Optimal spike coding

Jeroen Olieslagers, David Liu

February 23, 2021

Computational and Biological Learning Lab, Cambridge UK

Introduction: representing continuous signals with spikes

Universally agreed that neural firing rates communicate information
[Denève and Machens, 2016]

[Boerlin and Denève, 2011]

[Bourdoukan et al., 2012]

Outline

Learning optimal spike-based representations

Bourdoukan, Ralph, et al.; 2012, *Advances in neural information processing systems*

Spike-based population coding and working memory

Boerlin, Martin, and Sophie Denève; 2011, *PLoS Comput Biol*

Spatio-temporal Representations of Uncertainty in Spiking Neural Networks

Savin, Cristina, and Sophie Deneve; 2014, *Advances in neural information processing systems*

Learning to represent signals spike by spike

Brendel W, Bourdoukan R, Vertechi P, Machens CK, Denéve S.; 2020, *Advances in neural information processing systems*

Learning optimal spike-based representations

Representing continuous task variables

Consider the task of representing a continuous signal $x(t)$ with spikes

Similar to autoencoder framework $x \to o \to \hat{x}$, o is the output spike train(s)

Learning and information representation is usually treated separately, here it arises from the same mechanism

Experimental observations

- high trial-to-trial variability in neurons
- asynchronous firing at the network level (spike trains of different neurons weakly correlated)
- tight E-I balance (excitatory input met by inhibitory input of equal or greater size)
- spike-timing-dependent plasticity (strength of synapses change as function of presynaptic and postsynaptic spike times)

Single neuron mechanism

Decoder $\frac{d}{dt}\hat{x} = -\hat{x}(t) + \Gamma o(t)$ (low pass filter of $\delta(t)$ peaks spike train)

Autoencoding objective $L(t) = (x(t) - \hat{x}(t))^2$

Here comes optimality: *neurons only spike when it decreases the loss*

$L(t|\text{spike}) = (x(t) - (\hat{x}(t) + \Gamma))^2$

$L(t|\text{no spike}) = (x(t) - \hat{x}(t))^2$

Spiking condition

$$L(t|\text{spike}) > L(t|\text{no spike}) \tag{1}$$

Single neuron mechanism

Expanding squares, assigning $V(t) = \Gamma(x(t) - \hat{x}(t))$ and $T = \frac{\Gamma^2}{2}$

Spiking condition $(1) \Rightarrow V(t) > T$

Connect this normative approach to mechanistic interpretation: $V(t)$ is membrane potential, T is the (hard) spiking threshold

Single neuron dynamics

What is the dynamics of $V(t)$?

$\dot{V} = -\Gamma(x - \hat{x}) + \Gamma(\dot{x} + x) - \Gamma^2 o$, define $c = \dot{x} + x$ and interpret as the external input

$$\dot{V} = -V + \Gamma c - \Gamma^2 o \tag{2}$$

leaky integrate-and-fire (LIF) neuron with inhibitory autapse and feedforward weight Γ

Can interpret autapse as as spiking reset from $V \approx T$ to $V = -T$ (since $\Gamma^2 = 2T$

We have recovered the popular mechanistic LIF model!

Single neuron learning

So far no learning, but we have the optimal target

$\dot{V} = -V + \Gamma c - \omega o$, target $\omega \to \Gamma^2$

$\dot{V} + V = \Gamma(\dot{x} + x) - \omega(\dot{\bar{o}} + \bar{o})$, use $\dot{\bar{o}} = -\bar{o} + o$, integrate equation

$V = \Gamma x + \omega o$

Write loss $L = (x - \hat{x})^2 = \frac{1}{\Gamma^2}(V^2 + 2(\omega - \Gamma^2)\bar{o} + (\omega - \Gamma^2)^2\bar{o}^2)$

Note that at the optimal solution the last two terms vanish, thus

$$L_V = \frac{1}{2}V^2 \tag{3}$$

has the same minimum according to the authors

Note sure if true in rigorous sense, but note $V = \Gamma(x - \hat{x})$ and
$L = (x - \hat{x})^2$

Single neuron learning

Gradient of loss $\frac{\partial L_V}{\partial \omega} = V \frac{\partial V}{\partial \omega} = -V\bar{o} - V\omega \frac{\partial \bar{o}}{\partial \omega}$

Ignoring the second term gives online Hebbian-like synaptic plasticity from gradient descent

$$\tau \dot{\omega} = V\bar{o} \tag{4}$$

System of coupled ODEs in V, $\bar{o}$ and ω can be solved analytically between spikes, gives a recurrence relation

$$\omega_{i+1} = \omega_i + \frac{\omega_i(\omega_i - 2T)}{\tau(T - \Gamma c - \omega_i)} \tag{5}$$

which has a single stable fixed point at $\omega = 2T = \Gamma^2$, justifies ignoring second term

Now we introduce a population of identical neurons

$$\dot{\boldsymbol{V}} = -\boldsymbol{V} + \boldsymbol{\Gamma}c - (\boldsymbol{\Gamma}\boldsymbol{\Gamma}^T + \mu I)\boldsymbol{o} \qquad (6)$$

with as before $\dot{\hat{x}} = -\hat{x} + \boldsymbol{\Gamma} \cdot \boldsymbol{o}$

Loss with metabolic cost $L = (x - \hat{x})^2 + \mu||\bar{\boldsymbol{o}}||^2$

Weight scale with $\sim \frac{1}{N}$ so firing rates are $O(1)$

Spiking condition (1) leads to dynamics

$$\dot{\boldsymbol{V}} = -\boldsymbol{V} + \boldsymbol{\Gamma} c - (\boldsymbol{\Gamma}\boldsymbol{\Gamma}^T + \mu I)\boldsymbol{o} \tag{7}$$

with spiking threshold $V_i > T_i = \Gamma_i^2/2 + \mu/2$ with $V_i = \Gamma_i(x - \hat{x}) - \mu\bar{o}$

Take identical $\Gamma_i = \Gamma$

Network of LIF with homogeneous inhibitory connectivity
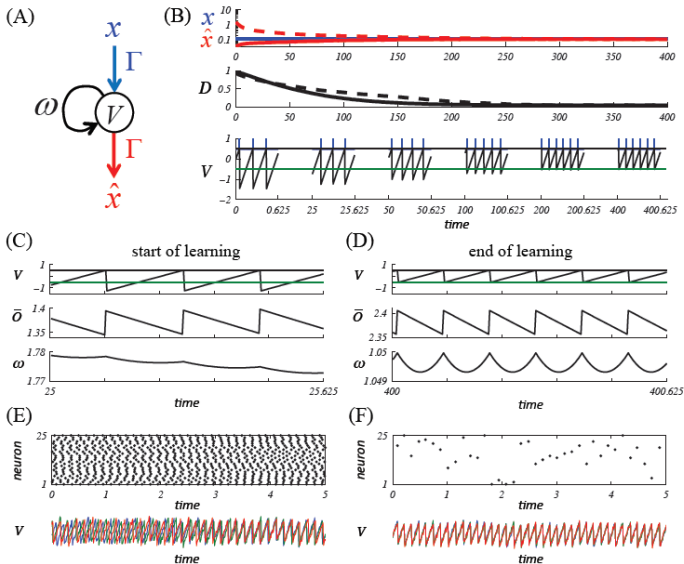
As before, $L = ||\boldsymbol{V}||^2/2$

Gradient $\frac{\partial L_V}{\partial \Omega_{ij}} = -\sum_k V_k \delta_{ki} \bar{o}_j - \sum_{kl} V_k \Omega_{kl} \frac{\partial \bar{o}_l}{\partial \Omega_{ij}}$

Ignoring the second term gives Hebbian-like synaptic plasticity again

$$\tau \dot{\Omega}_{ij} = V_i \bar{o}_j \tag{8}$$

Justified as first term $O(\frac{1}{N})$ because $\Gamma \sim \frac{1}{N}$, second term $\ll O(\frac{1}{N})$ when terms $\Omega_{kl} \frac{\partial \bar{o}_l}{\partial \Omega_{ij}}$ weakly correlated with zero mean at end of learning

(A)

(B)

(C) start of learning

(D) end of learning

(E)

(F)

Multidimensional signals

Multidimensional input

$$\dot{\boldsymbol{V}} = -\boldsymbol{V} + \Gamma^T \boldsymbol{c} - \Omega \boldsymbol{o} \tag{9}$$

with the neuron output kernels as the rows of Γ

The input $\boldsymbol{c} = \dot{\boldsymbol{x}} + \boldsymbol{x}$ can be generalized as control input for linear dynamical system

Optimal weights $\Omega = \Gamma^T \Gamma + \mu I$, learning with same plasticity rule

Note we can write $\hat{\boldsymbol{x}} = \Gamma \bar{\boldsymbol{o}}$

(A)

(B)

(C) start of learning

(D) end of learning

(E)

(F)

17
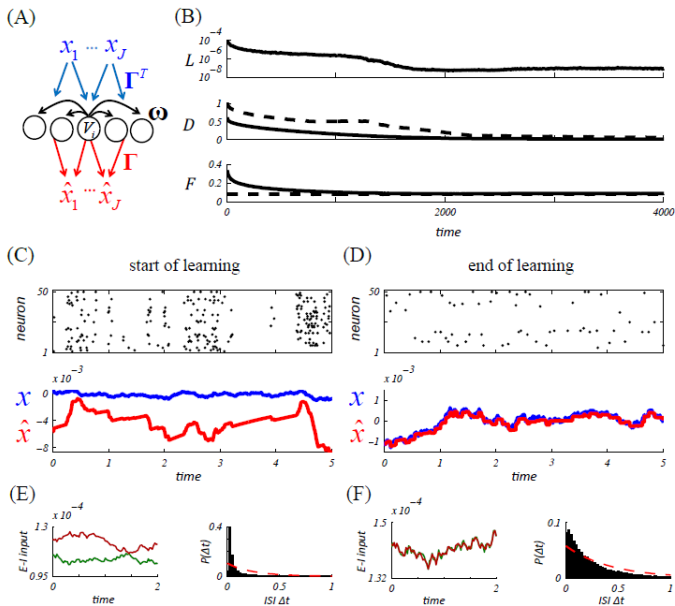
Optimal weights never stabilize, but oscillate around optimal value as long as neurons keep spiking

Recurrent weights Ω and decoding weights Γ are linked in optimal case

- Neural representations drift within coding subspace
  [Rule et al., 2020]: for the brain to decode itself, any downstream 'decoder' must continuously adapt
- We also know synapses are dynamic, in particular excitatory synapses

This framework is consistent with these observations

The spiking threshold T is a function of the decoding weights, adaptation related to changing readout weights?

Probabilistic spiking condition if interpret $p(o(t)) \propto e^{-\alpha L[o(t)]}$

$\frac{p(\text{spike at } t)}{p(\text{no spike at } t)} = e^{\alpha\left(2\Gamma(x(t)-\hat{x}(t))-\Gamma^2\right)}$

$\alpha \to \infty$ gives deterministic spiking as in [Bourdoukan et al., 2012]

Discussion

Note that learning only applicable to single-layer networks here

Separate topic but exciting developments in optimizing recurrent SNNs [Huh and Sejnowski, 2017, Bellec et al., 2018], online learning [Bellec et al., 2020]

Not biologically plausible, but shows similarities to STDP

Neural adaptation increases computational capacity

Recent advances have pushed performance close to ANNs

Spike-based population coding and working memory

Probabilistic computations with spikes

This work essentially combines PPCs [Ma et al., 2006] with optimal spike coding

Instead of tracking an arbitrary signal $x(t)$, now the system tracks the log posterior evaluated a discrete stimulus points

Published before the previous paper (which introduced learning), here they show optimal spike coding is consistent with cortical dynamics

Connection weights here are determined by optimal spiking as before, i.e. from the decoding weights Γ

Probabilistic population codes

Treat spiking as stochastic, in particular distribution from exponential family

$$\log p(\boldsymbol{S}_t | x_t) = \sum_j H_j(x_t) S_{tj} + \phi(x_t) + \Phi(\boldsymbol{S}_t) \qquad (10)$$

here $\boldsymbol{S}_t$ are the spike counts of population in $[t, t + \mathrm{d}t]$

$\boldsymbol{H}(x_t) = \log \boldsymbol{f}(x_t)$ kernel related to log tuning curves, $\phi(x_t)$ bias

Cue combination of 'auditory' and 'visual' input; just names given to independent groups of $\boldsymbol{S}$, $\boldsymbol{H}$, ... of the same tracked x_t

Assume Poisson spiking process for sensory input

Task

- Tracking of stimulus x_t given noisy sensory input
- Maintaining the estimate when input is absent using prior knowledge of stimulus dynamics

Drift-diffusion process

$$\mathrm{d}x_t = \delta\,\mathrm{d}t + \sigma\,\mathrm{d}W_t \tag{11}$$

Analytical solution of $p(x_t, t)$ is Fokker-Planck equation, equivalently for $l = \log p(x_t, t$

$$\partial_t l(x_t, t) = -\delta\partial_x l(x_t, t) + \frac{\sigma^2}{2}[\partial_{xx} l(x_t, t) + (\partial_x l(x_t, t))^2] \tag{12}$$

Ideal observer model

Given noisy sensory input from spike counts $\boldsymbol{S}_t^a$, $\boldsymbol{S}_t^v$, want to infer posterior $p(x_t|\boldsymbol{S}_{[0,t]}^a, \boldsymbol{S}_{[0,t]}^v)$

Denote log posterior $l(x_t, t)$, prior evolution from Fokker-Planck (12) combined with likelihood terms

$$\partial_t l(x_t, t) = \sum_j H_j^a(x_t) S_{tj}^a + \phi^a(x_t) + \sum_j H_j^v(x_t) S_{tj}^v + \phi^v(x_t) + u(x_t, t) \tag{13}$$

with $u(x_t, t)$ the prior evolution equation

$$u(x_t, t) = -\delta \partial_x l(x_t, t) + \frac{\sigma^2}{2}[\partial_{xx} l(x_t, t) + (\partial_x l(x_t, t))^2] \tag{14}$$

Neural implementations

Line of research dealing with 'neural' implementations of statistical inference [Pouget and Latham, 1999, Beck and Pouget, 2007, Deneve et al., 2007, Sun et al., 2018]

Typically feedforward noisy input neurons to recurrent line attractors which performs approximate cue combination and ML/MAP estimation

Here however we use sensory integration through optimal spiking dynamics instead of a recurrent attractor

Optimal spike coding

Decoder from recurrent population spikes $\boldsymbol{O}(t)$

$$\partial_t g(x_t, t) = -\lambda g(x_t, t) + \sum_j \Gamma_j(x_t) O_j(t) \tag{15}$$

with recurrent dynamics such that $g(x_t, t) \approx \log p(x_t | \boldsymbol{S}^a_{[0,t]}, \boldsymbol{S}^a_{[0,t]})$

Discretize stimulus space $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$, denote $L_i(t) = l(x_i, t)$, $G_i(t)$ similar, $\Gamma_{ij} = \Gamma_j(x_i)$

Spiking condition (only fire when new information)

$$\sum_j (L_j(t) - G_j(t))^2 > \sum_j (L_j(t) - (G_j(t) + \Gamma_{ji})^2) \tag{16}$$

Optimal spike coding

The squared distance error (opposed to e.g. KL divergence) leads to local integrate-and-fire neuron dynamics

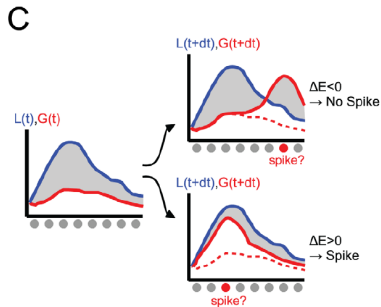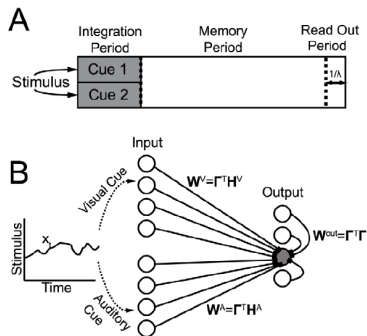Membrane potential $V(t) = \sum_j \Gamma_{ji}(L_j(t) - G_j(t))$

From the decoder and log posterior time evolution

$$\partial_t V_i = -\lambda V_i + \sum_{j,m \in \{a,v\}} W_{ij}^m S_j^m(t) - \sum_{j \neq i} W_{ij}^{out} O_j(t) + U_i(\boldsymbol{O}, t) - \phi_i \quad (17)$$

with feedforward $W^{a/v} = \Gamma^T H^{a/v}$ and recurrent $W^{out} = \Gamma^T \Gamma$

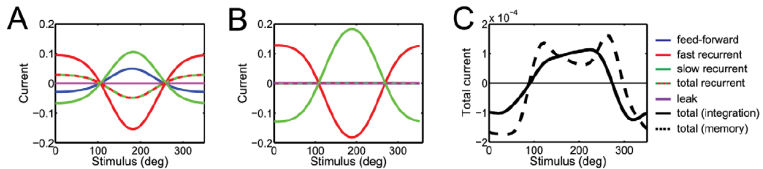Neurons fire when $V_i > T_i$, with $T_i = \sum_j \Gamma_{ji}^2/2$, reset to $-T_i$

$U_i(\boldsymbol{O}, t) = Y_i(t) + \Sigma_j \Gamma_{ij}^T Z_j(t)^2$ represents slow currents

$$\dot{Y}_i(t) = -\lambda Y_i(t) + \Sigma_j \Omega_{ij} O_j(t)$$
$$\dot{Z}_i(t) = -\lambda Z_i(t) + \Sigma_j K_{ij} O_j(t)$$

(18)

with $\Omega = \Gamma^T(\lambda \Gamma - \delta \Gamma' + \frac{\sigma^2}{2} \Gamma'')$ and $K = \frac{\sigma}{\sqrt{2}} \Gamma'$

Remember Γ is discrete stimulus points $I(x_t, t)$, Γ' derivate w.r.t. x

Interpretation

Feed-forward inputs transmit new sensory evidence about a stimulus to the output neurons

Feed-forward currents positive/negative for neurons with preferred stimuli similar/different from presented stimulus

Fast recurrent inputs subtract the output population's prediction from this sensory input and hence have opposite signs

Short-range fast inhibition and long-range fast excitation let one output neuron transmit information at a time

Interpretation

Slow recurrent connections, on the other hand, have two distinct roles.

- reintroduce information that has leaked out, hence making past information available to enable sustained bumps of activity in output layer (working memory)
- The second role of the slow currents is to take into account the non-stationary dynamics of the stimulus

Spike generation is a competition between (1) sensory integration from feed-forward and slow lateral inputs, (2) prediction from fast lateral connections

Important consequence is maintenance of tight balance between excitatory and inhibitory currents in output neurons

Diffusion

Nonlocal nonlinear term in $U_i(\boldsymbol{O}, t)$ when diffusion is $\sigma > 0$

$$\sum_j \Gamma_{ij}^T Z_j(t)^2 \tag{19}$$

May be implemented by nonlinear dendritic trees of output neurons

Interpretation

Free parameters leak λ and output kernel Γ

All other parameters are functions of λ, Γ, the stimulus dynamics or the input response tuning curves

Kernel Γ determines stimulus space meaning of a spike
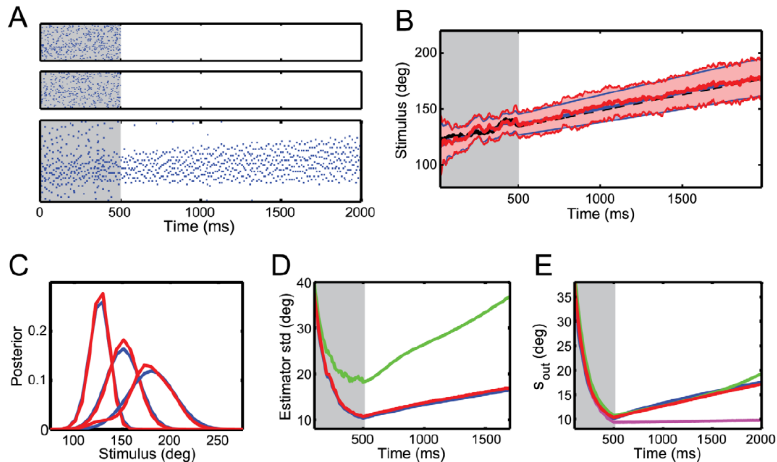
Leak λ determines temporal relevance of spike

Output kernels do not need to be known in advance by the decoder/downstream processing, can be estimated directly from the tuning curves
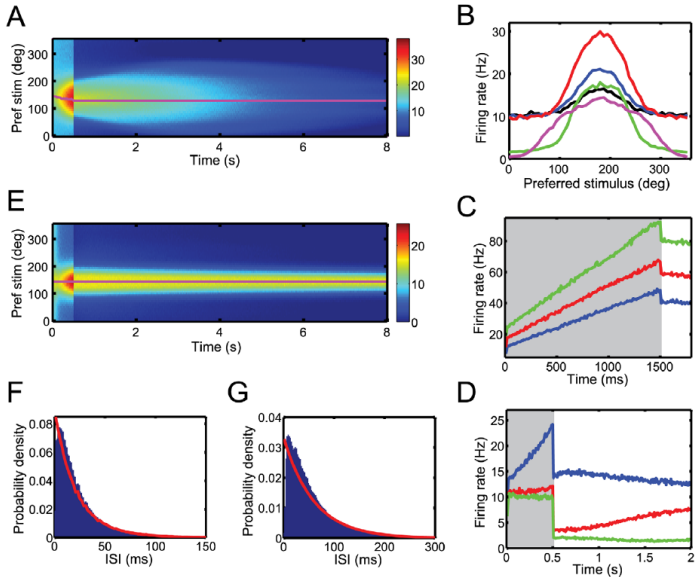
$$\mathbf{\Gamma}'(x_t) = \Sigma^{-1} \mathbf{f}^{out}(x_t)' \tag{20}$$

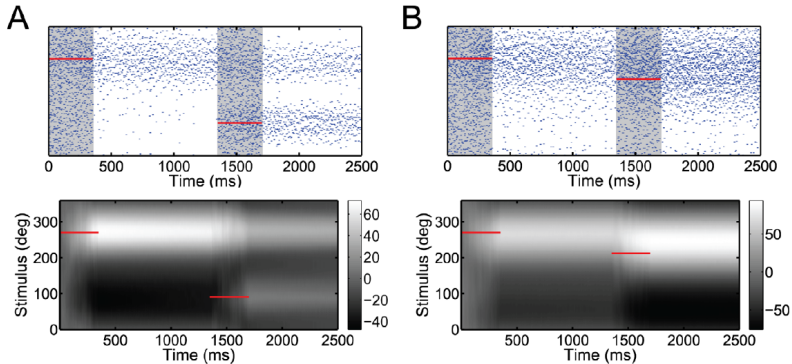(this relation holds for spiking processes in the exponential family)
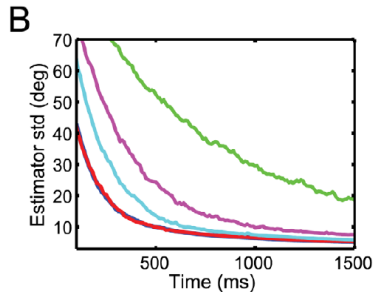
Simulations use this learned kernel, good approximation

Statistics

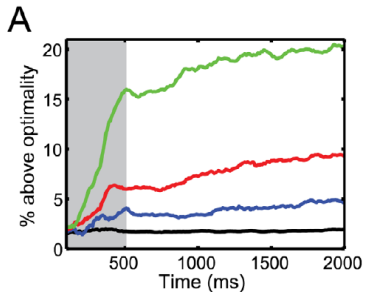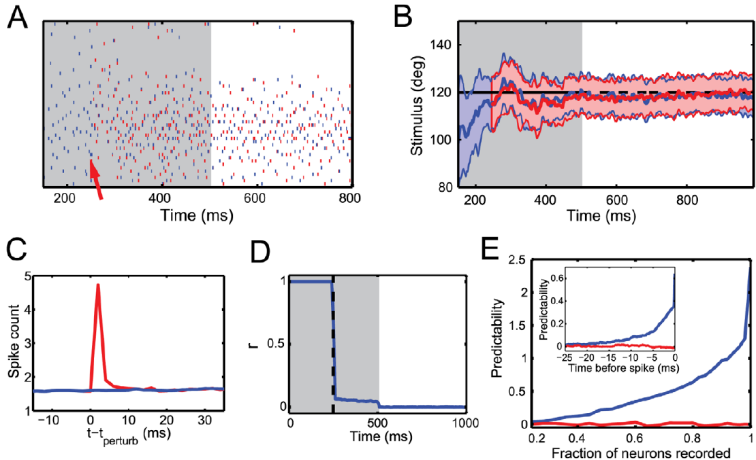Trial-to-trial variability (with initial Poisson input)

Both memory and integration periods have Fano factors ~ 1.4

ISI distribution quasi-exponential

Cross correlations coefficients $< 10^{-3}$

When 'recording' from subset of neurons and looking at peer-predictability, approaches independent Poisson rate code for $< 25\%$

Discussion

In vivo rat whisker barrel perturbation experiment similar to here, concluded cortex mainly uses a rate code [London et al., 2010]

Qualitatively identical result here, but not a rate code?

Alternative account for the origin of neural variability in cortex to stochastic firing

Literature suggests variability (partly) arises from chaotic dynamics of spiking recurrent networks with balanced excitation and inhibition

These dynamics here cannot be equated to noise

Can conserve and transmit information perfectly

The network is self-correcting and robust to types of noise reported in cortical neurons, e.g. spike generation noise or synaptic noise

Discussion

Authors propose following experiments to test between stochastic firing and their model

- The larger the simultaneously recorded population, the better one can predict individual spike times based on peer spike trains as opposed to Poisson rate networks

- On the behavioral level, humans should be able to memorize entire probability distributions. Simple cue combination experiment with temporal delay in presenting second cue

Related work

Spiking Boltzmann machine [Hinton and Brown, 1999]

Log probability codes and message passing with spikes [Rao, 2005]

Probabilistic population codes [Ma et al., 2006] is optimal for rate coding under Poisson noise

Fast population coding [Huys et al., 2007] uses Gaussian process prior of stimulus dynamics instead of Markovian, nonlocal due to KL divergence

Summary

With a spiking criterion of *maximally informative spikes* with respect to some loss criterion, one can derive mechanistic integrate-and-fire models with *interpretable components*

- Membrane potentials are proportional to prediction error
- Spiking threshold determined by spike 'meaning', i.e. decoding weights
- Learning can be achieved with Hebbian-like plasticity
- Resulting dynamics are asynchronous and show E-I balance, consistent with cortical recordings
- Deterministic dynamics still lead to variability due to chaotic dynamics

Unlike stochastic rate codes, each spike here 'counts'

Decoder is supposed to be abstraction of downstream processing

Beck, J. M. and Pouget, A. (2007).
Exact inferences in a neural implementation of a hidden markov model.
Neural computation, 19(5):1344–1361.

Bellec, G., Salaj, D., Subramoney, A., Legenstein, R., and Maass, W. (2018).
Long short-term memory and learning-to-learn in networks of spiking neurons.
arXiv preprint arXiv:1803.09574.

Bellec, G., Scherr, F., Subramoney, A., Hajek, E., Salaj, D., Legenstein, R., and Maass, W. (2020).
A solution to the learning dilemma for recurrent networks of spiking neurons.
Nature communications, 11(1):1–15.

Boerlin, M. and Denève, S. (2011).
Spike-based population coding and working memory.
PLoS computational biology, 7(2).

Bourdoukan, R., Barrett, D., Deneve, S., and Machens, C. K. (2012).
Learning optimal spike-based representations.
Advances in neural information processing systems, 25:2285–2293.

Deneve, S., Duhamel, J.-R., and Pouget, A. (2007).
Optimal sensorimotor integration in recurrent cortical networks: a neural implementation of kalman filters.
Journal of Neuroscience, 27(21):5744–5756.

Denève, S. and Machens, C. K. (2016).
Efficient codes and balanced networks.
Nature neuroscience, 19(3):375–382.

Hinton, G. E. and Brown, A. D. (1999).
Spiking boltzmann machines.
In *NIPS*, pages 122–128. Citeseer.

Huh, D. and Sejnowski, T. J. (2017).
Gradient descent for spiking neural networks.
arXiv preprint arXiv:1706.04698.

Huys, Q. J., Zemel, R. S., Natarajan, R., and Dayan, P. (2007).
Fast population coding.
Neural computation, 19(2):404–441.

London, M., Roth, A., Beeren, L., Häusser, M., and Latham, P. E. (2010).
Sensitivity to perturbations in vivo implies high noise and suggests rate coding in cortex.
Nature, 466(7302):123–127.

Ma, W. J., Beck, J. M., Latham, P. E., and Pouget, A. (2006).
Bayesian inference with probabilistic population codes.
Nature neuroscience, 9(11):1432–1438.

Pouget, S. D. A. and Latham, P. (1999).
Divisive normalization, line attractor networks and ideal observers.
In *Advances in Neural Information Processing Systems 11: Proceedings of the 1998 Conference*, volume 11, page 104. MIT Press.

Rao, R. P. (2005).
Hierarchical bayesian inference in networks of spiking neurons.
In *Advances in neural information processing systems*, pages 1113–1120.

Rule, M. E., Loback, A. R., Raman, D. V., Driscoll, L. N., Harvey, C. D., and O'Leary, T. (2020).
Stable task information from an unstable neural population.
Elife, 9:e51121.

Sun, X., Mangan, M., and Yue, S. (2018).
An analysis of a ring attractor model for cue integration.
In *Conference on Biomimetic and Biohybrid Systems*, pages 459–470. Springer.