Neural Field Models for Latent State Inference: Application to Large-Scale Neuronal Recordings

Michael E. Rule\textsuperscript{1}, David Schnoerr\textsuperscript{2}, Matthias H. Hennig\textsuperscript{3}, Guido Sanguinetti\textsuperscript{3}

\textsuperscript{*}Corresponding Author (mrule7404@gmail.com); \textsuperscript{1}University of Cambridge, Department of Engineering; \textsuperscript{2}Imperial College London, Theoretical Systems Biology; \textsuperscript{3}University of Edinburgh, Department of Informatics

November 8, 2019

Abstract

Large-scale neural recording methods now allow us to observe large populations of identified single neurons simultaneously, opening a window into neural population dynamics in living organisms. However, distilling such large-scale recordings to build theories of emergent collective dynamics remains a fundamental statistical challenge. The neural field models of Wilson, Cowan, and colleagues remain the mainstay of mathematical population modeling owing to their interpretable, mechanistic parameters and amenability to mathematical analysis. Inspired by recent advances in biochemical modeling, we develop a method based on moment closure to interpret neural field models as latent state-space point-process models, making them amenable to statistical inference. With this approach we can infer the intrinsic states of neurons, such as active and refractory, solely from spiking activity in large populations. After validating this approach with synthetic data, we apply it to high-density recordings of spiking activity in the developing mouse retina. This confirms the essential role of a long lasting refractory state in shaping spatiotemporal properties of neonatal retinal waves. This conceptual and methodological advance opens up new theoretical connections between mathematical theory and point-process state-space models in neural data analysis.

Author Summary

Developing statistical tools to connect single-neuron activity to emergent collective dynamics is vital for building interpretable models of neural activity. Neural field models relate single-neuron activity to emergent collective dynamics in neural populations, but integrating them with data remains challenging. Recently, latent state-space models have emerged as a powerful tool for constructing phenomenological models of neural population activity. The advent of high-density multi-electrode array recordings now enables us to examine large-scale collective neural activity. We show that classical neural field approaches can yield latent state-space equations and demonstrate that this enables inference of the intrinsic states of neurons from recorded spike trains in large populations.

Introduction

Neurons communicate using electrical impulses, or spikes. Understanding the dynamics and physiology of collective spiking in large networks of neurons is a central challenge in modern neuroscience, with immense translational and clinical potential. Modern technologies such as high-density multi-electrode arrays (HDMEA) enable the simultaneous recording of the electrical activity of thousands of interconnected neurons, promising invaluable insights into neural dynamics at the network level. However, the resulting data is high-dimensional and frequently exhibits complex, non-linear dynamics, presenting formidable statistical challenges.

Due to the complexity of the data, most analyses of neuronal population activity take a descriptive approach, adopting methods from statistical signal processing such as state-space models (SSM; Paninski et al., 2010; Zhao and Park, 2017b,a; Sussillo et al., 2016; Aghagolzadeh and Truccolo, 2016; Linderman et al., 2016; Gao et al., 2016) or autoregressive generalized-linear point-process models (PP-GLM; Paninski,
2004; Pillow et al., 2008; Truccolo et al., 2005; Truccolo, 2016). Such methods capture the population statistics of the system, but fail to provide mechanistic explanations of the underlying neural dynamics. While this phenomenological description is valuable and can aid many investigations, the inability to relate microscopic single-neuron properties to emergent collective dynamics limits the scope of these models to extract biological insights from these large population recordings.

Connecting single-neuron dynamics with population behavior has been the central focus of research within the theoretical neuroscience community over the last four decades. Neural field models (Amari, 1977; Wilson et al., 1972; Cowan, 2014; Bressloff, 2012) have been crucial in understanding how macroscopic firing dynamics in populations of neurons emerge from the microscopic state of individual neurons. Such models have found diverse applications including working memory (see Durstewitz et al., 2000 for a review), epilepsy (e.g. Zhang and Xiao, 2018; Proix et al., 2018; González-Ramírez et al., 2015; Martinet et al., 2017), and hallucinations (e.g. Ermentrout and Cowan, 1979; Bressloff et al., 2001; Rule et al., 2011), and have been successfully related to neuroimaging data such as Electroencephalography (EEG; Moran et al., 2013; Bojak et al., 2010; Pinotsis et al., 2012), Magnetoencephalography (MEG; Moran et al., 2013), Electromyography (EMG; Nazarpour et al., 2012), and Functional Magnetic Resonance Imaging (fMRI; Bojak et al., 2010), which measure average signals from millions of neurons. Nevertheless, using neural-field models to model HDMEA spiking data directly remains an open statistical problem: HDMEA recordings provide sufficient detail to allow modeling of individual neurons, yet the large number of neurons present prevents the adoption of standard approaches to non-linear data assimilation such as likelihood free inference.

In this paper, we bridge the data-model divide by developing a statistical framework for Bayesian modeling in neural field models. We build on recent advances in stochastic spatiotemporal modeling, in particular a recent result by Schnoerr et al. (Schnoerr et al., 2016) which showed that a spatiotemporal agent-based model of reaction-diffusion type, similar to the ones underpinning many neural field models, can be approximated as a spatiotemporal point process associated with an intensity (i.e. density) field that evolves in time. Subsequently, Rule and Sanguineti (Rule and Sanguineti, 2018) illustrated a moment-closure approach for mapping stochastic models of neuronal spiking onto latent state-space models, preserving the essential coarse-timescale dynamics. Here, we demonstrate that a similar approach can yield state-space models for neural fields derived directly from a mechanistic microscopic description. This enables us to leverage large-scale spatiotemporal inference techniques (Cseke et al., 2016; Zammit-Mangion et al., 2012) to efficiently estimate an approximate likelihood, providing a measure of fit of the model to the data that can be exploited for data assimilation. Our approach is in spirit similar to latent variable models such as the Poisson Linear Dynamical System (PLDS; (Macke et al., 2011; Aghagolzadeh and Truccolo, 2016; Smith and Brown, 2003)), with the important difference that the latent variables reflects non-linear neural field dynamics that emerge directly from a stochastic description of single-neuron activity (Bressloff, 2009; Buice et al., 2010; Touboul and Ermentrout, 2011).

We apply this approach to HDMEA recordings of spontaneous activity from ganglion cells in the developing mouse retina (Maccione et al., 2014), showing that the calibrated model effectively captures the non-linear excitatory phenomenon of coordinated, wave-like patterns of spiking (Meister et al., 1991) that have been considered in both discrete (Hennig et al., 2009a) and continuous neural-field models before (Lansdell et al., 2014).

Results

High level description of the approach

We would like to explain large-scale spatiotemporal spiking activity in terms of the intrinsic states of the participating neurons, which we cannot observe directly. Latent state-space models (SSMs) solve this problem by describing how the unobserved states of neurons relate to spiking observations, and predict how these latent states evolve in time. In this framework, one estimates a distribution over latent states from observations, and uses a forward model to predict how this distribution evolves in time, refining the latent-state estimate with new observations as they become available. This process is often called ‘data assimilation’. However, in order to achieve statistical tractability, SSMs posit simple (typically linear) latent dynamics, which cannot be easily related to underlying neuronal mechanisms. Emergent large-scale spatiotemporal phenomena such as traveling waves typically involve multiple, coupled
populations of neurons and nonlinear excitatory dynamics, both of which are difficult to incorporate into conventional state-space models.

Fortunately, mathematical neuroscience has developed methods for describing such dynamics using neural field models. Neural field models map microscopic dynamics to coarse-grained descriptions of how population firing rates evolve. This provides an alternative route to constructing latent state-space models for large-scale spatiotemporal spiking datasets. However, neural field models traditionally do not model statistical uncertainty in the population states they describe, which makes it difficult to deploy them as statistical tools to infer the unobserved, latent states of the neuronal populations. A model of statistical uncertainty is important for describing the uncertainty in the estimated latent states (posterior variance), as well as correlations between states or spatial regions. As we will illustrate, work over the past decades to address noise and correlations in neural field models also provides the tools to employ such models as latent SSMs in data-driven inference.

At a high level then, our approach follows the usual derivation of neural field models, starting with an abstract description of single-neuron dynamics, and considers how population averages evolve in time. Rather than deriving a neural-field equation for the population mean rate, we instead derive two coupled equations for the mean and covariance of population states. We interpret these two moments as a Gaussian-process estimate of the latent spatiotemporal activity, and derive updates for how this distribution evolves in time and how it predicts spiking observations. This provides an interpretation of neural-field dynamics amenable to state-space inference, which allows us to infer neural population states from spiking observations.

Neural field models for refractoriness-mediated retinal waves

Although Wilson and Cowan (Wilson and Cowan, 1972, 1973) considered refractoriness, most subsequent applications consider only two states: neurons may be either actively spiking (A state), or quiescent (Q state). In general, voltage and calcium gated conductances typically lead to refractory states, which can be short following individual spikes, or longer after more intensive periods of activity. An excellent example of the importance of a refractory mechanism is found in the developing retina, where a slow afterhyperpolarization (sAHP) current mediates the long-timescale refractory effects that strongly shapes the spatiotemporal dynamics of spontaneous retinal waves (Hennig et al., 2009b). To address this, we explicitly incorporate additional refractory (R) states into our neural field model (e.g. (Buice and Cowan, 2007, 2009); Figure 1). In the following, we first outline a non-spatial model for such system, before extending it to a spatial setting with spatial couplings. Finally, we develop a Bayesian inference scheme for inferring latent states from observational data.

A stochastic three-state neural mass model

We now consider the neural field model with three states as a generic model of a spiking neuron (Figure 1), where a neuron can be in either an actively spiking (A), refractory (R), or quiescent (Q) state. We assume that the neurons can undergo the following four transitions:

\[
\begin{align*}
Q & \xrightarrow{\rho_Q} A & \quad & Q + A & \xrightarrow{\rho_A} A + A \\
A & \xrightarrow{\rho_A} R & \quad & R & \xrightarrow{\rho_R} Q,
\end{align*}
\]

i.e. quiescent neurons transition spontaneously to the active state; active neurons excite quiescent neurons; active neurons become refractory, and refractory neurons become quiescent. The \( \rho_{\cdot,\cdot} \) denote corresponding rate constants.

For illustration, we first consider the dynamics of a local (as opposed to spatially-extended) population of neurons. In this case the state of the system is given by the non-negative number counts \( Q, A \) and \( R \) of the respective neuron types (we slightly abuse notation here and use \( Q, A, \) and \( R \) both as symbols for the neuron states and as variables counting the neurons in the corresponding states; see Figure 2 for an illustration). The time evolution of the corresponding probability distribution to be in a state \( (Q, A, R) \) at a certain time point is then given by a master equation ((Buice and Cowan, 2007; Ohira and Cowan, 1993; Bressloff, 2009); Methods: Moment-closure for a single population). Due to the nonlinear excitatory interaction \( Q+A\rightarrow A+A \) in Eq. (1), no analytic solutions to the master equation are known. To
get an approximate description of the dynamics, we employ the Gaussian moment closure method that approximates the discrete neural counts \((Q, A, R)\) by continuous variables, and assumes a multivariate normal distribution (Figure 2B; Goodman, 1953; Whittle, 1957; Gomez-Uribe and Vergheese, 2007; Bressloff, 2009; Buice et al., 2010; Schnoerr et al., 2017; Rule and Sanguinetti, 2018). This allows one to derive a closed set of ordinary differential equations for the mean and covariance of the approximate process which can be solved efficiently numerically (Methods: Moment-closure for a single population; Figure 2).

Applying this procedure to our system leads to the following evolution equations for the first moments (mean concentrations):

\[
\begin{align*}
\dot{Q} &= r_q - rq_a \\
\dot{A} &= rq_a - ra_r \\
\dot{R} &= ra_r - rq_q
\end{align*}
\]  

(2)

where the rate variables \(\dot{\cdot}\) describe the rates of the different transitions in Eq. (1), and \(\langle \cdot \rangle\) denotes expected-value with respect to the distribution over population states. Intuitively, Eq. (2) says that the mean number of neurons in each state evolves according to the difference between the rate that neurons enter, and the rate that neurons leave, said state. For spontaneous (Poisson) state transitions, these rates are linear and depend only on the average number of neurons in the starting state. The transition from \(Q\) to \(A\), however, has both a spontaneous and excito-excitatory component. The latter depends on the expected product of active and quiescent cells \(\langle AQ\rangle\), which is a second moment and can be expressed in terms of the covariance: \(\langle AQ\rangle = \langle A \rangle \langle Q \rangle + \Sigma_{AQ}\). We obtain similar equations for the covariance of the system (Eq. 6; Methods: Moment-closure for a single population). These can be solved jointly with Eq. (2) forward in time to give an approximation of the system’s dynamics.

**Generalization to spatial (neural field) system**

So far we have considered a single local population. We next extend our model to a two-dimensional spatial system. In this case the mean concentrations become density or mean fields (‘neural fields’) that depend on spatial coordinates \(x = (x_1, x_2)\), e.g. \(\langle Q \rangle\) becomes \(Q(x)\). Similarly, the covariances become two-point correlation functions. For example, \(\Sigma_{Q,A}(x, x')\) denotes the covariance between the number of neurons in the quiescent state at location \(x\) and the number of neurons in the active state at location \(x'\) (see Methods: Extension to spatial system for details).

By replacing the mean concentrations and covariances accordingly in Eqs. (2) and (6), we obtain spatial evolution equations for these space-dependent quantities. The terms arising from the linear transitions in Eq. (1) (i.e. \(r_q, rq_a\) and the first term in \(r_q\) in Eq. 2) do not introduce any spatial coupling and hence do not need to be modified (note also that neurons do not diffuse or move otherwise, which is why we do not obtain a dynamic term in the resulting equations). The nonlinear excitatory interaction \(Q + A \rightarrow A + A\) in Eq. (1), however, introduces a coupling that we need to specify further in a spatial setting. We assume that each quiescent neuron experiences an excitatory drive from nearby active neurons, and that the interaction strength can be described as a function of distance \(||\Delta x||\) by a Gaussian interaction kernel:

\[
k(\Delta x) \propto \exp(-||\Delta x||^2/2\sigma_e^2),
\]

(3)

where \(\sigma_e\) the standard deviation determining the length scale of the interaction, which decays exponentially as a function of distance squared. This kernel introduces a spatial coupling between the neurons, which could be mediated by synaptic interactions, diffusing neurotransmitters, gap junction coupling, or combinations thereof. With this coupling, the transition rate (compare to Eq. (2)) from the quiescent to active state at position \(x\) becomes the following integral:

\[
r_{qa}(x) = \rho_q \langle Q(x) \rangle + \rho_e \int k(x-x') \langle Q(x)A(x') \rangle \, dx',
\]

(4)

where the integral runs over the whole volume of the system (Methods: Extension to spatial system).

We thus obtain a ‘second-order’ neural field in terms of the mean fields and two-point correlation functions. We simulated the spatially-extended system by sampling. Figure 3 shows that it is indeed capable of producing multi-scale wave-like phenomena similar to the waves observed in the retina (Methods: Sampling from the model).
Neural field models as latent-variable state-space models

The equations for the mean fields and correlations can be integrated forward in time and used as a state-space model to explain population spiking activity (Figure 4; Methods: Bayesian filtering). In extracellular recordings, we do not directly observe the intensity functions \( \langle Q(x) \rangle \), \( \langle A(x) \rangle \), and \( \langle R(x) \rangle \). Instead, we observe the spikes that active neurons emit, or in the case of developmental retinal waves recorded via a HDMEA setup, we observe the spikes of retinal ganglion cells which are driven by latent wave activity. The spiking intensity should hence depend on the density \( A(x) \) of active neurons. Here, we assume that neural firing is a Poisson process conditioned on the number of active neurons, which allows us to write the likelihood of point (i.e. spike) observations in terms of \( A(x) \) (Truccolo et al., 2005, 2010; Truccolo, 2016; Methods: Point-process measurement likelihood).

The combination of this Poisson-process observation model with the state-space model derived in previous sections describes how hidden neural field states evolve in time and how these states drive neuronal spiking. Given spatiotemporal spiking data, the latent neural field states and correlations can then be inferred using a sequential Bayesian filtering algorithm. The latter uses the neural field model to predict how latent states evolve, and updates this estimate at each time point based on the observed neuronal spiking (Methods: Bayesian filtering). This provides estimates of the unobserved physiological states of the neurons.

We verified that this approach works using simulated data. We first simulated observations from the neural field equations (Figure 3; Methods: Sampling from the model), which generated waves qualitatively similar to those seen in the developing retina. We then sampled spiking as a conditionally-Poisson process driven by the number of active neurons in each location, with a baseline rate of \( \beta = 0 \) and gain of \( \gamma = 15 \) spikes/second per simulation area. We then applied Bayesian filtering to these spiking samples in order to recover a Gaussian estimate of the latent neural field states (Methods: Bayesian filtering). Figure 5 illustrates the latent states recovered via filtering using the known ground-truth model parameters, and shows that filtering can recover latent neural field states from the spiking observations. Overall, this indicates that moment-closure of stochastic neural field equations can yield state-space models suitable for state inference from spiking data. In the next section, we illustrate this approach applied to waves recorded from the developing retina.

State inference in developmental retinal waves

Having developed an interpretation of neural field equations as a latent-variable state-space model, we next applied this model to the analysis of spatiotemporal spiking data from spontaneous traveling wave activity occurring in the neonatal vertebrate retina (e.g. Figure 6; Sernagor et al., 2003; Hennig et al., 2009a; Blankenship et al., 2009; Meister et al., 1991; Zhou and Zhao, 2000; Feller et al., 1996; Maccione et al., 2014).

During retinal development, the cell types that participate in wave generation change (Maccione et al., 2014; Sernagor et al., 2003; Zhou and Zhao, 2000), but the three-state model globally describes dynamics in the inner retina at all developmental stages (Figure 6). The Active (A) state describes a sustained bursting state, such as the depolarization characteristic of starburst amacrine cells (Figure 6a) during acetylcholine-mediated early-stage (Stage 2) waves between P0 and P9 (Feller et al., 1996; Zhou and Zhao, 2000), and late-stage (Stage 3) glutamate-dependent waves (Bansal et al., 2000; Zhou and Zhao, 2000). For example, Figure 6c-d illustrates spontaneous retinal wave activity recorded from a postnatal day 6 mouse pup (Stage 2). In addition, at least for cholinergic waves, the slow refractory state R is essential for restricting wave propagation into previously active areas (Zheng et al., 2006). We note that the multi-scale wave activity exhibited in the three-state neural field model (e.g. Figure 3) recapitulates the phenomenology of retinal wave activity explored in the discrete three-state model of Hennig et al. (Hennig et al., 2009b).

Using RGC spikes recorded with a 4,096-electrode HDMEA (Figure 6), we demonstrate the practicality of latent-state inference using heuristic rate parameters and illustrate an example of inference for a retinal wave dataset from postnatal day 11 (Stage 3; Figure 7). For retinal wave inference, we normalize the model by population-size (Methods: System-size scaling) so that the gain and bias do not depend on the local neuronal population size.

The state inference (‘data assimilation’) procedure uses new observations to correct for prediction errors. Because of this, many different model parameters may give similar state estimates. Nevertheless,
it is important that the rate parameters approximately match the data. The rate of excitation ($\rho_e$) should be fast, and the rate at which active cells become refractory ($\rho_r$) should match the typical wave duration. Likewise, it is important that the recovery rate $\rho_r$ matches the inter-wave interval timescale.

In Figure 7, model parameters were set based on observed timescales, and then adjusted such that the simulated model dynamics match those recovered during state inference ($\rho_e=10$, $\rho_r=1.8$, and $\sigma=0.1$). These parameters were held fixed during subsequent state inference. The interaction radius $\sigma=0.15$ and excitation strength $\rho_e$ interact to determine how excitable the system is and how quickly waves propagate. The overall excitability should be small enough so that the system is stable, and does not predict wave events in the absence of spiking observations. As in Lansdell et al. (Lansdell et al., 2014), lateral interactions in our model reflect an effective coupling that combines both excitatory synaptic interactions and the putative effect of diffusing excitatory neurotransmitters, which has been shown to promote late-stage glutamatergic wave propagation (Blankenship et al., 2009).

The moment-closure system does not accurately approximate the rare and abrupt nature of wave initiation. We therefore model spontaneous wave initiation as an extrinsic noise source, and set the spontaneous excitation rate $\rho_q$ to zero in the neural field model that defines our latent state-space. The Poisson noise was re-scaled to reflect an effective population size of $16 \text{ neurons/mm}^2$, significantly smaller than the true population density (Jeon et al., 1998). However, due to the recurrent architecture and correlated neuronal firing, the effective population size is expected to be smaller than the true population size. Equivalently, this amounts to assuming supra-Poisson scaling of fluctuations for the neural population responsible for retinal waves.

Bayesian filtering recovers the expected features of the retinal waves (Figure 7): the excitatory-excitatory transition $Q\rightarrow A \rightarrow A + A$ and the onset of refractoriness $A \rightarrow R$ are rapid compared to the slow refractory dynamics, and therefore the A state is briefly occupied and mediates an effective $Q \rightarrow R$ transition during wave events. The second-order structure provided by the covariance is essential, as it allows us to model posterior variance (shaded regions in Figure 7), while also capturing strong anti-correlations due to the conservation of reacting agents, and the effect of correlated fluctuations on the evolution of the means. Furthermore, spatial correlations allow localized RGC spiking events to be interpreted as evidence of regional (spatially-extended) latent neuronal activity.

**Open challenges in model identification**

So far, we have demonstrated good recovery of states when the true rate parameters are known (Figure 5), and shown that plausible latent states can be inferred from neural point-process datasets using a priori initialized parameters (Figure 7). A natural question then is whether one can use the Bayesian state-space framework to estimate a posterior likelihood on the rate parameter values, and infer model parameters directly from data. Presently, model inference remains challenging for four reasons: under-constrained parameters, computational complexity, numerical stability, and non-convexity in the joint posterior. It is worth reviewing these challenges as they relate to important open problems in machine learning and data assimilation.

First, the effective population size, the typical fraction of units in quiescent vs. refractory states, and the gain parameter mapping latent activations to spiking, are all important to setting appropriate rates, and are not accessible from observation of RGC spiking alone. Recovering a physiologically realistic model would require direct measurement or appropriate physiological priors on these parameters. In effect, many equivalent systems can explain the observed RGC spiking activity, a phenomenon that has been termed "sloppiness" in biological systems (Transtrum et al., 2015; Panas et al., 2015). Indeed, Hennig et al. (Hennig et al., 2011) show that developmental waves are robust to pharmacological perturbations, suggesting that the retina itself can use different configurations to achieve similar wave patterns. Second, although state inference is computationally feasible, parameter inference requires many thousands of state-inference evaluations. A Matlab implementation of state-inference running on a 2.9 GHz 8-core Xeon CPU can process ~85 samples/s for a 3-state system on a 10x10 spatial basis. For a thirty-minute recording of retinal wave activity, state inference is feasible, but repeated state inference for parameter inference is impractical. Third, model likelihood must be computed recursively, and is subject to loss of numerical accuracy due to back-propagation through time (Pascanu et al., 2013; Bengio et al., 1994; Hochreiter et al., 2001). In other words, small errors in the past can have large effects in the future owing to the nonlinear and excitable nature of the system. Fourth and finally, the overall likelihood
surface need not be convex, and may contain multiple local optima. Additionally, regions of parameters space can exhibit vanishing gradient for one or model parameters. This can occur when the value of one parameter makes others irrelevant. For example, if the excito-excitatory interaction \( \rho_e \) is set to a low value, the interaction radius \( \sigma_e \) for excitation becomes irrelevant since the overall excitation is negligible.

Overall, parameter inference via Bayesian filtering presents a formidable technical challenge. Presently, it seems that traditional methods, based on mathematical expertise and matching observable physical quantities (e.g. wavefront speed, c.f. (Lansdell et al., 2014)), remain the best-available approach to model estimation. Despite these challenges, the innovation presented here, of applying moment-closure methods for data assimilation, is important per se, because it provides a snapshot of the activity of unobserved states which can greatly aid scientific investigation. The state-space formulation of neural field models enables Bayesian state inference from candidate neural field models, and opens the possibility of likelihood-based parameter inference in the future.

**Discussion**

In this work, we showed that classical neural-field models, which capture the activity of large, interacting neural populations, can be interpreted as state-space models, where we can explicitly model microscopic, intrinsic dynamics of the neurons. This is achieved by interpreting a second-order neural field model as defining equations on the first two moments of a latent-variable process, which is coupled to spiking observations. In the state-space model interpretation, latent neural field states can be recovered from Bayesian filtering. This allows inferring the internal states of neuronal populations in large networks based solely on recorded spiking activity, information that can experimentally only be obtained with whole cell recordings.

We demonstrated successful state inference for simulated data, where the correct model and parameters were known. Next, we applied the model to large-scale recordings of developmental retinal waves. Here the correct latent-state model is unknown, but a relatively simple three-state model with slow refractoriness is well-motivated by experimental observations (Zheng et al., 2006). Previous works (Feller et al., 1997; Zheng et al., 2006; Godfrey and Swindale, 2007; Hennig et al., 2009a) predict that activity-dependent refractoriness is important for restricting the spatial spreading of waves. Intuitively, one should expect the refractory time constant to be a highly sensitive parameter: very long refractory constants will impede the formation of waves, while short constants might lead to interference phenomena. These intuitions were borne out empirically by our simulation studies; additionally, we observed that long refractory constants led to ineffective data assimilation, as the model prior is too dissimilar from the data it is trained upon. In contrast to phenomenological latent state-space models, the latent states here are motivated by an (albeit simplified) description of single-neuron dynamics, and the state-space equations arise directly from considering the evolution of collective activity as a stochastic process.

In the example explored here, we use Gaussian moment-closure to arrive at a second-order approximation of the distribution of latent states and their evolution. In principle, other distributional assumptions may also be used to close the moment expansion. Other mathematical approaches that yield second-order models could also be employed, for example the linear noise approximation (Van Kampen, 1992), or defining a second cumulant in terms of the departure of the model from Poisson statistics (Buice et al., 2010). The approach applied here to a three-state system can generally be applied to systems composed of linear and quadratic state transitions. Importantly, systems with only linear and pairwise (quadratic) interactions can be viewed as a locally-quadratic approximation of a more general smooth nonlinear system (Ale et al., 2013), and Gaussian moment closure therefore provides a general approach to deriving approximate state-space models in neural population dynamics.

The state-space interpretation of neural field models opens up future work to leverage the algorithmic tools of SSM estimation for data assimilation with spiking point-process datasets. However, challenges remain regarding the retinal waves explored here, and future work is needed to address these challenges. Model likelihood estimation is especially challenging. Despite this, the connection between neural-field models and state-space models derived here will allow neural field modeling to incorporate future advances in estimating recursive, nonlinear, spatiotemporal models. We also emphasize that some of the numerical challenges inherent to high-dimensional spatially extended neural field models do not apply.
to simpler, low-dimensional neural mass models, and the moment-closure framework may therefore provide a practical avenue to parameter inference in such models.

In summary, this report connects neural field models, which are grounded in models of stochastic population dynamics, to latent state-space models for population spiking activity. This connection opens up new approaches to fitting neural field models to spiking data. We expect that this interpretation is a step toward the design of coarse-grained models of neural activity that have physically interpretable parameters, have physically measurable states, and retain an explicit connection between microscopic activity and emergent collective dynamics. Such models will be essential for building models of collective dynamics that can predict the effects of manipulations on single-cells on emergent population activity.

Materials & Methods

Data acquisition and preparation

Example retinal wave datasets are taken from Maccione et al. (Maccione et al., 2014). Spikes were binned at 100 ms resolution for analysis. Spikes were further binned into regions on a 20×20 spatial grid. For the three-state model, this resulted in a 1200-dimensional spatiotemporal system, which provided an acceptable trade-off between spatial resolution and numerical tractability.

Spiking activity in each region was segmented into wave-like and quiescent states using a two-state hidden Markov model with a Poisson observations. To address heterogeneity in the Retinal Ganglion Cell (RGC) outputs, the observation model was adapted to each spatial region based on firing rates. Background activity was used to establish per-region biases, defined as the mean activity in a region during quiescent periods. The scaling between latent states and firing rate (gain) was adjusted locally based on the mean firing rate during wave events. The overall (global) gain for the observation model was then adjusted so that no wave events exhibited a fraction of cells in the active (A) state greater than one.

Moment-closure for a single population

To develop a state-space formalism for inference and data assimilation in neural field models, we begin with a master equation approach. This approach has been used before to analyze various stochastic neural population models, often as a starting point to derive ordinary differential equations for the moments of the distribution of population states, as we do here (Ohira and Cowan, 1993; Buice and Cowan, 2007; Bressloff, 2009; El Boustani and Destexhe, 2009; Buice et al., 2010; Touboul and Ermentrout, 2011). In our case, we examine a three-state system of the kind proposed in Buice and Cowan (Buice and Cowan, 2007, 2009), and use a Gaussian moment-closure approach similar to Bressloff (Bressloff, 2009).

The master equation describes how the joint probability distribution of neural population states (in our example the active, quiescent and refractory states) evolves in time. However, modelling this full distribution is computationally prohibitive for a spatially-extended system, since the number of possible states scales exponentially with the number of neural populations. Instead, we approximate the time evolution of the moments of this distribution.

In principle, an infinite number of moments are needed to describe the full population activity. To limit this complexity, we consider only the first two moments (mean and covariance), and use a moment-closure approach to close the series expansion of network interactions in terms of higher moments ((Schnoerr et al., 2017; Gomez-Uribe and Verghese, 2007; Goodman, 1953; Whittle, 1957); for applications to neuroscience see (Ly and Tranchina, 2007; Bressloff, 2009; El Boustani and Destexhe, 2009; Buice et al., 2010; Touboul and Ermentrout, 2011; Rule and Sanguineti, 2018)). Using this strategy, we obtain a second-order neural field model that describes how the mean and covariance of population spiking evolve in time, and recapitulates spatiotemporal phenomena when sampled.

We may describe the number of neurons in each state in terms of a probability distribution \( \Pr(Q, A, R) \) (Figure 2A), where we slightly abuse notation and use \( Q, A, \) and \( R \) both as symbols for the neuron states and as variables counting the neurons in the corresponding states, i.e. non-negative integers. The time evolution of this probability distribution captures stochastic population dynamics, and is represented by a master equation that describes the change in density for a given state \( (Q, A, R) \) when neurons change
states. Accordingly, the master equation describes the change in probability of a given state \(\{Q, A, R\}\) in terms of the probability of entering, minus the probability of leaving the state:

\[
\begin{align*}
\partial_t \Pr(Q, A, R) &= \Pr(Q, A+1, R-1)\rho_q(A+1) \quad \text{(transition } A \rightarrow R) \\
&+ \Pr(Q-1, A, R+1)\rho_r(R+1) \quad \text{(transition } R \rightarrow Q) \\
&+ \Pr(Q+1, A-1, R)\left[\rho_q + \rho_e(A-1)\right] \quad \text{(transition } Q \rightarrow A \text{ and } A \rightarrow Q \rightarrow A) \\
&- \Pr(Q, A, R)\left[\rho_qA + \rho_q + \rho_qA + \rho_rR\right] \quad \text{(outgoing transitions)}
\end{align*}
\]  

(5)

Even in this simplified non-spatial scenario, no analytic solutions are known for the master equation. However, from Eq. (5) one can derive equations for the mean and covariance of the process.

The approach, generally, is to consider expectations of individual states, e.g. \(\langle Q \rangle\) (first moments, i.e. means), or \(\langle QA \rangle\) (second moments), taken with respect to the probability distribution \(\Pr(Q, A, R)\) described by the master equation (5). Differentiating these moments in time, and substituting in the time-evolution of the probability density as given by the master equation, yields expressions for the time-evolution of the moments. However, in general these expressions will depend on higher moments and are therefore not closed.

For our system, the nonlinear excitatory interaction \(Q+A \rightarrow A+A\) couples the evolution of the means of the covariance \(\Sigma_{AQ}\), and the evolution of the covariance is coupled to the third moment, and so on. The moment equations are therefore not closed, and require an infinite number of moments to describe the evolution of the mean and covariance. To address this complexity, we approximate \(\Pr(Q, A, R)\) with a multivariate normal distribution at each time-point (Figure 2B), thereby replacing counts of neurons with continuous variables. This Gaussian moment-closure approximation sets all cumulants beyond the variance to zero, yielding an expression for the third moment in terms of the mean and covariance, leading to closed ordinary differential equations for the means and covariances (Goodman, 1953; Whittle, 1957; Gomez-UrIBE and Verghese, 2007; Schnoerr et al., 2017).

For our model with transitions given in Eq. (1) this leads to the system of ODEs for the mean values given in Eq. (2) in the main text. For the evolution of the covariance we obtain

\[
\Sigma_{\text{noise}} = J \Sigma + \Sigma J^T + \Sigma_{\text{noise}},
\]

where \(J\) is the Jacobian of the equations for the deterministic means in Eq. (2), and the \(\Sigma_{\text{noise}}\) fluctuations are Poisson and therefore proportional to the mean reaction rates (Eq. 2). Intuitively, the Jacobian terms \(J\) describe how the covariance of the state distribution 'stretches' or 'shrinks' along with the deterministic evolution of the means, and the additional \(\Sigma_{\text{noise}}\) reflects added uncertainty due to the fact that state transitions are stochastic. Each state experiences Poisson fluctuations with variance equal to the mean transition rates, due to the sum of transitions into and away from the state. Because the number of neurons is conserved, a positive fluctuation into one state implies a negative fluctuation away from another, yielding off-diagonal anticorrelations in the noise.

Together, equations (2) and (6) provide approximate equations for the evolution of the first two moments of the master equation (Eq. 5), expressed in terms of ordinary differential equations governing the mean and covariance of a multivariate Gaussian distribution. Here, we have illustrated equations for a 3-state system, but the approach is general and can be applied to any system with spontaneous and pairwise state transitions.

**Extension to spatial system**

To extend the moment equations (2) and (6) to a neural field system, we consider a population of neurons at each spatial location. In this spatially-extended case, we denote the intensity fields as \(Q, A,\) and \(R\), which are now vectors with spatial indices (or, in the spatially-continuous case: scalar functions of
coordinates x). In the spatially-extended system, active (A) neurons can excite nearby quiescent (Q) neurons. We model the excitatory influence of active cells as a weighted sum over active neurons in a local neighborhood, defined by a coupling kernel $K(\Delta x)$ that depends on distance (Eq. 4). To simplify the derivations that follow, denote the convolution integral in equation (4) as a linear operator $K$ such that

$$KA = K(\Delta x) * A(x).$$  \(7\)

In this notation, one can think of $K$ as a matrix that defines excitatory coupling between nearby spatial regions. Using the notation of Eq. (7), the rate that active cells excite quiescent ones is given by the product

$$\rho_e(KA) * Q = \rho_e \text{Diag}(KAQ^\top),$$  \(8\)

where $\circ$ denotes element-wise (in the spatially-continuous case: function) multiplication. For the time evolution of the first moment (mean intensity) of $Q$ in the spatial system, one therefore considers the expectation $\langle KAQ^\top \rangle$, as opposed to $\langle AQ \rangle$ in the non-spatial system. Since $K$ is a linear operator, and the extension of the Gaussian state-space model over the spatial domain $x$ is a Gaussian process, the second moment of the nonlocal interactions $KA$ with $Q$ can be obtained in the same way as one obtains the correlation for a linear transformation of a multivariate Gaussian variable:

$$\langle KAQ^\top \rangle = K\langle AQ^\top \rangle = K(\Sigma_{A,Q} + \langle A \rangle \langle Q \rangle^\top).$$  \(9\)

The resulting equations for the spatial means are similar to the nonspatial system (Eq. 2), with the exception that we now include spatial coupling in the rate at which quiescent cells enter the active state:

$$r_{qa} = \rho_q \langle Q \rangle + \rho_e \text{Diag}(KAQ^\top)$$

$$= \rho_q \langle Q \rangle + \rho_e \text{Diag}(K(\Sigma_{A,Q} + \langle A \rangle \langle Q \rangle^\top))$$

$$= \rho_q \langle Q \rangle + \rho_e \text{Diag}(K\Sigma_{A,Q} + K \langle A \rangle \circ \langle Q \rangle).$$  \(10\)

The numbers of neurons in the quiescent verses active states are typically anti-correlated, because a neuron entering the active state implies that one has left the quiescent state. Therefore, the expected number of interactions between quiescent and active neurons is typically smaller than what one might expect from the deterministic mean field alone. The influence of correlations $\text{Diag}(K\Sigma_{A,Q})$ on the excitation is therefore important for stabilizing the excitatory dynamics.

To extend the equations for the second moment to the neural field case, we consider the effect of spatial couplings on the Jacobian (Eq. 6). The spontaneous first-order reactions remain local, and so the linear contributions are similar to the non-spatial case. However, nonlocal interaction terms emerge in the nonlinear contribution to the Jacobian:

$$J_{\text{nonlinear}} = \rho_e \begin{bmatrix}
-\text{Diag}(K\langle A \rangle) & -\text{Diag}(\langle Q \rangle K) \\
\text{Diag}(K\langle A \rangle) & \text{Diag}(\langle Q \rangle K)
\end{bmatrix},$$  \(11\)

where here the “Diag” operation refers to constructing a diagonal matrix from a vector. Intuitively, the first column of Eq. (11) reflects the fact that the availability of quiescent cells modulates the excitatory effect of active cells, and the second column reflects the fact that the density active of neurons in nearby spatial volumes contribute to the rate at which quiescent cells become active.

**Basis projection**

The continuous neural field equations are simulated by projection onto a finite spatial basis $\mathcal{B}$. Each basis element is an integral over a spatial volume. Means for each basis element are defined as an integral over this volume, and correlations are defined as a double integral. For example, consider the number of
quiescent neurons associated with the $i^{th}$ basis function $B_i$, which we will denote as $Q_i$. The mean ($\langle Q_i \rangle$) and covariance $\Sigma_{Q/A}^{ij}$ between the quiescent and active states are given by the projections:

\[
\begin{align*}
\langle Q_i \rangle &= \int B_i(x)Q(x) \, dx \\
\Sigma_{Q/A}^{ij} &= \int B_i(x)B_j(x')\Sigma_{QA}(x, x') \, dx \, dx',
\end{align*}
\]

where $x$ and $x'$ range over spatial coordinates as in Eq. (3) and (4). When selecting a basis $B$, assumptions must be made about the minimum spatial scale to model. A natural choice is the radius of lateral (i.e. spatially nonlocal) interactions in the model $\sigma$, (Eq. 3), since structure below this scale is attenuated by the averaging over many nearby neurons in the dendritic inputs.

**Sampling from the model**

For ground-truth simulations, we sample from a hybrid stochastic model derived from a Langevin approximation to the three-state neural field equation. In this approximation, the deterministic evolution of the state is given by the mean-field equations (Eq. (2) for a local system, Eq. (10) for the neural field system), and the stochastic noise arising from Poisson state transitions is approximated as Gaussian as given by second-order terms (i.e. $\Sigma_{\text{noise}}$ in Eq. (6); see also (Rieder and Buckwar, 2013; Schnoerr et al., 2017)). Spontaneous wave initiation events are too rare to approximate as Gaussian, and instead are sampled as Poisson (shot) noise, giving us a hybrid stochastic model:

\[
r_q(t) \sim \text{Poisson}(\rho_q \cdot dt) \cdot \delta(t),
\]

where $\delta(t)$ is a Dirac delta (impulse). To avoid uniform spontaneous excitation, the excito-excitatory reaction rate is adjusted by a small finite threshold $\vartheta$, i.e. $r_{qa} \leftarrow \max(0, r_{qa} - \vartheta)$ in Eq. (10). For our simulations (e.g. Figure 3), we let $\vartheta = 8 \times 10^{-3}$. For the non-spatial system, the hybrid stochastic differential equation is:

\[
\begin{bmatrix}
\frac{dQ}{dt} \\
\frac{dA}{dt} \\
\frac{dR}{dt}
\end{bmatrix} =
\begin{bmatrix}
-r_q(t) & 0 & \rho_r \\
\rho_a & -\rho_a & 0 \\
0 & -\rho_r & -\rho_r
\end{bmatrix}
\begin{bmatrix}
Q \\
A + \rho_e \\
R
\end{bmatrix} +
\Sigma_{\text{noise}}^{1/2} dW,
\]

where $\Sigma_{\text{noise}}$ is the fluctuation noise covariance as in Equation (6) (with $\rho_a$ excluded, as it is addressed by the shot noise, Eq. 13), and $dW$ is the derivative of a multidimensional standard Wiener process, i.e. a spherical (white) Gaussian noise source. The deterministic component of (14) equation can be compared to Equation (2) for the means of the non-spatial system in the moment-closure system (without the covariance terms).

The stochastic differential equation for the spatial system is similar, consisting to a collection of local populations coupled through the spatial interaction kernel (Eqs. 3-4), and follows the same derivation used when extending the moment-closure to the spatial case (Methods: Extension to spatial system, Eqs. 7-10). When applying the Euler-Maruyama method method to the spatiotemporal implementation, fluctuations were scaled by $\sqrt{\Delta t \Delta x}$, where $\Delta x$ is the volume of the spatial basis functions used to approximate the spatial system (see Methods: System-size scaling for further detail). The Euler-Maruyama algorithm samples noise from a Gaussian distribution, and can therefore create negative intensities due to discretization error. We addressed this issue by using the complex Langevin equation (Schnoerr et al., 2014), which accommodates transient negative states.

**Point-process measurement likelihood**

Similarly to generalized linear point-process models for neural spiking (Truccolo et al., 2005, 2010; Truccolo, 2016), we model spikes as a Poisson process conditioned on a latent intensity function $\lambda(x, t)$, which characterises the probability of finding a certain number of spikes $k$ in a small spatiotemporal interval $\Delta x \times \Delta t$ as:

\[
\Pr \left( \int_{t_0}^{t_0+\Delta t} \int_{x_0}^{x_0+\Delta x} y(x, t) \, dx \, dt = k \right) \sim \text{Poisson} \left( k; \int_{t_0}^{t_0+\Delta t} \int_{x_0}^{x_0+\Delta x} \lambda(x, t) \, dx \, dt \right).\]
In (15), \( y(x, t) \) denotes the experimentally-observed spiking output, and is a sum over Dirac delta distributions corresponding to each spike with an associated time \( t_i \) and spatial location \( x_i \), i.e. \( y(x, t) = \sum_{i \in 1..N} \delta(x_i) \delta(t_i) \). We use a linear Poisson likelihood for which the point-process intensity function

\[
\lambda(x, t) = \gamma(x) A(x, t) + \beta(x)
\]

depends linearly on the number of active neurons \( A(x, t) \) with spatially-varying gain \( \gamma(x) \) and bias \( \beta(x) \). In other words, the observed firing intensity in a given spatiotemporal volume should be proportional to the number of active neurons, with some additional offset or bias \( \beta \) to capture background spiking unrelated to the neural-field dynamics.

**Bayesian filtering**

Having established an approach to approximate the time-evolution of the moments of a neural field system, we now discuss how Bayesian filtering allows us to incorporate observations in the estimation of the latent states. Suppose we have measurements \( y_0, \ldots, y_N \) of the latent state \( x \) at time \( t_0, \ldots, t_N \), given by a measurement process \( \Pr(y_i|x_i) \), which in our case is given by the point-process likelihood (Eq. 16). Bayesian filtering allows us to recursively estimate the filtering distribution \( \Pr(x_t|y_i, \ldots, y_0) \) at time \( t_i \), i.e. the posterior state probability at time \( t_i \) given the current and all previous observations. The procedure works by the following iterative scheme: i) suppose we know the filtering distribution \( \Pr(x_{t_i}|y_i, \ldots, y_0) \) at time \( t_i \). Solving the model dynamics forward in time up to \( t_{i+1} \) gives the predictive distribution \( \Pr(x_t|y_i, \ldots, y_0) \) for all times \( t_i < t \leq t_{i+1} \). ii) at the time \( t_{i+1} \), the measurement \( y_{i+1} \) needs to be taken into account which can be done by means of the Bayesian update:

\[
\Pr(x_{t+1}|y_{t+1}, \ldots, y_0) = \frac{\Pr(y_{i+1}|x_{t+1}) \Pr(x_{t+1}|y_i, \ldots, y_0)}{\Pr(y_{i+1}|y_i, \ldots, y_0)},
\]

where we have used the Markov property and \( \Pr(y_{i+1}|x_{t+1}, y_i, \ldots, y_0) = \Pr(y_{i+1}|x_{t+1}) \) to obtain the right hand side. Eq. (17) gives the filtering \( \Pr(x_{t+1}|y_i, \ldots, y_0) \) at time \( t_{i+1} \) which serves as the input of the next \( i \) step. Performing steps i) and ii) iteratively hence provides the filtering distribution for all times \( b_0 \leq t \leq t_n \).

For our neural field model we must compute both steps approximately: to obtain the predictive distribution in step i) we integrate forward the differential equations for mean and covariance derived from moment-closure (Eq. 2-6 and Methods: Extension to spatial system). In practice, we convert the continuous-time model to discrete time. If \( F_{\Delta t} \) denotes the local linearization of the mean dynamics in continuous time such that \( \partial_{\Delta t} \mu(t) = F_{\Delta t} \mu(t) \), then the approximated discrete-time forward operator is

\[
F_{\Delta t} = \exp(F_{\Delta t} \Delta t) \approx I + F_{\Delta t} \Delta t.
\]

We update the covariance using this discrete-time forward operator, combined with an Euler integration step for the Poisson fluctuations. A small constant diagonal regularizion term \( \Sigma_{\text{reg}} \) can be added, if needed, to improve stability. The resulting equations read:

\[
\mu_{t+\Delta t} = F_{\Delta t} \mu_t
\]

\[
\Sigma_{t+\Delta t} = F_{\Delta t} \Sigma_t F_{\Delta t}^T + \Sigma_{\text{noise}} \Delta t + \Sigma_{\text{reg}}.
\]

This form is similar to the update for a discrete-time Kalman filter (Kalman et al., 1960; Kalman and Bucy, 1961), the main difference being that the dynamics between observation times are taken from the nonlinear moment equations.

Consider next the measurement update of step ii) in Eq. (17). Since the Gaussian model for the latent states \( x \) is not conjugate with the Poisson distribution for observations \( y \), we approximate the posterior \( \Pr(x_{t+1}|y_{t+1}, \ldots, y_0) \) using the Laplace approximation (c.f. (Paninski et al., 2010; Macke et al., 2011)). The Laplace-approximated measurement update is computed using a Newton-Raphson algorithm. The measurement update is constrained to avoid negative values in the latent fields by adding a \( \epsilon/x \) potential (compare to the log-barrier approach; (Nazarpour et al., 2012)), which ensures that the objective function gradient points away from this constraint boundary, where \( x \) is the intensity of any of the three fields.
The gradients and Hessian for the posterior measurement log-likelihood $\ln L$ are

$$
-\ln L = \frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) + v y x + \beta - y \ln(y x + \beta)
$$

$$
-\frac{\partial \ln L}{\partial x} = \Sigma^{-1} (x - \mu) + \nu y - y \frac{y}{y x + \beta}
$$

$$
-\frac{\partial^2 \ln L}{\partial x^2} = \Sigma^{-1} + y \left( \frac{y}{y x + \beta} \right)^2,
$$

where $x$ is the latent state with prior mean $\mu$ and covariance $\Sigma$, and couples to point-process observations $y$ linearly with gain $\gamma$ and bias $\beta$ as in Eq. (16). The parameter $v = \Delta x^2 / \Delta t$ is the spatiotemporal volume of the basis function or spatial region over which the counts are observed.

### System-size scaling

For clarity, the derivations in this paper are presented for a population of neurons with a known size, such that the fields $Q(x)$, $A(x)$, and $R(x)$ have units of neurons. In practice, the population size $\Omega$ of neurons is unknown, and it becomes expedient to work in normalized intensities, where $Q(x)$, $A(x)$, and $R(x)$ represent the fraction of neurons in a given state between 0 and 1, and are constrained such that $Q(x) + A(x) + R(x) = 1$. In this normalized model for population size $\Omega$, quadratic interaction parameters (like $\rho_{ij}$) as well as the gain are multiplied by $\Omega$, to reflect the re-scaled population. In contrast, noise variance should be divided by $\Omega$ to account for the fact that the coefficient of variation decreases as population size increases. Although rescaling by $\Omega$ is well-defined for finite-sized populations, the infinitesimal neural-field limit for the second-order model is not. This is because, while the mean-field equations scale with the population size $O(\Omega)$, the standard deviation of Poisson fluctuations scales with the square root of the population size $O(\sqrt{\Omega})$. The ratio of fluctuations to the mean (coefficient of variation) therefore scales as $O(1/\sqrt{\Omega})$, which diverges as $\Omega \rightarrow 0$.

This divergence is not an issue in practice as all numerical simulations are implemented on a set of basis functions with finite nonzero volumes, and each spatial region is therefore associated with finite nonzero population size. Even in the limit where fluctuations would begin to diverge, one can treat the neural field equations as if defined over a continuous set of overlapping basis functions with nonzero volume. Conceptually, this can be viewed as setting a minimum spatial scale for the neural field equations, which is defined by spatial extent of each local population. If the model is defined over a set of overlapping spatial regions, then these populations experience correlated fluctuations. Consider Poisson fluctuations as entering with some rate-density $\sigma^2(x)$ per unit area. The observed noise variances and covariances, projected onto basis functions $B_i(x)$ and $B_j(x)$, are:

$$
\Sigma_{i,j}^{\text{noise}} = \int B_i(x) B_j(x) \sigma^2(x) dx
$$

If the neuronal population density is given as $\rho(x)$ per unit area, then the effective population size for a given basis function is:

$$
\Omega_i = \int B_i(x) \rho(x) dx
$$

If the population density is uniform, and if basis functions have a constant volume $v$, we can write this more simply as $\Omega = v \rho$. In the system-size normalized model, the contributions of basis function volume cancel and the noise variance should be scaled simply as $1/\rho$.

### Acknowledgements:

We thank Gerrit Hilgen for important discussions in establishing biologically-plausible parameter regimes for the three-state model. We thank Evelyne Sernagor for the retinal wave datasets, as well as ongoing advice and invaluable feedback on the manuscript. We also thank Dimitris Milios and Botond Cseke for helpful technical discussions.
References


Bojak I, Oostendorp TF, Reid AT, Köter R. Connecting Mean Field Models of Neural Activity to EEG and fMRI Data. Brain Topography 2010 jun;23(2):139–149.


Figure Legends

Fig 1. 3-state Quiescent-Active-Refractory (QAR) neural-field model. Cells in the developing retina are modeled as having three activity states. Active cells (A; red) fire bursts of action potentials, before becoming refractory (R; green) for an extended period of time. Quiescent (Q; blue) cells may burst spontaneously, or may be recruited into a wave by other active cells. These three states are proposed to underlie critical multi-scale wave dynamics (Hennig et al., 2009b).

Fig 2. Summarizing estimated neural state as population moments. (A) The activity within a local spatial region (encircled, left) can be summarized by the fraction of cells (represented by colored dots) in the quiescent (blue), active (red), and refractory (green) states (Q, A, R; right). (B) An estimate of the population state can be summarized as a probability distribution \( \Pr(Q, A, R) \approx \mathcal{N}(\mu, \Sigma) \) over the possible proportions of neurons in each state. A Gaussian moment-closure approximates this distribution as Gaussian, with given mean and covariance (orange crosshairs).

Fig 3. Spatial 3-state neural-field model exhibits self-organized multi-scale wave phenomena. Simulated example states at selected time-points on a \([0, 1]^2\) unit interval using a 20×20 grid with effective population density of \( \rho=50 \) cells per unit area, and rate parameters \( \sigma=0.075, \rho_a=0.4, \rho_r=3.2\times10^{-3}, \rho_e=0.028, \) and \( \rho_q=0.25 \) (Methods: Sampling from the model). As, for instance, in neonatal retinal waves, spontaneous excitation of quiescent cells (blue) lead to propagating waves of activity (red), which establish localized patches in which cells are refractory (green) to subsequent wave propagation. Over time, this leads to diverse patterns of waves at a range of spatial scales.
Fig 4. Hidden Markov model for latent neural fields. For all time-points $T$, state transition parameters $\theta=(\rho_q, \rho_a, \rho_r, \rho_e, \sigma)$ dictate the evolution of a multivariate Gaussian model $\mu, \Sigma$ of latent fields $Q, A, R$. The observation model $(\beta, \gamma)$ is a linear map with adjustable gain and threshold, and reflects how field $A$ couples to firing intensity $\lambda$. Point-process observations (spikes) $y$ are Poisson with intensity $\lambda$.

Fig 5. Filtering recovers latent states in ground-truth simulated data. Spatially averaged state occupancy (blue, red, and green: $Q, A, R$) (vertical axis) is plotted over time (horizontal axis). Solid lines represent true values sampled from the model, and shaded regions represent the 95% confidence interval estimated by filtering. The active ($A$) state density has been scaled by $\times 25$ for visualization. Colored plots (below) show the qualitative spatial organization of quiescent (blue), active (red), and refractory (green) neurons. Model parameters are the same as Figure 3, with the exception of the spatial resolution, which has been reduced to $9 \times 9$. 
Fig 6. Retinal waves recorded via high-density multi-electrode arrays. (A) Spontaneous retinal waves are generated in the inner retina via laterally interacting bipolar (blue) and amacrine (red) cells, depending on the developmental age. These waves activate retinal ganglion cells (yellow), the output cells of the retina. Electrical activity is recorded from the neonatal mouse retina via a 64×64-electrode array with 42 μm spacing. (B) Average spiking rate recorded across the retina (the central region devoid of recorded spikes is the optic disc). This example was recorded on postnatal day 6. (C) Spikes were binned at 100 ms resolution, and assigned to 10×10 spatial regions for analysis. Spiking activity on each electrode was segmented into "up" states (during wave activity) and "down" states (quiescent) using a two-state hidden Markov model with Poisson observations. In this example, most waves and inter-wave intervals lasted between one and ten seconds. (D) Example wave event, traveling across multiple spatial regions and lasting for a duration of 16-20 seconds.
Fig 7. **State inference via filtering: retinal datasets.** We apply a calibrated model to spiking observations from retinal waves (postnatal day 11) to infer latent neural-field states. In all plots, red, green, and blue indicate (normalized) densities of active, refractory, and quiescent cells. **(top)** Solid lines indicate inferred spatial means, and shaded regions the 95% confidence bound. The A state has been scaled-up by $\times5$. Example time slices are shown in the colored plots below. Dark regions indicate areas absent from the recording. Summary statistics are shown on the right, with power spectra (averaged over all included regions and states) indicating periodic $\sim5$ waves/min, and the typical fraction of Q/A/R states, pooled over all times and regions, summarized in histograms below. **(bottom)** Forward simulation of the calibrated model without data recapitulates the retinal wave activity. Solid lines indicate sampled spatial means. Colored plots show example time slices. Wave frequency is comparable to the data ($\sim5$ waves/min), and occupancy statistics are similar. The model was initialized with 70% of cells quiescent and 30% refractory, with a 25 s burn-in to remove initial transients.