Boosting the Speed and Precision of SNN Parameter Estimation

Huimiao Chen

Jun. 9, 2023

Unveiling Neuronal Activity with SNNs

Question: Can we use spiking neural networks (SNNs) as generative models of multi-neuronal recordings, while taking into account that most neurons are unobserved?

- If we could estimate the parameters of SNNs using limited experimental data, it would greatly enhance our ability to analyze cortical circuits.
- It can be utilized to simulate and predict various neuronal activities, analyze neuronal dynamics, and could potentially be extended to model the cerebral cortex and analyze neurological disorders.

Unveiling Neuronal Activity with SNNs

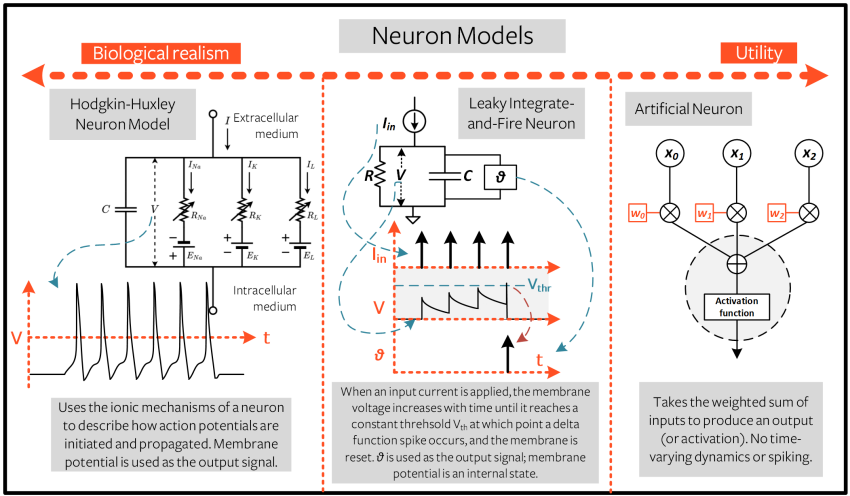
- Advancements in electrophysiological recording technologies help capture spiking activity from more neurons – enabling us to consider reverse-engineering neuronal parameters and simulate neural networks in specific brain regions.
- However, the immense number of neurons in the brain means that the recorded data represents only a small fraction of the total, presenting challenges for parameter estimation.
- In reality, a large proportion of neurons can be viewed as belonging to several groups, each exhibiting similar characteristics.
- In studies simulating cortical networks with spiking neurons, the number of different cell types or neuronal populations in each cortical column ranges from about 10 to 200, yet a single cortical column may contain tens of thousands of simulated neurons, and multiple cortical columns can have millions of neurons.

Unveiling Neuronal Activity with SNNs

Consequently,

- To address this, we can group neurons exhibiting similar characteristics, as a large proportion of neurons can be categorized into several distinct groups.
- Avoiding the computational intensity of complete microscopic models, mesoscopic models provide an efficient and interpretable way to simulate neuronal activity by using effective statistical methods.
- Coarser scale models, such as neural mass models, capture certain characteristics well but may not align closely with microscopic models in dynamic situations.

Neuron Models: ‘Artificial’ vs ‘Biological’



Simplifying HH Model

The HH Model is usually too complicated with three gating variables. So we have many simplified models, like the reduced HH model and the FitzHugh-Nagumo model.

The Hodgkin-Huxley equations include a dynamical equation for voltage:

$$C_{\mathrm{m}}\dot{V} = \bar{g}_{\mathrm{Na}} m^3 h (E_{\mathrm{Na}} - V) + \bar{g}_{\mathrm{K}} n^4 (E_{\mathrm{K}} - V) + g_{\mathrm{L}}(E_{\mathrm{L}} - V) + I \tag{4.7}$$

and three kinetic equations for the gating variables:

$$\dot{m} = \alpha_m(V)(1-m) - \beta_m(V)m, \tag{4.8}$$
$$\dot{h} = \alpha_h(V)(1-h) - \beta_h(V)h, \tag{4.9}$$
$$\dot{n} = \alpha_n(V)(1-n) - \beta_n(V)n, \tag{4.10}$$

where the rate functions $\alpha_x > 0$ and $\beta_x > 0$ (with x standing for m, h and n) are empirical voltage-dependent functions obtained from the experimental data:

$$\alpha_m(V) = \frac{0.1(25-V)}{\exp\big((25-V)/10\big) - 1}, \qquad \beta_m(V) = 4\exp(-V/18), \tag{4.11}$$

$$\alpha_h(V) = 0.07\exp(-V/20), \qquad \beta_h(V) = \frac{1}{\exp\big((30-V)/10\big) + 1}, \tag{4.12}$$

$$\alpha_n(V) = \frac{0.01(10-V)}{\exp\big((10-V)/10\big) - 1}, \qquad \beta_n(V) = 0.125\exp(-V/80). \tag{4.13}$$

The reduced Hodgkin-Huxley model is governed by two equations:

$$C\dot{V} = \bar{g}_{\mathrm{Na}} m_\infty(V)^3 (c-n)(E_{\mathrm{Na}} - V) + \bar{g}_{\mathrm{K}} n^4 (E_{\mathrm{K}} - V) + g_{\mathrm{L}}(E_{\mathrm{L}} - V) + I, \tag{5.11}$$

$$\tau_n(V)\dot{n} = n_\infty(V) - n, \tag{5.12}$$

where all parameters are identical to the original Hodgkin-Huxley model. The dynamics of the system resembles the original Hodgkin-Huxley model (Fig. 5.15). The reduced system is able to generate an action potential although it rises faster and also falls faster than the original model.

Since a continuous one-dimensional system cannot oscillate, to model the action potential we need to add another variable to restore the voltage V after spiking. In FitzHugh-Nagumo model, a slow recovery variable W is added to Eq. 5.2 to bring down voltage V, and W satisfies a linear equation (Eq. 5.3) with a large time constant τ. The dynamical equations read:

$$\dot{V} = V - V^3/3 - W + I, \tag{5.2}$$

$$\tau\dot{W} = -W + AV + B. \tag{5.3}$$

Complicating LIF model

The LIF Model is usually too simple. So we have more complicated models, like the generalized integrate-and-fire (GIF) model and the Adaptive Exponential Integrate-and-Fire (AdEx) model.

The leaky integrate-and-fire model is equivalent to a simple RC circuit with resistance R and capacitance C. The dynamics is described by

$$C\dot{V} = -V/R + I \tag{5.25}$$

where I is the current input. In the absence of input ($I = 0$), the voltage decays exponentially to the resting state $V = 0$. An equivalent form of the equation is

$$\tau\dot{V} = -V + RI \tag{5.26}$$

where $\tau = RC$ is the time constant. In the limit of large resistance $R \to \infty$, Eq. 5.25 becomes $C\dot{V} = I$, which is a *perfect integrator* because the value of V faithfully reflects the time integral of input I.

General SNNs of GIF neurons

We consider a general SNN of GIF neurons.

- The GIF model includes additional features to make the model more biologically accurate.
- It accounts for spike-triggered adaptation, where the neuron's response changes based on spike history.
- This is done by adding extra terms to the differential equation of the LIF model that account for spike-triggered adaptation of a dynamic threshold for firing and a escape noise mechanism.

General SNNs of GIF neurons

Considering a single neuron, the synaptic input current of the neuron can be represented as the aggregate of post-synaptic currents initiated by each pre-synaptic neuron's spike.

$$RI_{\mathrm{syn},i}(t) = \tau_{\mathrm{m}} \sum_j w^{ij} \left(\epsilon^{ij} * s_j\right)(t)$$

$s_i(t)$: the spikes of a neuron i as time goes forward. $s_i(t) = \sum_k \delta\left(t - t_{i,k}\right)$
$t_{i,k}$: the time of the k-th spike.
δ: the Dirac function representing a spike.
R: the membrane resistance.
τ_{m}: the membrane time constant.
w^{ij}: the synaptic weights between neurons i and j.

$\epsilon^{ij}(t)$: the post-synaptic current, normalized by its charge (or, equivalently, its integral).
$\epsilon^{ij}(t) = \mathcal{H}(t - \Delta^{ij}) e^{-(t-\Delta^{ij})/\tau^j_{\mathrm{syn}}} / \tau^j_{\mathrm{syn}}$ ($1/\tau^j_{\mathrm{syn}}$ normalizes the kernel).
$\mathcal{H}$: the Heaviside step function.
τ^j_{syn}: the synaptic time constant.
Δ^{ij}: the synaptic delay.
$*$: the convolution operation.

General SNNs of GIF neurons

If considering a network reasonably approximated by multiple neuronal populations, the synaptic input current of neuron i can be recast as the summation of connections originating from all neuronal populations, including the population to which neuron i itself belongs.

$$RI_{\mathrm{syn},i}(t) = \tau_{\mathrm{m}} \sum_{\xi} w^{i\xi} \sum_{j \in \Pi_i^{\xi}} \left(\epsilon^{i\xi} * s_j^{\xi} \right)(t)$$

ξ: the index of neuronal populations.
$i\xi$: the relationship between neuron i and neurons belonging to population ξ, wherein neurons within a population are assumed to have identical parameters.
Π_i^{ξ}: the set of neurons that are part of population ξ and are interconnected with neuron i.

Conditional Intensity Function in GIF Model

The GIF model incorporates a conditional intensity function to capture the probability of neuron firing given a specific threshold and membrane potential.

$$\lambda_i(t) = f_i\left(u_i(t) - \vartheta_i(t)\right)$$

$\vartheta_i(t)$: the threshold.
$f_i(x)$: the exponential link function, $f_i(x) = c_i e^{x/\Delta_{u,i}}$.
c_i: the escape rate at the threshold (the base rate of the exponential link function).
$\Delta_{u,i}$: the degree of threshold softness, $\Delta_{u,i} > 0$.

Firing Probability in GIF Model

Given $\lambda_i(t)$, we can have the firing probability of neuron i at time t:

$$P_i(t) = 1 - e^{-\int_t^{t+\Delta t} \lambda_i(\tau)\mathrm{d}\tau} \approx 1 - e^{-\bar{\lambda}_i(t)\Delta t} \approx \bar{\lambda}_i \Delta t$$

$P_i(t) = 1 - e^{-\int_{t_l}^{t_l+\Delta t} \lambda_i(t)\mathrm{d}t}$ is held as valid since $\lambda_i(t)$ is essentially the rate of a Poisson distribution for the firing of neuron i.

How to derive:

(1) The probability of x events occurring in a unit time is given by $P\{X = x\} = \frac{\lambda^x e^{-\lambda}}{x!}$.

(2) Extending this to x events in time Δt, we have $P\{X = x, \lambda\Delta t\} = \frac{(\lambda\Delta t)^x e^{-\lambda\Delta t}}{x!}$.

(3) The probability of no events occurring within the time interval $[t, t + \Delta t)$ can be written as $P\{X = 0, \lambda\Delta t\} = e^{-\lambda\Delta t}$.

(4) For an inhomogeneous Poisson process, the probability of observing no events within the time interval $[t, t + \Delta t)$ is represented by the second term in the above formula.

Note: Δt is assumed to be sufficiently small.

Subthreshold Dynamics

The escape rate $\lambda_i(t)$ needs to work with the potential from the subthreshold dynamics.

$$\tau_{\mathrm{m}} \frac{\mathrm{d}u_i}{\mathrm{d}t} = -u_i + \mu_i(t) + RI_{\mathrm{syn},i}(t)$$

$\mu_i(t) = u_{\mathrm{rest},i} + RI_{\mathrm{ext},i}(t)$: the convergent value of the potential when there is no synaptic current input, composed of a static resting potential $u_{\mathrm{rest},i}$ and an external stimulus $I_{\mathrm{ext},i}$.

Threshold Adaptation and Potential Adaptation

When considering threshold adaptation, every spike $t_{i,k}$ contributes to the dynamic threshold $\vartheta_i(t)$ through a spike-triggered threshold kernel $\theta_i\left(t - t_{i,k}\right)$.

$$\vartheta_i(t) = u_{\mathrm{th},i} + \sum_{t_{i,j}<t} \theta_i\left(t - t_{i,k}\right) = u_{\mathrm{th},i} + (\theta_i * s_i)(t)$$

where $(\theta_i * s_i)(t) = \int_{-\infty}^{t} \theta_i\left(t-\tau\right) s_i\left(\tau\right) \mathrm{d}\tau$.

Note: If to avoid the manual reset of the membrane potential after firing, we can introduces a spike-triggered potential kernel $\eta_i(t)$, similar to the spike-triggered threshold kernel, to accommodate effects including the refractoriness.
Fortunately, $\eta_i(t)$ can be integrated into $\theta_i(t)$, that is, $\theta \to (\theta - \eta)$ and $\eta \to 0$. The resulting membrane potential then acts as a free membrane potential, unaffected by any previous spike history. The final mathematical form is the same.

Population Activity

We transition from the microscopic model to a neuronally-grounded mesoscopic model, shifting our perspective from the level of individual neurons to the aggregate level of neuronal populations and thereby effectively reducing the statistical dimensionality and computational complexity of the model.

The population activity can be represented as follows.

$$A(t) = \frac{\Delta n(t)}{N\Delta t},\ \bar{A}(t) = \frac{\Delta \bar{n}(t)}{N\Delta t}$$

N: the number of interconnected neurons in a population.
p^{ξ}: the probability of current population connecting to a neuron from a population ξ.
A: the population activity.
$\bar{A}$: the mean population activity.
$\Delta n(t)$: the number of neuron firings in the population during a time period t. t now stands for a small time interval Δt.

Mean-Field Approximation

By employing the mean-field approximation, we can express the current and membrane potential as follows.

$$RI(t) = \tau_{\mathrm{m}} \sum_{\xi=1}^{X} p^{\xi} N^{\xi} w^{\xi} \left(\epsilon^{\xi} * A^{\xi} \right)(t)$$

$$\tau_{\mathrm{m}} \frac{\partial u}{\partial t} = -u + \mu(t) + \tau_{\mathrm{m}} \sum_{\xi=1}^{X} p^{\xi} N^{\xi} w^{\xi} \left(\epsilon^{\xi} * A^{\xi} \right)(t)$$

The subscript and superscript i are omitted here.

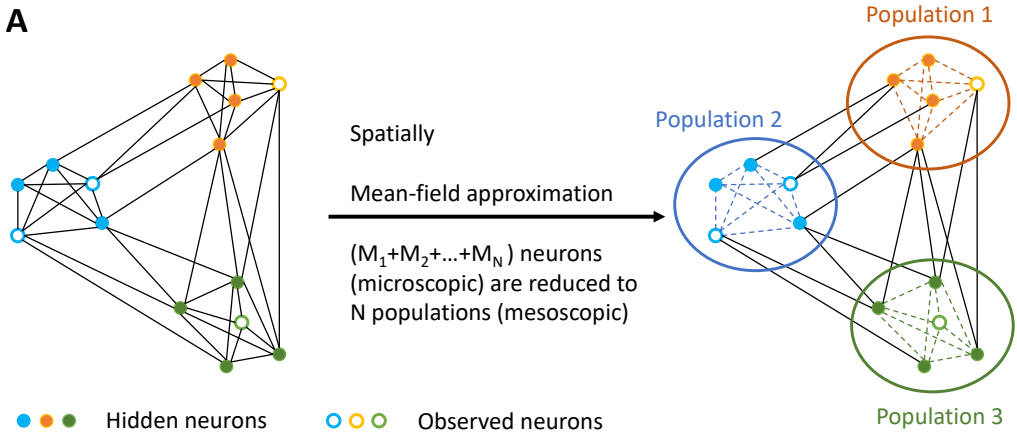
Quasi-Renewal Approximation

With the quasi-renewal approximation, the population firing rate $\lambda(t)$ is reformulated into a simpler yet effective format, $\lambda_i(t \mid \hat{t}_i)$, which considers only the most recent spike $\hat{t}_i$ of the neuron and the historical population activity.

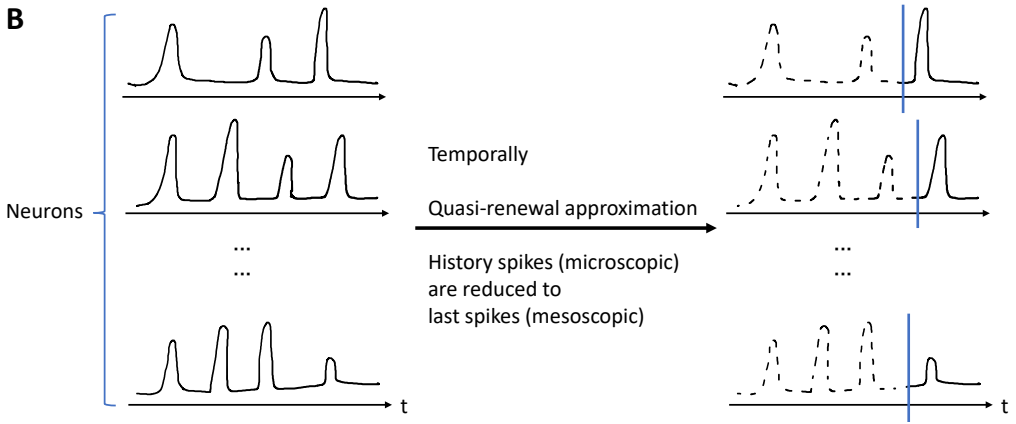
$$\lambda_i(t) = f_i\left(u_i(t) - \vartheta_i(t)\right) \approx f(u_A(t, \hat{t}_i) - \vartheta_i(t)) \rightarrow \lambda_A(t \mid \hat{t}_i) \approx f(u_A(t, \hat{t}_i) - \vartheta_A(t, \hat{t}_i))$$

The quasi-renewal approximation uses an adaptation kernel to describe the effect of spikes preceding the last spike, which can be considered as the "average" of those spikes.
$\vartheta_A(t, \hat{t}) = u_{\text{th}} + \theta(t - \hat{t}) + \int_{-\infty}^{\hat{t}} \tilde{\theta}\left(t - \tau\right) A(\tau) \mathrm{d}\tau$ and $\tilde{\theta}(t) = \Delta_u[1 - e^{-\theta(t)/\Delta_u}]$.
Δ_u: the softness coefficient. u_{th}: the potential threshold.
$\theta(t)$: the basic kernel of the adaptation of spikes on the potential threshold, same as that in the microscopic model, $\theta(t) = \frac{J_\theta}{\tau_\theta} e^{-t/\tau_\theta}$, where $\frac{1}{\tau_\theta} e^{-t/\tau_\theta}$ is a normalized term with area under the curve being 1, and J_θ and τ_θ are adaptation strength and adaptation time scale, respectively.
Note: $\theta(t)$ can include multiple exponential terms, i.e., $\theta(t) = \sum_z \frac{J_{\theta,z}}{\tau_{\theta,z}} e^{-t/\tau_{\theta,z}}$.

Schematic Diagram of Mean-Field Approximation



Schematic Diagram of Quasi-Renewal Approximation



Mesoscopic Description

Introduce essential intermediate terms for mesoscopic description.

- $m(t_1, t_2)$: the count of neurons at time t_1 having their last spikes at time t_2, with $\sum_{k=-\infty}^{t_{l-1}} m(t_1, t_2) = N$.
- $S(t_1 \mid t_2)$: the survival rate, $S(t_1 \mid t_2) = \frac{\langle \hat{m}(t_1, t_2) \rangle}{\Delta n(t_2)}$.
- $v(t_l, t_k)$: the variance, $v(t_l, t_k) = \frac{\langle \Delta \hat{m}^2(t_1, t_2) \rangle}{N \Delta t}$
- The mean $\langle \hat{m}(t_1, t_2) \rangle$ represents a mesoscopic variable. The hat over m denotes non-normalized variables. All these variables require initial conditions to commence the simulation of the mesoscopic model.

Mesoscopic Description

By transitioning from microscopic to mesoscopic variables, we reach a point of low-dimensional variable space without being overly coarse.

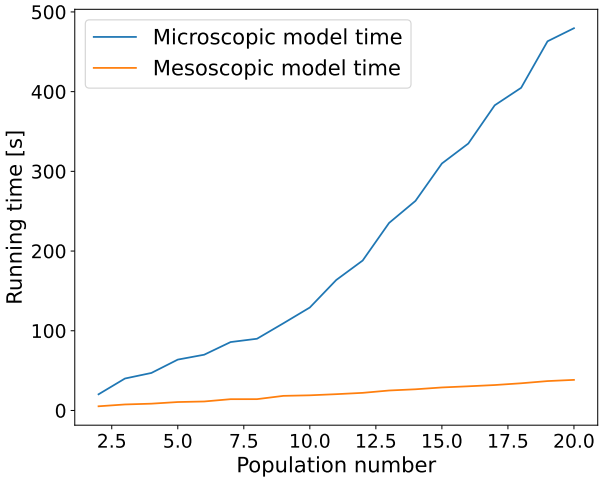
$$A(t) = \bar{A}(t) + \sqrt{\frac{\bar{A}(t)}{N}}\zeta(t)$$

where $\zeta(t)$ is a Gaussian white noise, and

$$\bar{A}(t) = \int_{-\infty}^{t} \lambda_A(t \mid \hat{t}) S(t \mid \hat{t}) A(\hat{t}) \mathrm{d}\hat{t} + \Lambda(t) \left(1 - \int_{-\infty}^{t} S(t \mid \hat{t}) A(\hat{t}) \mathrm{d}\hat{t}\right)$$

$$\lambda_A(t \mid \hat{t}) = c \exp\left(\frac{u(t,\hat{t}) - \vartheta_A(t,\hat{t})}{\Delta_u}\right), \quad \Lambda(t) = \frac{\int_{-\infty}^{t} \lambda_A(t \mid \hat{t}) v(t,\hat{t}) \mathrm{d}\hat{t}}{\int_{-\infty}^{t} v(t,\hat{t}) \mathrm{d}\hat{t}}$$

Simulation Time Comparison



As population increases, mesoscopic model demonstrates significantly shorter running time compared to microscopic model, making it the preferred choice for data collection.

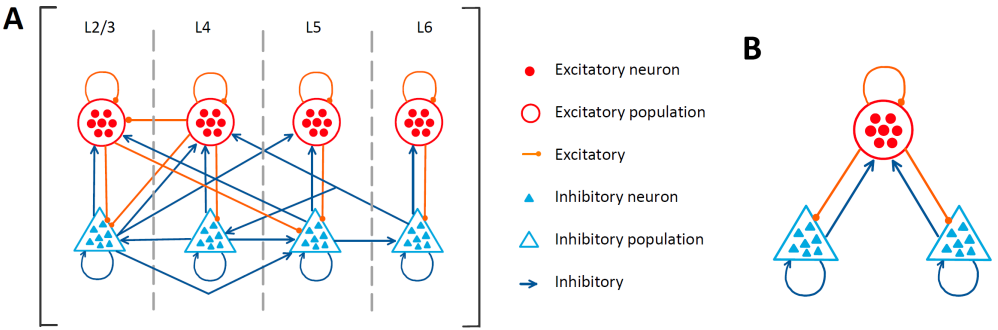
Parameter Settings

Key parameters that require estimation, including

- synaptic weights between each pair of populations (J_{syn})
- resting potential (μ, which combines the resting potential and the constant external input due to their similar effects on neuronal activities)
- membrane time constant (τ_m)
- baseline threshold voltage (u_{th})
- adaptation strength (J_θ)
- adaptation time scale (τ_θ)

We hold other parameters constant. We consider two scenarios: networks with 3 populations and 8 populations.

Schematic of Network Architecture



(A) The architecture of the eight-population is composed of four layers (L2/3, L4, L5, and L6). Each layer comprises both excitatory and inhibitory populations. **(B)** The architecture of the three-population comprises two competing excitatory populations and a shared inhibitory population, representing a winner-take-all SNN.

Parameter Settings: 3 Populations (Default)

- The number of neurons in the three populations is 400, 200, and 400, respectively.
- The base rate for the exponential link function is set to 10 Hz.
- The reset potential is at 0 mV, the refractory period is 4 ms.
- The threshold softness for adaptation is 2.5 mV.
- The transmission delay constant is 1 ms.
- The time constants for excitatory and inhibitory synapses are 3 ms and 6 ms, respectively.
- The connection probabilities are all set to 0.6.
- If the biological simulation time is 60 s, we have a step input of 20 mV starting at the 30 s mark; otherwise, there is no step input.

Parameter Settings: 8 Populations

Parameters of the modified Potjans-Diesmann model.

population	L2/3e	L2/3i	L4e	L4i	L5e	L5i	L6e	L6i
synaptic time constants [s]	0.0005							
transmission constant delay [s]	0.0015							
refractory period [s]	0.002							
softness of threshold adaptation [mV]	5.0							
external step input time	[0.06s, 0.09s]							
step stimulus [mV]	0	0	19	11.964	0	0	9.896	3.788
neuron number	20683	5834	21915	5479	4850	1065	14395	2948

Parameter Settings: 8 Populations

Connection probability for 8-population scenario.

	L2/3e	L2/3i	L4e	L4i	L5e	L5i	L6e	L6i
L2/3e	0.1009	0.1689	0.0437	0.0818	0.0323	0	0.0076	0
L2/3i	0.1346	0.1371	0.0316	0.0515	0.0755	0	0.0042	0
L4e	0.0077	0.0059	0.0497	0.135	0.0067	0.0003	0.0453	0
L4i	0.0691	0.0029	0.0794	0.1597	0.0033	0	0.1057	0
L5e	0.1004	0.0622	0.0505	0.0057	0.0831	0.3726	0.0204	0
L5i	0.0548	0.0269	0.0257	0.0022	0.06	0.3158	0.0086	0
L6e	0.0156	0.0066	0.0211	0.0166	0.0572	0.0197	0.0396	0.2252
L6i	0.0364	0.001	0.0034	0.0005	0.0277	0.008	0.0658	0.1443

Data Collection and Processing

Data collection: For the key parameters that need to be estimated, we generate data randomly within a reasonable range.

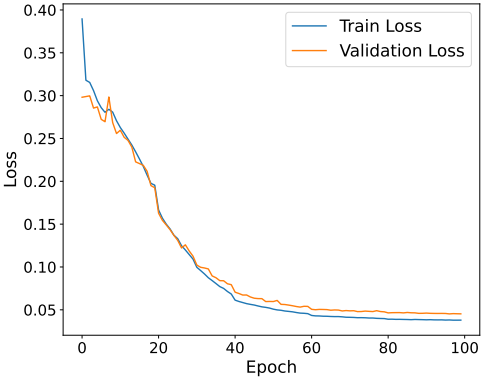
Data processing: The six parameters – the synaptic connectivity matrix J_{syn}, resting potential μ, membrane time constant τ_m, baseline threshold voltage u_{th}, adaptation strength J_θ, and adaptation time scale τ_θ – have dimensions of $Y \times Y$, Y, Y, Y, $Y \times Z$, and $Y \times Z$, respectively. (1) Averaging operation is applied to down-sample the data into shorter sequences. (2) Data normalization involves dividing all data points by the maximum absolute value in the dataset. (3) Maximum value for J_{syn} data is calculated independently for each population. (4) Denormalization is necessary to revert the output data to its original form using the previously determined normalization factors.

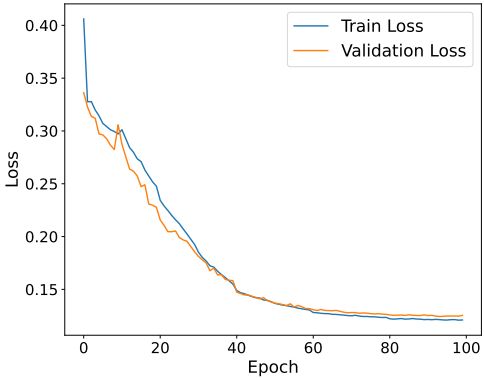
Y: the total number of populations.
Z: the number of exponential kernels used to approximate the adaptation from the spikes before the last spike.

RNN Model for Estimiation

- We train a simple artificial neural network to reverse-engineer the SNN parameters from the neuronal activity information.
- We opt for the relatively simpler gated recurrent unit (GRU) instead of long short-term memory (LSTM) in a recurrent neural network (RNN).
- It reads neural activity sequences from multiple populations and ultimately predicts the parameters with multi-head output.

Training Loss and Validation Loss





Numerical Comparison (3-Population, Sample 1)

		Labels			Estimations		
		P1	P2	P3	P1	P2	P3
J_{syn} **[mV]**	P1	-0.56	-0.27	0.00	-0.76	-0.13	0.20
	P2	-0.58	-0.27	0.07	-0.42	-0.30	-0.09
	P3	-0.60	-0.27	0.12	-0.58	-0.21	0.20
μ **[mV]**		48.94	50.31	30.73	49.66	49.87	30.55
τ_m **[ms]**		27.12	19.11	35.79	27.14	20.30	35.60
u_{th} **[mV]**		28.26	23.84	12.17	28.03	24.84	11.85
J_θ **[mV·ms]**		971.74	971.74	144.08	977.28	986.23	185.08
τ_θ **[ms]**		804.97	804.97	1340.50	799.58	790.15	1345.05
Normalized MSE loss				0.0053			

Numerical Comparison (3-Population, Sample 2)

		Labels			Estimations		
		P1	P2	P3	P1	P2	P3
J_{syn} **[mV]**	P1	-0.32	0.00	0.07	-0.46	-0.46	-0.06
	P2	-0.29	-0.32	0.00	-0.31	-0.43	-0.01
	P3	0.00	0.00	0.10	-0.03	-0.09	-0.14
μ **[mV]**		34.07	55.31	37.28	33.62	53.50	38.11
τ_m **[ms]**		27.27	18.28	20.93	27.75	18.82	21.21
u_{th} **[mV]**		11.42	16.40	26.81	12.17	16.43	26.87
J_θ **[mV·ms]**		674.67	674.67	1231.18	665.73	691.55	1246.66
τ_θ **[ms]**		508.75	508.75	879.77	496.08	498.71	888.23
Normalized MSE loss				0.0040			

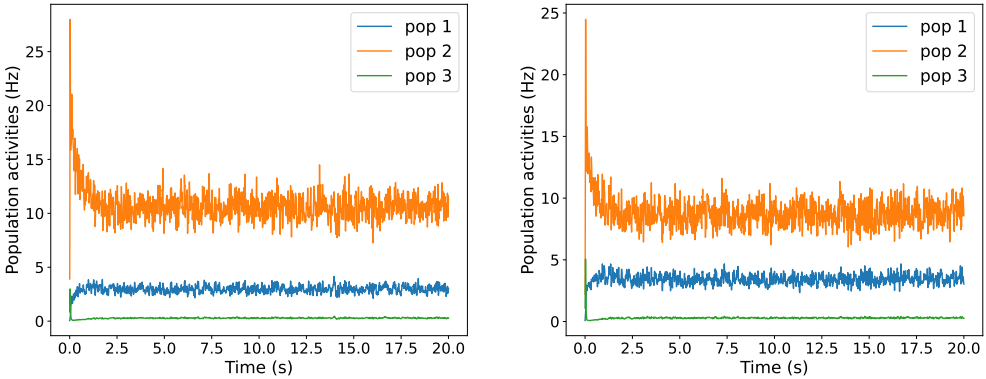
Numerical Comparison (3-Population, Sample 3)

		Labels			Estimations		
		P1	P2	P3	P1	P2	P3
J_{syn} **[mV]**	P1	0.16	-0.64	0.00	0.16	-0.70	-0.10
	P2	0.16	-0.64	0.16	0.18	-0.62	0.17
	P3	0.00	-0.64	0.16	0.05	-0.67	0.19
μ **[mV]**		36.00	36.00	36.00	35.04	35.98	35.41
τ_m **[ms]**		20.00	20.00	20.00	20.62	20.64	20.84
u_{th} **[mV]**		15.00	15.00	15.00	14.98	15.03	14.97
J_θ **[mV·ms]**		100.00	100.00	100.00	112.60	96.41	108.94
τ_θ **[ms]**		1000.00	1000.00	1000.00	1074.90	1037.52	1144.01
Normalized MSE loss				0.0051			

Explanation of the Weight Estimation

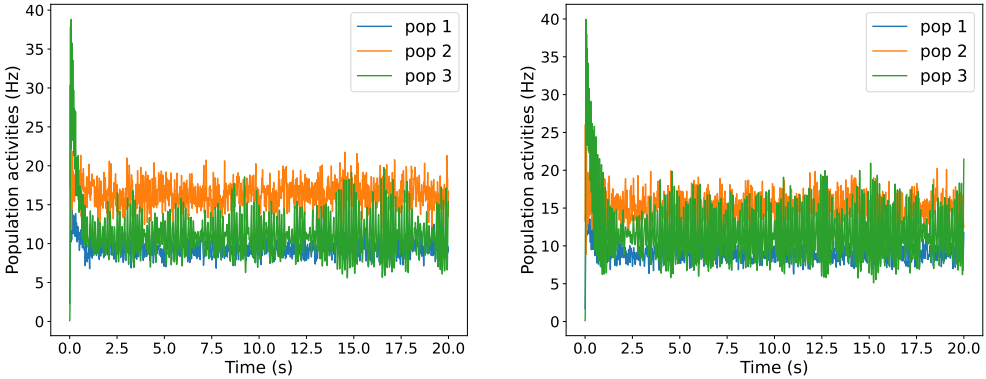
- Considering that minor variations in J_{syn} can lead to significant differences in neuronal activities, we utilize the J_{syn} from the label and other parameters from estimation to plot activities for estimation.
- This does not render the prediction of J_{syn} irrelevant. We can still employ the columns of J_{syn} to discern whether the populations are excitatory or inhibitory.
- We could establish a threshold for elements in J_{syn}; if an element in J_{syn} exceeds the threshold, it signifies a connection between the corresponding two populations, providing a raw connectivity matrix composed of 0's and 1's.
- Further theoretical insights and engineering techniques need exploration to either enhance the estimation of J_{syn} or to prove its inherent complexity.

Neuron Activities Comparison (3-Population, Sample 1)



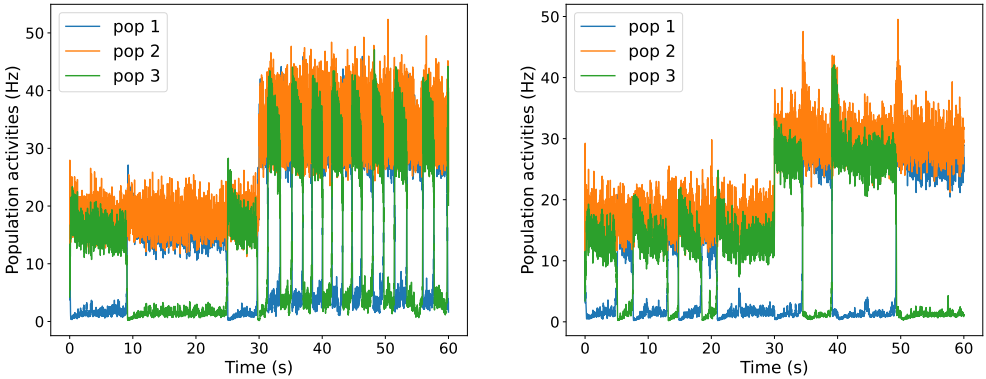
The left figure displays the label result, and the right figure displays the estimation result.

Neuron Activities Comparison (3-Population, Sample 2)



The left figure displays the label result, and the right figure displays the estimation result.

Neuron Activities Comparison (3-Population, Sample 3)



The left figure displays the label result, and the right figure displays the estimation result.

Numerical Comparison (8-Population, Sample 1 Part 1)

		Labels							
		P1	P2	P3	P4	P5	P6	P7	P8
J_{syn} **[mV]**	P1	-0.66	0.20	0.18	0.00	0.00	-0.58	0.00	0.00
	P2	-0.40	0.15	0.00	0.00	0.17	0.00	0.13	0.00
	P3	-0.43	0.17	0.18	-0.41	0.14	-0.71	0.00	0.00
	P4	-0.43	0.22	0.00	-0.87	0.00	0.00	0.00	0.21
	P5	-0.92	0.00	0.19	0.00	0.17	-1.03	0.00	0.09
	P6	-0.49	0.00	0.12	-1.11	0.00	-0.90	0.12	0.16
	P7	-0.95	0.00	0.00	-0.75	0.15	-0.40	0.11	0.22
	P8	0.00	0.00	0.00	0.00	0.13	-0.57	0.00	0.11

Numerical Comparison (8-Population, Sample 1 Part 2)

		Estimations							
		P1	P2	P3	P4	P5	P6	P7	P8
J_{syn} **[mV]**	P1	-0.84	0.14	-0.08	-0.51	0.29	-0.45	0.01	0.21
	P2	-0.52	-0.25	-0.14	-0.48	-0.04	-0.03	0.01	0.57
	P3	-0.36	-0.05	0.03	-0.51	-0.16	-0.79	-0.08	0.03
	P4	-0.34	0.19	-0.13	-1.26	-0.21	-0.45	-0.17	0.04
	P5	-1.03	-0.39	-0.07	-0.48	0.03	-0.66	-0.12	0.30
	P6	-0.40	-0.43	0.09	-1.40	0.05	-0.92	0.10	-0.06
	P7	-1.02	-0.31	0.01	-0.82	0.19	-0.20	-0.02	0.46
	P8	-0.23	-0.04	0.04	-0.14	0.20	-0.31	0.05	0.38

Numerical Comparison (8-Population, Sample 1 Part 3)

	Labels							
	P1	P2	P3	P4	P5	P6	P7	P8
μ **[mV]**	51.52	54.30	26.38	20.55	32.86	53.42	22.34	42.00
τ_m **[ms]**	30.92	19.33	15.19	28.60	23.90	31.52	23.31	16.49
u_{th} **[mV]**	13.61	24.40	28.83	12.28	19.78	10.91	24.40	18.69
J_θ **[mV·ms]**	982.48	43.80	43.80	982.48	43.80	982.48	43.80	43.80
τ_θ **[ms]**	1080.22	577.58	577.58	1080.22	577.58	1080.22	577.58	577.58

Numerical Comparison (8-Population, Sample 1 Part 4)

	Estimations							
	P1	P2	P3	P4	P5	P6	P7	P8
μ **[mV]**	51.41	50.04	33.05	23.89	28.71	56.63	19.23	42.77
τ_m **[ms]**	30.87	15.79	20.34	31.98	25.12	31.58	19.62	20.35
u_{th} **[mV]**	12.61	25.36	29.32	13.90	19.09	11.99	21.24	23.26
J_θ **[mV·ms]**	872.58	180.64	70.87	1204.65	133.53	1017.42	22.94	87.04
τ_θ **[ms]**	998.88	539.75	485.23	1209.86	723.77	1187.07	785.79	828.82
Normalized MSE loss	0.0440							
Normalized MSE loss without J_{syn}	0.0303							

Numerical Comparison (8-Population, Sample 2 Part 1)

		Labels							
		P1	P2	P3	P4	P5	P6	P7	P8
J_{syn} **[mV]**	P1	0.05	0.06	0.06	0.00	0.00	-0.32	0.00	0.11
	P2	0.00	0.11	0.12	0.00	0.00	0.00	0.09	0.07
	P3	0.00	0.08	0.08	-0.35	0.11	0.00	0.06	0.10
	P4	0.06	0.08	0.00	-0.45	0.09	0.00	0.00	0.00
	P5	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00
	P6	0.12	0.00	0.00	-0.46	0.00	-0.46	0.09	0.00
	P7	0.07	0.00	0.06	-0.19	0.12	-0.37	0.12	0.04
	P8	0.11	0.08	0.10	0.00	0.06	0.00	0.00	0.07

Numerical Comparison (8-Population, Sample 2 Part 2)

		Estimations							
		P1	P2	P3	P4	P5	P6	P7	P8
J_{syn} **[mV]**	P1	-0.30	0.14	-0.05	-0.23	-0.05	-0.54	-0.13	-0.11
	P2	0.10	-0.12	-0.16	-0.10	-0.06	-0.63	-0.04	0.22
	P3	0.07	0.13	-0.29	-0.15	0.19	-0.16	-0.18	-0.15
	P4	-0.15	0.09	0.03	-0.46	0.02	-0.39	0.03	0.12
	P5	-0.03	0.06	-0.01	-0.16	-0.30	-0.23	0.08	0.03
	P6	-0.06	-0.14	-0.07	-0.18	0.01	-0.58	0.05	0.28
	P7	0.02	-0.22	-0.28	-0.14	0.02	-0.16	-0.15	-0.01
	P8	-0.09	-0.13	-0.04	-0.31	0.06	-0.32	-0.01	0.10

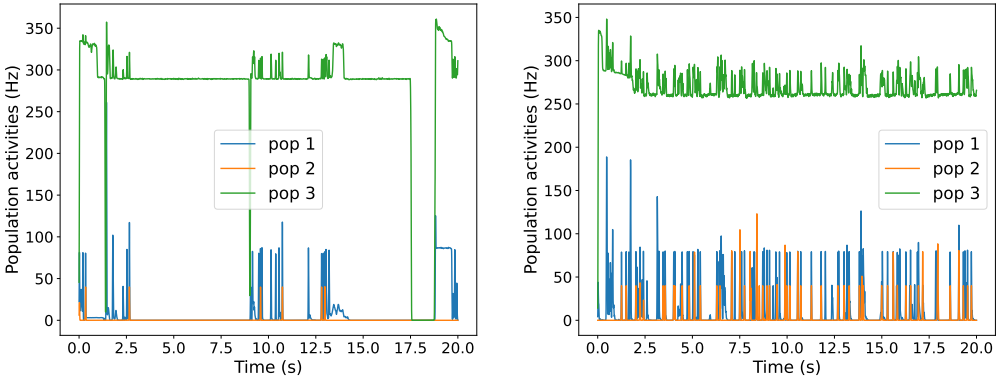
Numerical Comparison (8-Population, Sample 2 Part 3)

	Labels							
	P1	P2	P3	P4	P5	P6	P7	P8
μ **[mV]**	43.73	48.50	49.12	46.45	57.00	55.25	55.04	56.00
τ_m **[ms]**	35.51	17.44	24.96	29.30	26.67	11.86	24.74	27.86
u_{th} **[mV]**	19.90	12.80	24.18	26.43	21.54	18.83	20.23	18.44
J_θ **[mV·ms]**	1152.54	1152.54	1152.54	99.35	1152.54	99.35	1152.54	1152.54
τ_θ **[ms]**	1146.80	1146.80	1146.80	651.40	1146.80	651.40	1146.80	1146.80

Numerical Comparison (8-Population, Sample 2 Part 4)

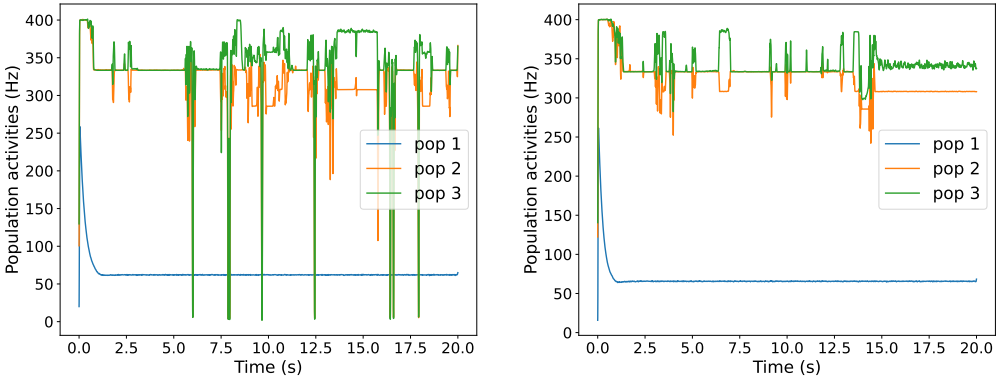
	Estimations							
	P1	P2	P3	P4	P5	P6	P7	P8
μ **[mV]**	49.85	47.78	48.46	42.75	48.12	55.19	54.77	58.31
τ_m **[ms]**	39.89	24.80	25.22	32.39	27.76	12.87	26.32	29.04
u_{th} **[mV]**	20.78	10.85	21.97	23.94	19.03	18.65	16.40	16.98
J_θ **[mV·ms]**	970.41	1270.93	764.44	473.48	1075.95	317.52	1180.89	1042.70
τ_θ **[ms]**	1311.60	1301.01	1140.90	845.82	926.21	815.45	1348.87	1116.72
Normalized MSE loss	0.0617							
Normalized MSE loss without J_{syn}	0.0518							

Neural Activities Comparison (8-Population, Sample 1)



3 populations are randomly selected from 8 populations for display. The left figure displays the label result, and the right figure displays the estimation result.

Neural Activities Comparison (8-Population, Sample 2)



3 populations are randomly selected from 8 populations for display. The left figure displays the label result, and the right figure displays the estimation result.

References I

[1] N. A. Steinmetz, C. Koch, K. D. Harris, and M. Carandini, Challenges and opportunities for large-scale electrophysiology with neuropixels probes, Current opinion in neurobiology, vol. 50, pp. 92-100, 2018.
[2] P. Gao and S. Ganguli, On simplicity and complexity in the brave new world of large-scale neuroscience, Current opinion in neurobiology, vol. 32, pp. 148-155, 2015.
[3] S. Wang, V. Schmutz, G. Bellec, and W. Gerstner, Mesoscopic modeling of hidden spiking neurons, arXiv preprint arXiv:2205.13493, 2022.
[4] T. C. Potjans and M. Diesmann, The cell-type specific cortical microcircuit: relating structure and activity in a full-scale spiking network model, Cerebral cortex, vol. 24, no. 3, pp. 785-806, 2014.
[5] H. Markram, E. Muller, S. Ramaswamy, M. W. Reimann, M. Abdellah, C. A. Sanchez, A. Ailamaki, L. Alonso-Nanclares, N. Antille, S. Arsever, et al., Reconstruction and simulation of neocortical microcircuitry, Cell, vol. 163, no. 2, pp. 456-492, 2015.
[6] E. M. Izhikevich and G. M. Edelman, Large-scale model of mammalian thalamocortical systems, Proceedings of the national academy of sciences, vol. 105, no. 9, pp. 3593-3598, 2008.
[7] T. Schwalger, M. Deger, and W. Gerstner, Towards a theory of cortical columns: From spiking neurons to interacting neural populations of finite size, PLoS computational biology, vol. 13, no. 4, p. e1005507, 2017.
[8] W. J. Freeman et al., Mass action in the nervous system, vol. 2004. Citeseer, 1975.
[9] O. David and K. J. Friston, A neural mass model for meg/eeg:: coupling and neuronal dynamics, NeuroImage, vol. 20, no. 3, pp. 1743-1755, 2003.
[10] R. Moran, D. A. Pinotsis, and K. Friston, Neural masses and fields in dynamic causal modeling, Frontiers in computational neuroscience, vol. 7, p. 57, 2013.

References II

[11] V. K. Jirsa and H. Haken, A derivation of a macroscopic field theory of the brain from the quasi-microscopic neural dynamics, Physica D: Nonlinear Phenomena, vol. 99, no. 4, pp. 503-526, 1997.
[12] W. Gerstner, W. M. Kistler, R. Naud, and L. Paninski, Neuronal dynamics: From single neurons to networks and models of cognition. Cambridge University Press, 2014.
[13] G. Deco, V. K. Jirsa, P. A. Robinson, M. Breakspear, and K. Friston, The dynamic brain: from spiking neurons to neural masses and cortical fields, PLoS computational biology, vol. 4, no. 8, p. e1000092, 272 2008.
[14] W. Gerstner, Population dynamics of spiking neurons: fast transients, asynchronous states, and locking, Neural computation, vol. 12, no. 1, pp. 43-89, 2000.
[15] N. Brunel and V. Hakim, Fast global oscillations in networks of integrate-and-fire neurons with low firing rates, Neural computation, vol. 11, no. 7, pp. 1621-1671, 1999.
[16] A. René, A. Longtin, and J. H. Macke, Inference of a mesoscopic population model from population spike trains, Neural computation, vol. 32, no. 8, pp. 1448-1498, 2020.
[17] W. Gerstner and R. Naud, How good are neuron models?, Science, vol. 326, no. 5951, pp. 379-380, 2009.
[18] S. Mensi, R. Naud, C. Pozzorini, M. Avermann, C. C. Petersen, and W. Gerstner, Parameter extraction and classification of three cortical neuron types reveals two distinct adaptation mechanisms, Journal of neurophysiology, vol. 107, no. 6, pp. 1756-1775, 2012.
[19] S. Hochreiter and J. Schmidhuber, Long short-term memory, Neural computation, vol. 9, no. 8, pp. 1735-1780, 1997.

References III

[20] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, Empirical evaluation of gated recurrent neural networks on sequence modeling, arXiv preprint arXiv:1412.3555, 2014.
[21] W. M. Kistler, W. Gerstner, and J. L. v. Hemmen, Reduction of the hodgkin-huxley equations to a single-variable threshold model, Neural computation, vol. 9, no. 5, pp. 1015-1045, 1997.
[22] R. Jolivet, T. J. Lewis, and W. Gerstner, Generalized integrate-and-fire models of neuronal activity approximate spike trains of a detailed model to a high degree of accuracy, Journal of neurophysiology, vol. 92, no. 2, pp. 959-976, 2004.
[23] R. Brette and W. Gerstner, Adaptive exponential integrate-and-fire model as an effective description of neuronal activity, Journal of neurophysiology, vol. 94, no. 5, pp. 3637-3642, 2005.
[24] W. Gerstner, Time structure of the activity in neural network models, Physical review E, vol. 51, no. 1, p. 738, 1995.
[25] T. Schwalger and A. V. Chizhov, Mind the last spikefiring rate models for mesoscopic populations of spiking neurons, Current opinion in neurobiology, vol. 58, pp. 155-166, 2019.
[26] H. R. Wilson and J. D. Cowan, Excitatory and inhibitory interactions in localized populations of model neurons, Biophysical journal, vol. 12, no. 1, pp. 1-24, 1972.
[27] R. Naud and W. Gerstner, Coding and decoding with adapting neurons: a population approach to the peri-stimulus time histogram, 2012.
[28] Sandrine Lefort, Christian Tomm, J-C Floyd Sarria, and Carl CH Petersen. The excitatory neuronal network of the c2 barrel column in mouse primary somatosensory cortex. Neuron, 61(2):301-316, 2009.

References IV

[29] Christian Pozzorini, Skander Mensi, Olivier Hagens, Richard Naud, Christof Koch, and Wulfram Gerstner. Automated high-throughput characterization of single neurons by means of simplified spiking models. PLoS computational biology, 11(6):e1004275, 2015.
[30] Scott J Cruikshank, Timothy J Lewis, and Barry W Connors. Synaptic basis for intense thalamocortical activation of feedforward inhibitory cells in neocortex. Nature neuroscience, 10(4):462-468, 2007.
[31] Mahesh M Karnani, Jesse Jackson, Inbal Ayzenshtat, Jason Tucciarone, Kasra Manoocheri, William G Snider, and Rafael Yuste. Cooperative subnetworks of molecularly similar interneurons in mouse neocortex. Neuron, 90(1):86-100, 2016.
[32] Anita Tusche, Anne Böckler, Philipp Kanske, Fynn-Mathis Trautwein, and Tania Singer. Decoding the charitable brain: empathy, perspective taking, and attention shifts differentially predict altruistic giving. Journal of Neuroscience, 36(17):4719-4732, 2016.
[33] Carsten K Pfeffer, Mingshan Xue, Miao He, Z Josh Huang, and Massimo Scanziani. Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. Nature neuroscience, 16(8):1068-1076, 2013.
[34] Adam M Packer and Rafael Yuste. Dense, unspecific connectivity of neocortical parvalbumin-positive interneurons: a canonical microcircuit for inhibition? Journal of Neuroscience, 31(37):13260-13271, 2011.
[35] Luc J Gentet, Yves Kremer, Hiroki Taniguchi, Z Josh Huang, Jochen F Staiger, and Carl CH Petersen. Unique functional properties of somatostatin-expressing gabaergic neurons in mouse barrel cortex. Nature neuroscience, 15(4):607-612, 2012.

References V

[36] Michael Avermann, Christian Tomm, Celine Mateo, Wulfram Gerstner, and Carl CH Petersen. Microcircuits of excitatory and inhibitory neurons in layer 2/3 of mouse barrel cortex. Journal of neurophysiology, 107 (11):3116-3134, 2012.
[37] Tilo Schwalger, Moritz Deger, and Wulfram Gerstner. Towards a theory of cortical columns: From spiking neurons to interacting neural populations of finite size. PLoS computational biology, 13(4):e1005507, 2017.
[38] Y Liu. H, wang xj. Spike-frequency adaptation of a generalized leaky integrate-and-fire model neuron. J Comput Neurosci, 10(1):25-45, 2001.
[39] Maurice J Chacron, André Longtin, Martin St-Hilaire, and Len Maler. Suprathreshold stochastic firing dynamics with memory in p-type electroreceptors. Physical Review Letters, 85(7):1576, 2000.
[40] C Daniel Geisler and Jay M Goldberg. A stochastic model of the repetitive activity of neurons. Biophysical journal, 6(1):53-69, 1966.
[41] Christian Pozzorini, Richard Naud, Skander Mensi, and Wulfram Gerstner. Temporal whitening by power-law adaptation in neocortical neurons. Nature neuroscience, 16(7):942-948, 2013.
[42] Moritz Deger, Tilo Schwalger, Richard Naud, and Wulfram Gerstner. Fluctuations and information filtering in coupled populations of spiking neurons with adaptation. Physical Review E, 90(6):062704, 2014.
[43] Taro Toyoizumi, Kamiar Rahnama Rad, and Liam Paninski. Mean-field approximations for coupled populations of generalized linear model spiking neurons with markov refractoriness. Neural computation, 21(5):1203-1243, 2009.

References VI

[44] Alison I Weber and Jonathan W Pillow. Capturing the dynamical repertoire of single neurons with generalized linear models. Neural computation, 29(12):3260-3289, 2017.

[45] Wilson Truccolo, Uri T Eden, Matthew R Fellows, John P Donoghue, and Emery N Brown. A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. Journal of neurophysiology, 93(2):1074-1089, 2005.

[46] Jonathan W Pillow, Jonathon Shlens, Liam Paninski, Alexander Sher, Alan M Litke, EJ Chichilnisky, and Eero P Simoncelli. Spatio-temporal correlations and visual signalling in a complete neuronal population. Nature, 454(7207):995-999, 2008.

[47] Valentin Schmutz, Eva Löcherbach, and Tilo Schwalger. On a finite-size neuronal population equation.

arXiv preprint arXiv:2106.14721, 2021.

Thanks!