect from incorporating the LFP-driven pathway. Estimated LFP and MUA temporal filters for additional example neurons are shown in S6, S9 Figs.

Improvement in estimated spatiotemporal receptive fields

Given the observed VAF improvement for the second example neuron compared to the first neuron, we examined whether we could also observe parallel improvements in the estimated receptive fields. We evaluated the spatiotemporal receptive fields (i.e.,

restorations by convolving the spatiotemporal filter weights by the Gaussian map), while gradually incorporating the cortical state effects from both LFP and MUA signals.

Fig 9 shows, for the two example neurons, estimated receptive fields when using models with the stimulus-driven pathway alone (base model) in the first row, with LFPs from a nearby channel (Model 1) in the second row, with LFP from all the channels (Model 2) in the third row, and with the full model incorporating LFP from all the channels and MUA from two nearby channels in the fourth row (Model 3). For the first example neuron, we did not observe qualitative improvements in the receptive field (Fig 9A), which was not surprising in view of the absence of significant improvements in predictive performance (VAF=48.9% with the stimulus-driven pathway alone and only VAF=49.6% with the full model). However, for the second example neuron, we observed clear qualitative improvements in the receptive field after incorporating the cortical state effects (Fig 9B). This is evident from both the clear localization of the inhibitory and excitatory regions in the peak time lag at -26.7 ms, and clearer delineation of excitatory and inhibitory regions at the larger time lags (e.g. appearance of temporal reversal of the receptive field at time lag of -66.7 ms), after incorporating LFP- and MUA-driven pathways rather than simply using the stimulus-driven pathway (Fig 9B). This is also evident from the z-score values for the receptive fields of the two example neurons at the peak time lag. Neuron 1 had a z-score of 1.13 from the base model and 1.17 from model 3 (z-score improvement=0.04), while Neuron 2 had a z-score of 1.55 from the base model and 1.79 from model 3 (z-score improvement=0.24). These examples suggest that for neurons with higher variability, accounting for the cortical

state dynamics can not only improve the model predictive accuracy (VAF), but provide better visualizations of the spatial receptive fields and their temporal dynamics.

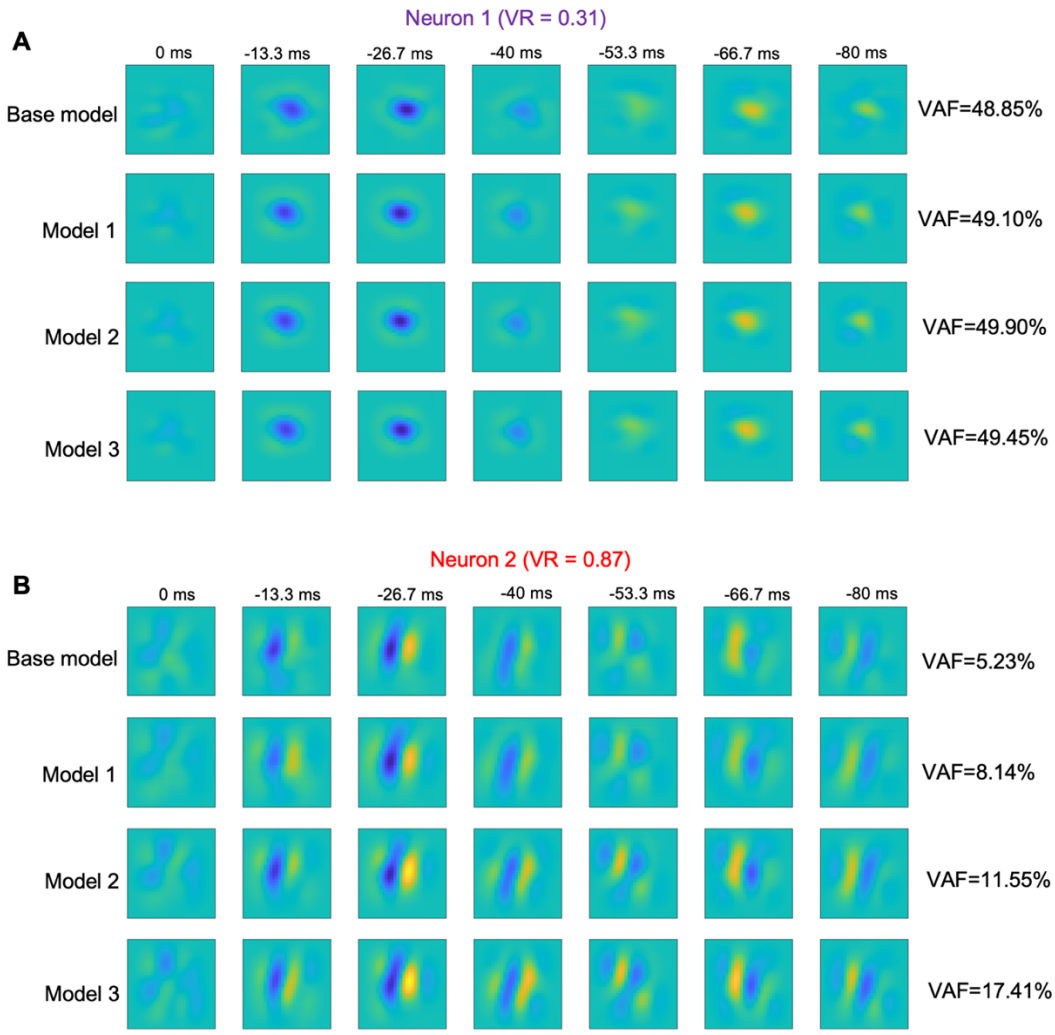

Fig 9: Improvements in receptive field estimates with cortical state. A) Improvement of spatiotemporal receptive fields with the incorporation of the brain state effects for example neuron 1. Base model is the system identification model with only the stimulus-driven pathway, Model 1 incorporates cortical state effects from one LFP channel located 100μm away, Model 2 incorporates all the LFP channels and Model 3 additionally incorporates two nearby MUA channels. **B)** Similar to A, for example neuron 2.

VAF improvements across the population of neurons

The initial VAF of the neurons with only the stimulus-driven pathway (base model) ranged from 0% to nearly 50%. At this stage, we discarded neurons with VAF<1% from further analysis to help ensure reliability of the estimated model parameters, leaving n=154 neurons. The improvements in model predictive power for this sample population of neurons is shown in Fig 10. With the stimulus-driven pathway alone (base model), neurons exhibiting a lower variability ratio (*VR*) had higher predictive performance, whereas those with a higher variability had lower predictive performance (Fig 10A). This inverse relationship between VR and the VAF predictive performance was confirmed statistically (Fig 10A, r=-0.89, p=2.29e-54).

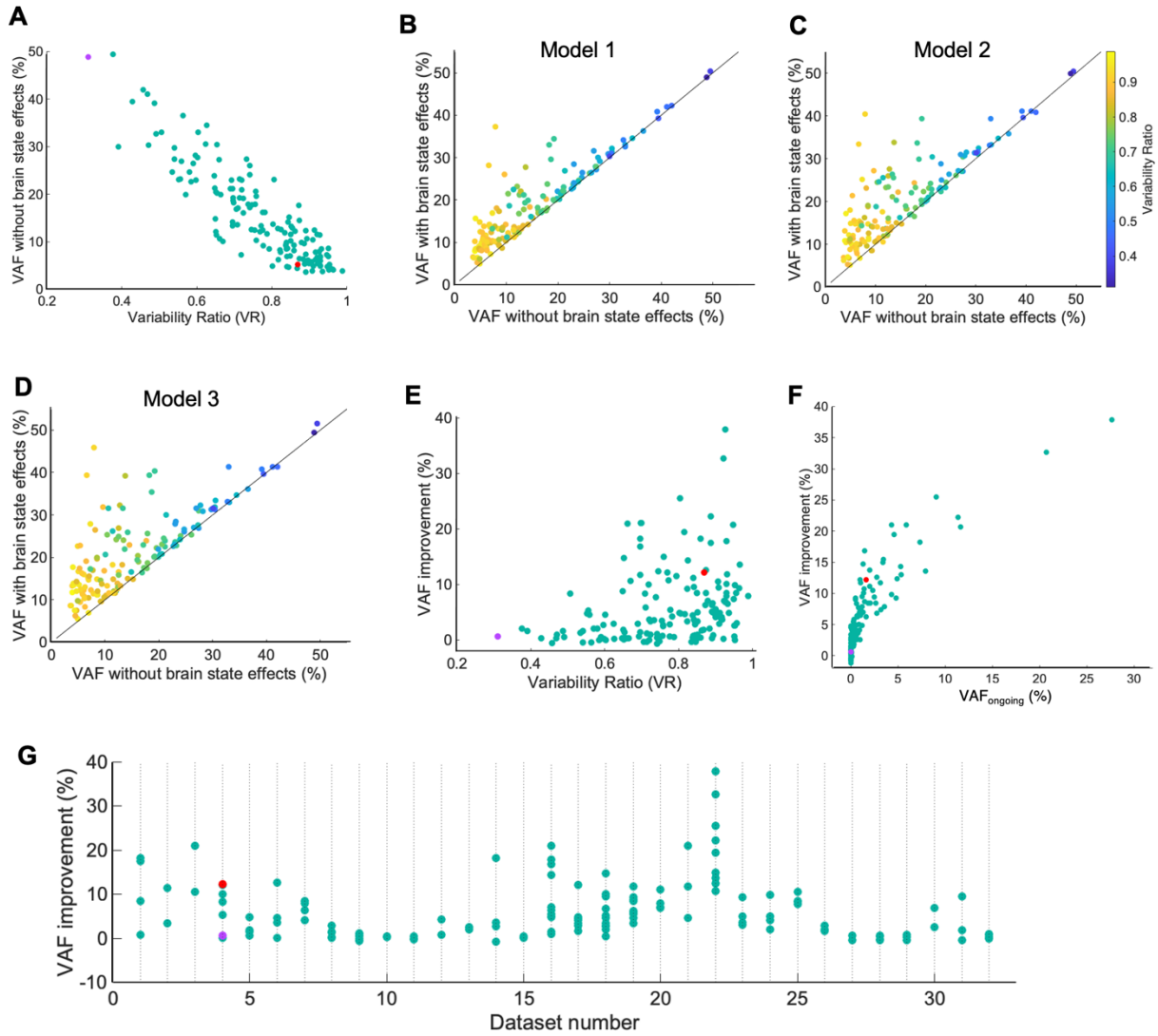

Fig 10: Improvements in predictive performance for the models with cortical state indicators, across the population of recorded neurons. A) Scatter plot for the population of neurons (n=154) showing the VAF from the stimulus-driven pathway against response variability ratio (*VR*). **B)** Scatter plot for the population of neurons showing VAF from model with LFP from the above channel, against VAF from only the stimulus-driven pathway. Data points for each neuron are color coded according to their variability ratio. **C)** Similar to B, with VAF from the model with LFP from all channels. **D)** Similar to B, but showing VAF from model with LFP from all channels and MUA from above and below channels. **E)** Improvement in VAF after incorporating cortical state effects in Model 3, plotted against response variability ratio for each neuron. **F)**

Improvement in predictive VAF plotted against the Model Accounted Variability (MAV) for each neuron. **G)** Distribution of VAF improvements of neurons across 32 datasets, each containing at least 2 simultaneously recorded neurons. Each data point on a column shows a simultaneously recorded neuron in that dataset.

Given the observed differential VAF improvements in two example neurons (Fig 9), we examined the improvements of VAF across the population of 154 neurons. We analyzed the gradual improvements in the VAF predictive accuracy when we incorporate LFP signals from a channel located 100μm away (Model 1, Fig 10B), LFP from all the channels (Model 2, Fig 10C), LFP from all the channels and also MUA signals from two nearby channels (Model 3, Fig 10D), compared to the base model (stimulus-driven pathway alone, abscissa of plots in Fig 10B,C,D). Each data point of these scatter plots represents single neurons, color-coded according to its variability ratio. The abscissa shows the VAF accuracy with stimulus-driven pathway alone (base model) and ordinate shows the VAF accuracy from three models incorporating cortical state effects (model 1-3). These neurons showed statistically significantly higher VAFs compared to the base model using LFP from a nearby channel (Fig 10B; VAF from Model 1 compared to base model: $p=9.53e-05$, $n=154$), LFP from all the channels (Fig 10C; VAF from Model 2 compared to base model $p=8.34e-07$, $n=154$) and with the complete model including both LFP and MUA parallel pathways (Fig 10D; VAF from Model 3 compared to base model $p=9.95e-09$, $n=154$).

These improvements were higher for neurons with higher response variability, but relatively minor for neurons with a lower response variability (Fig 10E: $r=0.49$,

p=1.52e-10). Overall, we observed improvements of the predictive performance (VAF) in the population of neurons improving from 0% to 40% (Fig 10E). Given the range of observed VAF improvements, we explored if the VAF improvements and the amount of variability explained by our model are related. For this, we calculated the measure "$VAF_{ongoing}$" where we compare each neurons' variable part of the response to the additional response predicted by model 3 compared to the base model (see methods). Higher $VAF_{ongoing}$ shows that the additional responses predicted by the Model 3 compared to the base model are accounting for most of the variable part of the neuronal response. The observed overall VAF improvements correlated well with the $VAF_{ongoing}$, i.e., the neuronal response variability was predicted well by our cortical state-driven pathway (Fig 10F: r=0.85, p= 1.65e-44). Thus, these results show the capability of the model to account for the observed trial-to-trial response variability by incorporating cortical state dynamics along with the stimulus-driven pathway.

Simultaneously recorded neurons in the population showed quite different amounts of response variability (Fig 4E). Therefore, we explored whether these simultaneously recorded neurons also show different amounts of VAF improvements after accounting for the cortical state effects using LFP and MUA signals. Across the sets of simultaneously recorded neurons, we observed different amounts of VAF improvements ranging from a maximum difference of 27.2% to a minimum of 0.07% with a median of 4.39%.

Given the observed apparent improvements in the receptive fields of the second example neuron (Fig 9B), we examined whether this improvement is observed throughout our population of neurons. To quantify the apparent qualitative improvements in estimated receptive fields for the population of neurons, we calculated z-score values for the receptive fields at the peak time lag (i.e., -26.7ms; see Methods). The z-score values after incorporating cortical state dynamics (median 1.08, inter-quartile range 0.86) were slightly larger than those from the base model with the stimulus pathway alone (median 0.91, inter-quartile range 0.72), and the difference was statistically significant (Fig 11A, p=1.38e-02, n=154). Fig 11B plots the z-scores for the two models against one another. Among the 154 neurons, 100 neurons (64.94%) have higher z-scores for the full model (dots above the diagonal line), while 54 neurons (35.06%) have lower z-scores for the full model (dots below the diagonal line). However, Fig 11C shows that the improvements in the receptive fields were not statistically correlated with the observed VAF improvements (r=0.13, p=9.6e-02).

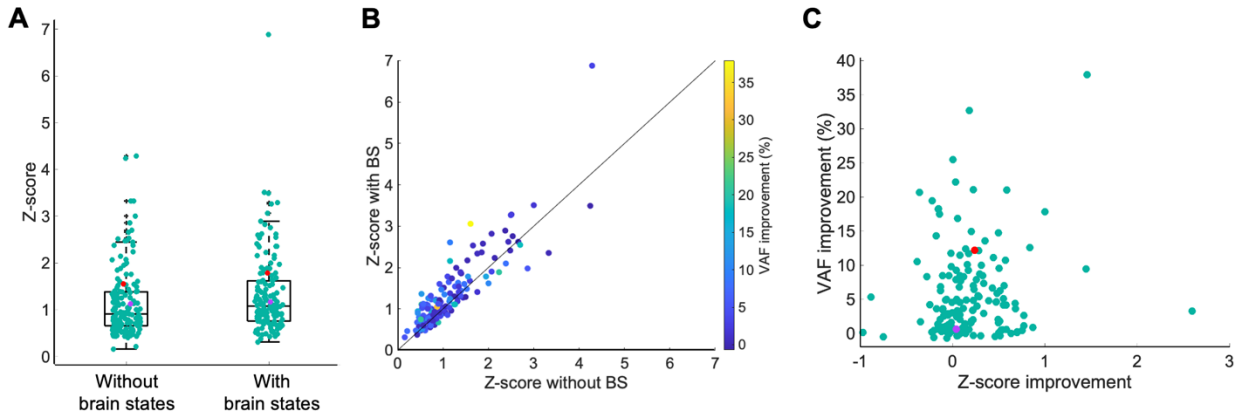

Fig 11: Receptive field improvements with cortical state across the sample population of neurons. A) Box plot of z-scores for the receptive fields estimated

without and with cortical state-driven pathway. **B)** Similar to A, shown as a scatter plot. C) Scatter plot of z-score improvement and VAF improvement.

Discussion

A major percentage of responses recorded from primary visual cortex neurons cannot be explained as due to visual stimuli alone, due to trial-to-trial response variability. The present study evaluated the hypothesis that the observed response variability arises from cortical state fluctuations. To test this idea, we proposed a compact convolutional neural network model of each neuron's visual receptive field, that also incorporated temporally filtered LFP and MUA signals as measures of cortical state dynamics (Fig. 3). The fitted model showed significantly improved accuracies in predicting neuronal responses and in estimating receptive fields, compared to a model with a stimulus-driven pathway alone (Fig 10B-D, Fig 11). The improvements were greater for neurons with higher trial-to-trial response variability than for those showing more reliable responses across trials (Fig 10E-F). These results support the idea that LFP and MUA carry information about response variability caused by the fluctuations in cortical states (Cui et al., 2016; Kelly et al., 2010).

Furthermore, neurons recorded simultaneously could evidently be affected differentially by cortical state variations. Some of the simultaneously recorded neurons showed substantially higher response variability across trials and higher predictive accuracy improvements with incorporated cortical state effects in the model, while other neurons produced relatively reliable responses across trials and showed lower predictive accuracy improvements (Fig. 4E, 10G). In addition, the estimated temporal

filters for LFP and MUA were neuron-specific and filter different frequencies of the LFP signal, demonstrating an important factor in how different neurons are affected differentially by cortical state dynamics (Fig. 7, S6, S9 Figs).

Inference of cortical state fluctuations

Our study in evaluating the effects of cortical states on response variability is based on the prevalence of ongoing cortical state dynamics within a data recording session. We confirm this by deriving a global fluctuation index (GFI, adapted from (Schölvinck et al., 2015)) from binned MUA during the blank inter-sweep periods, and synchrony index (SI, adapted from (Spacek & Swindale, 2016) from LFP across each data recording session. These fluctuations showed a similar pattern when compared at the same time points (Fig 4, red dots in panel E vs. panel C data points, S8 Fig). Thus, even though the animals are under anesthesia, their cortical state fluctuated continuously, consistent with the idea of states varying in a continuum (Harris & Thiele, 2011).

The estimates of cortical state could affect the performance of our proposed model. We used simultaneously recorded MUA and LFP brain signals to quantify cortical state dynamics. Since MUA and LFP signals contain information in the relatively high and low frequency ranges, respectively, we expected that they might carry complementary or independent information about cortical states. Cortical state characteristics have often been regarded as slow fluctuations on a time scale of seconds or more (Harris & Thiele, 2011) - these characteristics should be reflected in low frequency LFP signals. However, our estimated temporal filters for both LFP and MUA contain an optimal time window in the millisecond scale (Fig. 7). Further,

preliminary analysis showed that increasing the temporal window length by including more time points or including an additional parallel pathway for ultra-low LFP signals (<0.5Hz) don't bring additional improvements in predictive accuracies. Therefore, our millisecond scale time window may be primarily capturing rapid up-down phases (Harris & Thiele, 2011; Munro et al., 2015) embedded within the slow fluctuations. Further, these dynamics could also reflect millisecond-scale resolution of brain state embeddings including substates and rapid events shown in recent studies (Parks et al., 2023). In spite of similar time scales of the estimated filters, incorporating the MUA pathway does yield improved predictive performance, suggesting that it provides some additional higher temporal resolution information related to cortical states.

Variation in LFP frequencies of importance

Previous studies showed that different LFP frequency bands of EEG and LFPs correspond to distinct behavioral states (Saby & Marshall, 2012). For example, low frequency delta signals are associated with unconsciousness or deep-dreamless sleep, and high frequency gamma signals are associated with motor functions and higher mental activity (Saby & Marshall, 2012). Across our sample of neurons, we observed a wide diversity in the estimated LFP filters, and consequently of the LFP frequencies effective for the cortical state-driven pathway (Fig 7A-D, S6, S9 Figs). Thus, cortical state dynamics can affect individual neurons differentially. The filtered LFP frequency bands from the system identification model aligned well and showed substantial similarities to values of phase locking strength computed using LFP signals and spiking response for the corresponding neuron (i.e. similar frequency ranges in Fig 2B, C vs Fig

7C, D and in S9 Fig B vs C). This validated the capability of our model in estimating temporal filters to extract the significant LFP frequency bands when the whole LFP signal was given as an input to the model. Therefore, our approach works flexibly without requiring the use of LFP signals segmented into specific frequency bands.

Previous studies in area MT of awake behaving macaque monkeys showed strong links of spiking activity to specific frequency bands of LFP including high frequency gamma (30-70Hz) and low frequency delta (1-4Hz) bands (Cui et al., 2016). This study used wavelet transformation of LFPs to 16 logarithmically spaced scales from 0.5-70Hz, finding that gamma and delta frequencies were primarily relevant. This difference from our study, in which we saw a range of relevant frequencies, could be due to various factors, including the difference of species and brain areas, and our approach of estimating filters applied to raw LFPs rather than wavelet coefficients.

Cortical state-driven pathway

Our model architecture for the cortical state pathway is with a simple pair of linear filters that are unique for individual neurons and are capable of inferring cortical states from LFP and MUA. We evaluated several alternative approaches including filter-rectify-filter (FRF) cascades to capture envelopes of high frequency LFPs, using dimensionality-reduced LFP signals, and changing the brain state modulation to act at intermediate stages along the visual processing pathway. However, the predictive accuracy of each of these was comparatively lower. Further, we measured LFPs and MUA only for sites along our linear array, but signals from other columns and brain areas might carry relevant and distinct cortical state-related information. Including these additional signals

in parallel pathways in a model like ours might further improve the prediction accuracies.

Other studies have used measures of cortical states derived from other brain signals including scalp EEG (Hobson & Pace-Schott, 2002) and hippocampal LFPs (Pedrosa et al., 2022). Cortical state can also be reflected in behavioural signals such as whisker movements (Poulet & Petersen, 2008), locomotion velocity and pupil diameter (McGinley et al., 2015). Including such signals in parallel cortical state-driven pathways could enrich model's representation of brain state fluctuations. The extent to which these more elaborate models could boost the neural response predictive accuracy is uncertain since many of these brain state-indicative signals may correlate with one another (Larsen & Waters, 2018). Nevertheless, it would be interesting to examine whether using these behavioural signals in additional parallel pathways in a model like ours could further improve the predictive accuracies.

Spike contamination in LFP signals

Extracellular signals arise as a combination of many sources including synaptic currents, action potentials, calcium spikes, etc. and we record these biophysical processes in terms of brain signals such as EEG and LFPs depending on the recording method (Buzsáki et al., 2012). Consequently, LFPs recorded from an electrode could be contaminated by spikes from nearby neurons. Therefore, if we use the LFP from a given channel to predict the spiking activity of a neuron recorded from the same channel, that LFP signal could contain a significant contribution from the spike itself and yield misleading predictive accuracy.

While previous work showed that although spikes can be removed from LFP using a complex spike removal algorithm (Zanos et al., 2011), residual spike contamination effects may still remain (Pesaran et al., 2018). Hence one remedy is to use spikes and LFPs from different recording sites separated by several hundred microns (Pesaran et al., 2018). Consistent with this, (Cui et al., 2016) found similar system identification model weights by using spike-removed LFPs versus LFPs 100μm away from the spike recording electrode. Consequently, we excluded LFPs from the same channel as that of the recorded neuron, to mitigate effects of residual spike signals from that neuron.

Complexity in accounting for response variability

Our results demonstrated substantial trial-to-trial response variability in many of the neurons recorded from the primary visual cortex (Fig. 4). By incorporating cortical state effects in the system identification model, the predictive accuracy (VAF) was greatly improved (Fig. 10A-D), suggesting that fluctuations in cortical state contributed to the trial-to-trial variability. However, the present model did not completely account for the observed response variability (Fig. 10F). Even with the incorporation of brain state-related signals in our model, the predictive performance remains well below 100%. An elementary reason for predictive error is perhaps in part due to intrinsic neuronal noise, including cellular, electrical, and synaptic noise that arise at different stages of sensory processing (Faisal et al., 2008). The effects of the intrinsic neuronal noise sources are challenging to model due to their complexity. In addition, variability can arise due to stimulus-dependent factors including stimulus onset (Churchland et al., 2010), stimulus features including orientation and contrast (Kohn & Smith, 2005) and adaptation from

previous visual stimuli (Snow et al., 2017). More elaborate model architectures, and/or additional brain state indicators, might be required to account for these effects.

Previous studies have found that neurons' response variability increase across the visual hierarchy from retina to LGN to V1 and other higher order cortical areas, and is also inversely correlated with the average firing rates across these three brain regions (Goris et al., 2014; Kara et al., 2000). Consistent with this, we did observe a negative correlation between the firing rates and response variability of neurons, which was relatively weak though statistically significant (Fig 4C). The observed weak correlation might be because our population contains only areas 17 and 18 neurons, but not recordings across the visual hierarchy. However, if this correlation to be higher, adding a firing rate dependency into the model architecture would likely to be useful.

Stimulus-driven pathway

The stages in our stimulus-driven pathway is easily interpretable and applicable for both simple and complex cells in the primary visual cortex (Nguyen et al., 2024). However, important improvements in predictive performance might be gained by developing more elaborate model architectures to handle other known aspects of visual processing, particularly nonlinearities such as gain control (Ohzawa et al., 1985), surround suppression, adaptation, or second-order processing (Li et al., 2014). Some studies have reported that deep neural networks, e.g. with more filters and/or filter layers, can provide improved predictive power for cortical neurons' visual responses (Cadena et al., 2019; Ukita et al., 2019; Zhang et al., 2019). While such approaches might be desirable and interesting, they would entail greater challenges in optimizing the additional model

parameters while avoiding overfitting and would entail a substantial cost in interpretability.

Previous studies reported that the structures of V1 receptive fields could substantially change when measured in different contexts including general cortical states (Wörgötter et al., 1998). These include receptive fields becoming wider in synchronized cortical states and narrower during desynchronized states. Preliminary analysis with a subset of neurons revealed that for some neurons the receptive field size increased with synchronized states, consistent with the (Wörgötter et al., 1998) report, while others exhibited an opposite result - in any case, these effects were relatively small (S10 Fig). Therefore, in our system identification approach we treated the receptive fields as having a common structure and size throughout a given data acquisition session, regardless of cortical state variations. If this were to be shown to be a more serious effect, then this limitation in our method could conceivably be addressed in future approaches by using a model architecture that incorportates cortical state effects inside the stimulus-driven pathway.

Improvements in receptive field estimation

Incorporating brain state fluctuations often led to improvements in predictions of neurons' responses, and also to better estimates of RFs in terms of z-scores – however, across the sample population these two kinds of improvement were not statistically correlated (Fig. 11). This discrepancy could be due to multiple factors. Our model is trained based on optimizing the accuracy in predicting neural responses. In doing so, the optimization process might focus less on fitting the most accurate RF, but attaining

the higher VAFs. Conversely, inadequacies of our visual pathway model, which would degrade the overall predictive accuracy, could vary substantially across neurons, in a manner not necessarily related to the z-scores. Thus, a poor visual response model applied to a neuron with little brain state dependence, might give a "clean" (high z-score) RF estimate which is quite biased, resulting in poor predictive performance.

In addition, our z-score measure might not be an ideal measure of the quality of an estimated RF. For example, collapsing the spatial z-score values into a single number (Eq 15) in deriving our z-score measure would not be completely adequate for some neurons. As a result, comparing quantitative improvements in VAF to qualitative improvements in RFs in terms of z-score might lead to the observed discrepancies.

Conclusions

Our results demonstrated varying amounts of response variability, and of improvement in predictive performance, for different neurons - most interestingly, even for simultaneously recorded neurons. This heterogeneity of brain state effects across neurons places an important caveat on interpretations of global or large-scale (population) measures of brain state, such as scalp-recorded EEG. It would be interesting to further explore this heterogeneity, by examining whether there are systematic differences between neurons affected differentially by brain state, e.g. receptive field properties or laminar location.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., & Devin, M. (2016). Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467*.

Arieli, A., Sterkin, A., Grinvald, A., & Aertsen, A. (1996). Dynamics of Ongoing Activity: Explanation of the Large Variability in Evoked Cortical Responses. *Science*, *273*(5283), 1868-1871. https://doi.org/doi:10.1126/science.273.5283.1868

Armstrong, K. M., & Moore, T. (2007). Rapid enhancement of visual cortical response discriminability by microstimulation of the frontal eye field. *Proceedings of the National Academy of Sciences*, *104*(22), 9499-9504. https://doi.org/doi:10.1073/pnas.0701104104

Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nature Reviews Neuroscience*, *7*(5), 358-366. https://doi.org/10.1038/nrn1888

Banerjee, A., Dean, H., & Pesaran, B. (2012). Parametric models to relate spike train and LFP dynamics with neural information processing [Original Research]. *Frontiers in Computational Neuroscience*, *6*. https://doi.org/10.3389/fncom.2012.00051

Berger, H. (1929). Über das Elektrenkephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten*, *87*(1), 527-570. https://doi.org/10.1007/BF01797193

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*, 10008. https://doi.org/10.1088/1742-5468/2008/10/p10008

Borst, A., & Theunissen, F. E. (1999). Information theory and neural coding. *Nature Neuroscience*, *2*(11), 947-957. https://doi.org/10.1038/14731

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat Vis*, *10*(4), 433-436. https://www.ncbi.nlm.nih.gov/pubmed/9176952

Brown, R. E., Basheer, R., McKenna, J. T., Strecker, R. E., & McCarley, R. W. (2012). Control of sleep and wakefulness. *Physiological reviews*, *92*(3), 1087-1187. https://doi.org/10.1152/physrev.00032.2011

Buzsáki, G., Anastassiou, C. A., & Koch, C. (2012). The origin of extracellular fields and currents — EEG, ECoG, LFP and spikes. *Nature Reviews Neuroscience*, *13*(6), 407-420. https://doi.org/10.1038/nrn3241

Buzsáki, G., & Draguhn, A. (2004). Neuronal Oscillations in Cortical Networks. *Science*, *304*(5679), 1926-1929. https://doi.org/doi:10.1126/science.1099745

Cadena, S. A., Denfield, G. H., Walker, E. Y., Gatys, L. A., Tolias, A. S., Bethge, M., & Ecker, A. S. (2019). Deep convolutional models improve predictions of macaque V1 responses to natural images. *PLOS Computational Biology*, *15*(4), e1006897. https://doi.org/10.1371/journal.pcbi.1006897

Carandini, M. (2004). Amplification of Trial-to-Trial Response Variability by Neurons in Visual Cortex. *PLOS Biology*, *2*(9), e264. https://doi.org/10.1371/journal.pbio.0020264

Carandini, M., & Heeger, D. J. (2011). Normalization as a canonical neural computation. *Nat Rev Neurosci*, *13*(1), 51-62. https://doi.org/10.1038/nrn3136

Chollet, F., & others. (2018). *Keras: The Python Deep Learning library*. https://ui.adsabs.harvard.edu/abs/2018ascl.soft06022C

Churchland, M. M., Yu, B. M., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., Newsome, W. T., Clark, A. M., Hosseini, P., Scott, B. B., Bradley, D. C., Smith, M. A., Kohn, A., Movshon, J. A., Armstrong, K. M., Moore, T., Chang, S. W., Snyder, L. H., Lisberger, S. G., . . . Shenoy, K. V. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nature Neuroscience*, *13*(3), 369-378. https://doi.org/10.1038/nn.2501

Cohen, M. X. (2020). A data-driven method to identify frequency boundaries in multichannel electrophysiology data. *bioRxiv*, 2020.2007.2009.195784. https://doi.org/10.1101/2020.07.09.195784

Constantinople, Christine M., & Bruno, Randy M. (2011). Effects and Mechanisms of Wakefulness on Local Cortical Networks. *Neuron*, *69*(6), 1061-1068. https://doi.org/10.1016/j.neuron.2011.02.040

Cui, Y., Liu, L. D., McFarland, J. M., Pack, C. C., & Butts, D. A. (2016). Inferring Cortical Variability from Local Field Potentials. *J Neurosci*, *36*(14), 4121-4135. https://doi.org/10.1523/jneurosci.2502-15.2016

DeAngelis, G. C., Ghose, G. M., Ohzawa, I., & Freeman, R. D. (1999). Functional micro-organization of primary visual cortex: receptive field analysis of nearby neurons. *J Neurosci*, *19*(10), 4046-4064. https://www.ncbi.nlm.nih.gov/pubmed/10234033

Destexhe, A., Hughes, S. W., Rudolph, M., & Crunelli, V. (2007). Are corticothalamic 'up' states fragments of wakefulness? *Trends Neurosci*, *30*(7), 334-342. https://doi.org/10.1016/j.tins.2007.04.006

Ecker, A. S., Berens, P., Cotton, R. J., Subramaniyan, M., Denfield, G. H., Cadwell, C. R., Smirnakis, S. M., Bethge, M., & Tolias, A. S. (2014). State dependence of noise correlations in macaque primary visual cortex. *Neuron*, *82*(1), 235-248. https://doi.org/10.1016/j.neuron.2014.02.006

Gharat, A., & Baker, C. L., Jr. (2017). Nonlinear Y-Like Receptive Fields in the Early Visual Cortex: An Intermediate Stage for Building Cue-Invariant Receptive Fields from Subcortical Y Cells. *J Neurosci*, *37*(4), 998-1013. https://doi.org/10.1523/jneurosci.2120-16.2016

Goris, R. L. T., Movshon, J. A., & Simoncelli, E. P. (2014). Partitioning neuronal variability. *Nature Neuroscience*, *17*(6), 858-865. https://doi.org/10.1038/nn.3711

Harris, K. D., & Thiele, A. (2011). Cortical state and attention. *Nature Reviews Neuroscience*, *12*(9), 509-523. https://doi.org/10.1038/nrn3084

Hobson, J. A., & Pace-Schott, E. F. (2002). The cognitive neuroscience of sleep: neuronal systems, consciousness and learning. *Nature Reviews Neuroscience*, *3*(9), 679-693. https://doi.org/10.1038/nrn915

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol*, *160*, 106-154. https://doi.org/10.1113/jphysiol.1962.sp006837

Kara, P., Reinagel, P., & Reid, R. C. (2000). Low response variability in simultaneously recorded retinal, thalamic, and cortical neurons. *Neuron*, *27*(3), 635-646. https://doi.org/10.1016/s0896-6273(00)00072-6

Kelly, R. C., Smith, M. A., Kass, R. E., & Lee, T. S. (2010). Local field potentials indicate network state and account for neuronal response variability. *Journal of Computational Neuroscience*, *29*(3), 567-579. https://doi.org/10.1007/s10827-009-0208-9

Kleiner, M., Brainard, D. H., & Pelli, D. G. (2007). "What's new in Psychtoolbox-3?" *Perception 36*, 1-16.

Kohn, A., & Smith, M. A. (2005). Stimulus Dependence of Neuronal Correlation in Primary Visual Cortex of the Macaque. *The Journal of Neuroscience*, *25*(14), 3661-3673. https://doi.org/10.1523/jneurosci.5106-04.2005

Lachaux, J. P., Rodriguez, E., Martinerie, J., & Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Hum Brain Mapp*, *8*(4), 194-208. https://doi.org/10.1002/(sici)1097-0193(1999)8:4<194::aid-hbm4>3.0.co;2-c

Larsen, R. S., & Waters, J. (2018). Neuromodulatory Correlates of Pupil Dilation. *Front Neural Circuits*, *12*, 21. https://doi.org/10.3389/fncir.2018.00021

Lesica, N. A., Jin, J., Weng, C., Yeh, C. I., Butts, D. A., Stanley, G. B., & Alonso, J. M. (2007). Adaptation to stimulus contrast and correlations during natural visual stimulation. *Neuron*, *55*(3), 479-491. https://doi.org/10.1016/j.neuron.2007.07.013

Li, G., Yao, Z., Wang, Z., Yuan, N., Talebi, V., Tan, J., Wang, Y., Zhou, Y., & Baker, C. L., Jr. (2014). Form-cue invariant second-order neuronal responses to contrast modulation in primate area V2. *J Neurosci*, *34*(36), 12081-12092. https://doi.org/10.1523/jneurosci.0211-14.2014

Lin, I. C., Okun, M., Carandini, M., & Harris, K. D. (2015). The Nature of Shared Cortical Variability. *Neuron*, *87*(3), 644-656. https://doi.org/10.1016/j.neuron.2015.06.035

Liu, J., Dong, K., Sun, Y., Kakkos, I., Huang, F., Wang, G., Qi, P., Chen, X., Zhang, D., Bezerianos, A., & Sun, Y. (2021). Progress of Brain Network Studies on Anesthesia and Consciousness: Framework and Clinical Applications. *Engineering*. https://doi.org/https://doi.org/10.1016/j.eng.2021.11.013

Mainen, Z. F., & Sejnowski, T. J. (1995). Reliability of Spike Timing in Neocortical Neurons. *Science*, *268*(5216), 1503-1506. https://doi.org/10.1126/science.7770778

McGinley, M. J., Vinck, M., Reimer, J., Batista-Brito, R., Zagha, E., Cadwell, C. R., Tolias, A. S., Cardin, J. A., & McCormick, D. A. (2015). Waking state: rapid variations modulate neural and behavioral responses. *Neuron*, *87*(6), 1143-1161.

McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv:1802.03426. Retrieved February 01, 2018, from https://ui.adsabs.harvard.edu/abs/2018arXiv180203426M

Mendoza-Halliday, D., Torres, S., & Martinez-Trujillo, J. C. (2014). Sharp emergence of feature-selective sustained activity along the dorsal visual pathway. *Nature Neuroscience*, *17*(9), 1255-1262. https://doi.org/10.1038/nn.3785

Munro, E., Khodai, T., Sakata, S., & Toyoizumi, T. (2015). Two different mechanisms alternate during cortical synchronized states. *BMC Neuroscience*, *16*(1), P264. https://doi.org/10.1186/1471-2202-16-S1-P264

Murray, S. O. (2008). The effects of spatial attention in early human visual cortex are stimulus independent [Article]. *J. Vis.*, *8*(2), 1-11. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85069218299&partnerID=40&md5=b988515d8d8273882d7744cd3edb1d75

Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S., & Churchland, A. K. (2019). Single-trial neural dynamics are dominated by richly varied movements. *Nature Neuroscience*, *22*(10), 1677-1686. https://doi.org/10.1038/s41593-019-0502-4

Nguyen, P., Sooriyaarachchi, J., Huang, Q., & Baker, C. L., Jr. (2024). Estimating receptive fields of simple and complex cells in early visual cortex: A convolutional neural network model with parameterized rectification. *PLOS Computational Biology*, *20*(5), e1012127. https://doi.org/10.1371/journal.pcbi.1012127

Ohzawa, I., Sclar, G., & Freeman, R. D. (1985). Contrast gain control in the cat's visual system. *J Neurophysiol*, *54*(3), 651-667. https://doi.org/10.1152/jn.1985.54.3.651

Okun, M., Steinmetz, N. A., Cossell, L., Iacaruso, M. F., Ko, H., Barthó, P., Moore, T., Hofer, S. B., Mrsic-Flogel, T. D., Carandini, M., & Harris, K. D. (2015). Diverse coupling of neurons to populations in sensory cortex. *Nature*, *521*(7553), 511-515. https://doi.org/10.1038/nature14273

Olmos, A., & Kingdom, F. A. (2004). A biologically inspired algorithm for the recovery of shading and reflectance images. *Perception*, *33*(12), 1463-1473. https://doi.org/10.1068/p5321

Pachitariu, M., Steinmetz, N., Kadir, S., Carandini, M., & Kenneth D., H. (2016). Kilosort: realtime spike-sorting for extracellular electrophysiology with hundreds of channels. *bioRxiv*, 061481. https://doi.org/10.1101/061481

Parks, D. F., Schneider, A. M., Xu, Y., Brunwasser, S. J., Funderburk, S., Thurber, D., Blanche, T., Dyer, E. L., Haussler, D., & Hengen, K. B. (2023). A non-oscillatory, millisecond-scale embedding of brain state provides insight into behavior. *bioRxiv*. https://doi.org/10.1101/2023.06.09.544399

Pedrosa, R., Nazari, M., Mohajerani, M. H., Knöpfel, T., Stella, F., & Battaglia, F. P. (2022). Hippocampal gamma and sharp wave/ripples mediate bidirectional interactions with cortical networks during sleep. *Proceedings of the National Academy of Sciences*, *119*(44), e2204959119. https://doi.org/doi:10.1073/pnas.2204959119

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis*, *10*(4), 437-442. https://www.ncbi.nlm.nih.gov/pubmed/9176953

Pesaran, B., Vinck, M., Einevoll, G. T., Sirota, A., Fries, P., Siegel, M., Truccolo, W., Schroeder, C. E., & Srinivasan, R. (2018). Investigating large-scale brain dynamics using field potential recordings: analysis and interpretation. *Nat Neurosci*, *21*(7), 903-919. https://doi.org/10.1038/s41593-018-0171-8

Pettersen, K. H., Devor, A., Ulbert, I., Dale, A. M., & Einevoll, G. T. (2006). Current-source density estimation based on inversion of electrostatic forward solution: effects of finite extent of neuronal activity and conductivity discontinuities. *J Neurosci Methods*, *154*(1-2), 116-133. https://doi.org/10.1016/j.jneumeth.2005.12.005

Poulet, J. F. A., & Petersen, C. C. H. (2008). Internal brain state regulates membrane potential synchrony in barrel cortex of behaving mice. *Nature*, *454*(7206), 881-885. https://doi.org/10.1038/nature07150

Reimer, J., Froudarakis, E., Cadwell, C. R., Yatsenko, D., Denfield, G. H., & Tolias, A. S. (2014). Pupil fluctuations track fast switching of cortical states during quiet wakefulness. *Neuron*, *84*(2), 355-362. https://doi.org/10.1016/j.neuron.2014.09.033

Roohi-Azizi, M., Azimi, L., Heysieattalab, S., & Aamidfar, M. (2017). Changes of the brain's bioelectrical activity in cognition, consciousness, and some mental disorders. *Med J Islam Repub Iran*, *31*, 53. https://doi.org/10.14196/mjiri.31.53

Saalmann, Y. B., Pigarev, I. N., & Vidyasagar, T. R. (2007). Neural Mechanisms of Visual Attention: How Top-Down Feedback Highlights Relevant Locations. *Science*, *316*(5831), 1612-1615. https://doi.org/doi:10.1126/science.1139140

Saby, J. N., & Marshall, P. J. (2012). The utility of EEG band power analysis in the study of infancy and early childhood. *Dev Neuropsychol*, *37*(3), 253-273. https://doi.org/10.1080/87565641.2011.614663

Saleem, A. B., Chadderton, P., Apergis-Schoute, J., Harris, K. D., & Schultz, S. R. (2010). Methods for predicting cortical UP and DOWN states from the phase of deep layer local field potentials. *J Comput Neurosci*, *29*(1-2), 49-62. https://doi.org/10.1007/s10827-010-0228-5

Schölvinck, M. L., Saleem, A. B., Benucci, A., Harris, K. D., & Carandini, M. (2015). Cortical State Determines Global Variability and Correlations in Visual Cortex. *The Journal of Neuroscience*, *35*(1), 170-178. https://doi.org/10.1523/jneurosci.4994-13.2015

Snow, M., Coen-Cagli, R., & Schwartz, O. (2017). Adaptation in the visual cortex: a case for probing neuronal populations with natural stimuli. *F1000Res*, *6*, 1246. https://doi.org/10.12688/f1000research.11154.1

Spacek, M. A., & Swindale, N. V. (2016). Cortical state and natural movie responses in cat visual cortex. *bioRxiv*, 031765. https://doi.org/10.1101/031765

St-Amand, D., & Baker, C. L., Jr. (2023). Model-Based Approach Shows ON Pathway Afferents Elicit a Transient Decrease of V1 Responses. *J Neurosci*, *43*(11), 1920-1932. https://doi.org/10.1523/jneurosci.1220-22.2023

Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C. B., Carandini, M., & Harris, K. D. (2019). Spontaneous behaviors drive multidimensional, brainwide activity. *Science*, *364*(6437), eaav7893. https://doi.org/doi:10.1126/science.aav7893

Talebi, V., & Baker, C. L., Jr. (2012). Natural versus synthetic stimuli for estimating receptive field models: a comparison of predictive robustness. *J Neurosci*, *32*(5), 1560-1576. https://doi.org/10.1523/JNEUROSCI.4661-12.2012

Thiele, A., Pooresmaeili, A., Delicato, L. S., Herrero, J. L., & Roelfsema, P. R. (2009). Additive Effects of Attention and Stimulus Contrast in Primary Visual Cortex. *Cerebral Cortex*, *19*(12), 2970-2981. https://doi.org/10.1093/cercor/bhp070

Tolhurst, D. J., Movshon, J. A., & Dean, A. F. (1983). The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Research*, *23*(8), 775-785. https://doi.org/https://doi.org/10.1016/0042-6989(83)90200-6

Tootell, R. B., Silverman, M. S., & De Valois, R. L. (1981). Spatial frequency columns in primary visual cortex. *Science*, *214*(4522), 813-815. https://doi.org/10.1126/science.7292014

Ukita, J., Yoshida, T., & Ohki, K. (2019). Characterisation of nonlinear receptive fields of visual neurons by convolutional neural network. *Scientific Reports*, *9*(1), 3791. https://doi.org/10.1038/s41598-019-40535-4

Vinck, M., Batista-Brito, R., Knoblich, U., & Cardin, J. A. (2015). Arousal and locomotion make distinct contributions to cortical activity patterns and visual encoding. *Neuron*, *86*(3), 740-754. https://doi.org/10.1016/j.neuron.2015.03.028

Vogels, R., Spileers, W., & Orban, G. A. (1989). The response variability of striate cortical neurons in the behaving monkey. *Experimental Brain Research*, *77*(2), 432-436. https://doi.org/10.1007/BF00275002

Ward-Flanagan, R., Lo, A. S., Clement, E. A., & Dickson, C. T. (2022). A Comparison of Brain-State Dynamics across Common Anesthetic Agents in Male Sprague-Dawley Rats. *International journal of molecular sciences*, *23*(7), 3608. https://doi.org/10.3390/ijms23073608

Wilson, C. (2008). Up and down states. *Scholarpedia J*, *3*(6), 1410. https://doi.org/10.4249/scholarpedia.1410

Wörgötter, F., Suder, K., Zhao, Y., Kerscher, N., Eysel, U. T., & Funke, K. (1998). State-dependent receptive-field restructuring in the visual cortex. *Nature*, *396*(6707), 165-168. https://doi.org/10.1038/24157

Zanos, T. P., Mineault, P. J., & Pack, C. C. (2011). Removal of spurious correlations between spikes and local field potentials. *J Neurophysiol*, *105*(1), 474-486. https://doi.org/10.1152/jn.00642.2010

Zhang, Y., Lee, T. S., Li, M., Liu, F., & Tang, S. (2019). Convolutional neural network models of V1 responses to complex patterns. *Journal of Computational Neuroscience*, *46*(1), 33-54. https://doi.org/10.1007/s10827-018-0687-7

Zohary, E., Shadlen, M. N., & Newsome, W. T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature*, *370*(6485), 140-143. https://doi.org/10.1038/370140a0