The Computational Structure of Spike Trains

Robert Heinz Haslinger[1,2] Cosma Rohilla Shalizi[3]

[1] Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown MA [2] Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge MA
[3] Department of Statistics, Carnegie Mellon University, Pittsburgh, PA

INTRODUCTION

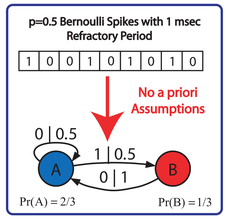
Neurons are computational devices, transforming synaptic input (hard to measure) into output spikes (routinely measured).

We quantify the computational structure of a spike train in the form of an optimally predictive hidden Markov model or *Causal State Model* solely though analysis of the output spikes.

In an experiment we measure spikes generated by some unknown dynamical process.

CSMs represent this process as a directed graph in which the nodes specify the dynamical systems states {S}.

The directed edges correspond to possible transitions between the states and are labled with the symbol (spike or no spike) emitted during the transition.



The Causal State Splitting Reconstruction (CSSR) algorithm is a non-parametric method for generating CSMs. No assumptions about the output structure are made.

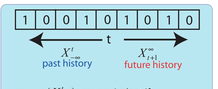
Using CSM's we objectively quantify the number of bits needed to describe both the complexity (structure) and randomness of both simulated and experimentally recorded spike trains.

WHAT ARE CAUSAL STATES?

A dynamical system's "state" consists of the information relevant for predicting its future evolution.

Two past histories "x" and "y" are equivalent when they have the same conditional distribution of future histories.

These equivalence classes are the *Causal States*.



$x, y \in \{X_{-\infty}^{t}\}$ are equivalent if

$$\Pr(X_{t+1}^{\infty} \mid X_{-\infty}^{t} = x) = \Pr(X_{t+1}^{\infty} \mid X_{-\infty}^{t} = y)$$

$$S_t = \epsilon(X_{-\infty}^{t} = x) = \{y : \Pr(X_{t+1}^{\infty} \mid X_{-\infty}^{t} = y) = \Pr(X_{t+1}^{\infty} \mid X_{-\infty}^{t} = x)\}$$

Causal States are Minimal and Predictively Sufficient

Causal states maximize the mutual information between the set of all possible predictive statistics and the future history. $I[X_{t+1}^{\infty}, S_t] = I[X_{t+1}^{\infty}; X_{-\infty}^{t}]$

CSMs are sufficient because they capture all of the probablistic dependencies present in the spike train and they are minimal because they are constructed to be the most compact representation which does so.

CSMs contain within them the history dependent instantaneous spike rate, e.g the probablility of emitting a "1" given the current state. $\lambda(t+1 \mid H_t) = \Pr(x_{t+1} = 1 \mid S_t)$

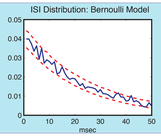
CAUSAL STATE MODELS ARE GENERATED FROM DATA

CSMs can be generated using the *Causal State Splitting Reconstruction (CSSR)* algorithm freely downloadable from *http://bactra.org/CSSR/*

CSSR is a non-parametric method and makes no assumptions as to the form of the CSM.

The only input (other than the data) is a *History Length "L"* which is chosen by minimizing the Baysian Information Criterion and bootstrapping 99% confidence bounds on the ISI distribution from the CSM.

CSM Generated from 200 sec of a 40 Hz Bernoulli Process



CAUSAL STATE MODELS QUANTIFY STRUCTURE

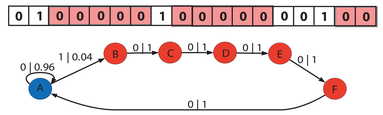
CSMs are minimal, optimally predictive hidden Markov models (HMM) and constitute a "program" which captures the statistical structure of the spike train.

(1) 40 Hz Bernoulli Process

Random spiking (40 Hz, p=0.04 spikes/msec IID) has no structure. The state of the spike generative process is always the same, regardless of the past spiking history.

(2) 40 Hz Bernoulli Process, 5 ms Refractory Period

Imposing a 5 msec refractory period after each spike imposes statistical structure on the spike train. Six states are required to describe the structure.



The structure can be sumarized by the ***COMPLEXITY***

$$C = H[S_t] = -E[\log(\Pr(S_t))]$$

ALGORITHMIC INFORMATION AND RANDOMNESS

The Algorithmic Information "K" of a string is the length of the shortest program which reproduces the string exactly and then stops. For a stochastic string this coincides with the Shannon Entropy.

Given a symbol string x_1^n, $K(x_1^n)$ can be written as:

$$E[K(x_1^n)] = H(x_1^n) + o(n) = \underbrace{H[T]}_{\text{structure}} + \underbrace{H[x_1^n \mid T]}_{\text{randomness}}$$

where T is a minimal sufficient statistic which describes the structure

the sequence of causal states S_0^n is a minimal sufficient statistic

Decomposition into Structure and Randomness

At time "t" the spike generating process is described by 1) the current state, 2) the next state, 3) the next emitted symbol. This suggests a natural decomposition of the entropy.

Complexity: Amount of information (in bits) about the system's past relevant for predicting its future.

Internal Entropy Rate: The average number of bits per time step, beyond the complexity, needed to describe the exact state sequence.

Residual Randomness: Randomness in the symbol emission which is not described by the CSM .

$$H[S_t, S_{t+1}, X_{t+1}] = H[S_t] + H[S_{t+1} \mid S_t] + H[X_{t+1} \mid S_t, S_{t+1}] = C + J + R$$

Complexity

$$C = H[S_t] = -E[\log(\Pr(S_t))]$$

Internal Entropy Rate

$$J = H[S_{t+1} \mid S_t] = -E[\log(\Pr(S_{t+1} \mid S_t))]$$

Residual Randomness

$$R = H[X_{t+1} \mid S_t, S_{t+1}] = -E[\log(\Pr(X_{t+1} \mid S_t, S_{t+1}))]$$

Complexity Only Needs to be Described Once!

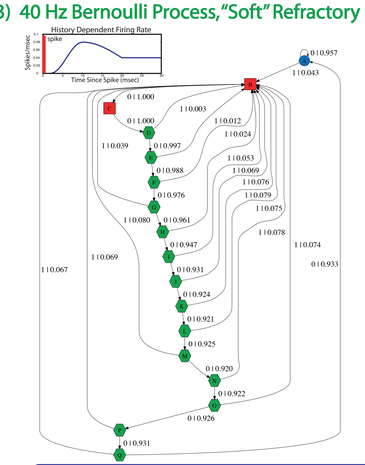
Because the states are Markovian, the complexity only needs to be described once, while the internal entropy rate and residual randomness are updated at each time step. The algorithmic information of the spike train is:

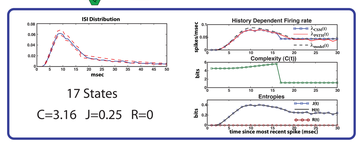
$$E[K(x_1^n)] = H[S_t] + N\{H[S_{t+1} \mid S_t] + H[X_{t+1} \mid S_t, S_{t+1}]\} = \underbrace{C}_{\text{structure}} + \underbrace{N(J+R)}_{\text{randomness}}$$

SIMULATED DATA

All CSMs were constructed by applying the CSSR algorithm and ISI bootstrapping to 200 seconds of simulated data with 1 millisecond time steps (N=200,000). No assumptions as to model form were made.

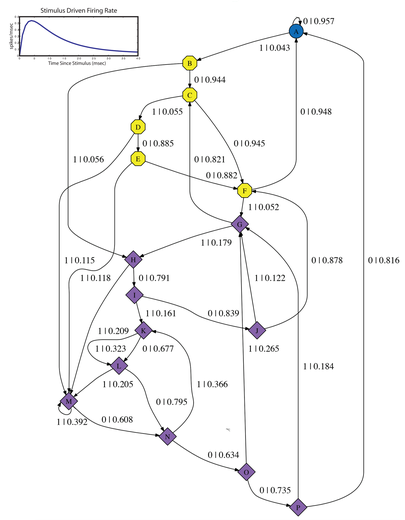
(3) 40 Hz Bernoulli Process, "Soft" Refractory Period

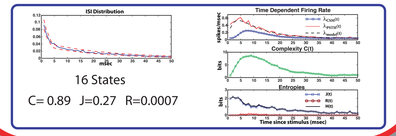




17 States
C=3.16 J=0.25 R=0

(4) "Stimulated" Inhomogeneous Bernoulli



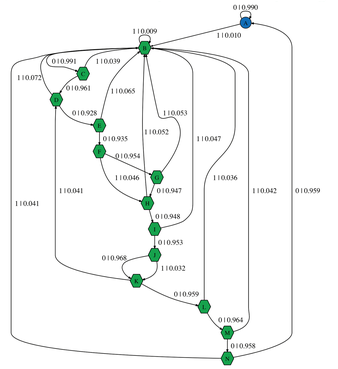


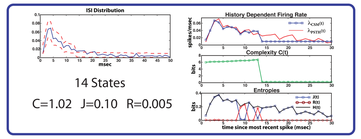
16 States
C= 0.89 J=0.27 R=0.0007

RAT BARREL CORTEX DATA

CSMs were constructed from spike data recorded by tetrodes in alpha chloralose anesthetized rat primary somatosensory SI (barrel) cortex. CSMs were generated for both spontaneous spiking and stimulus evoked (by vibrissa deflection) spiking.

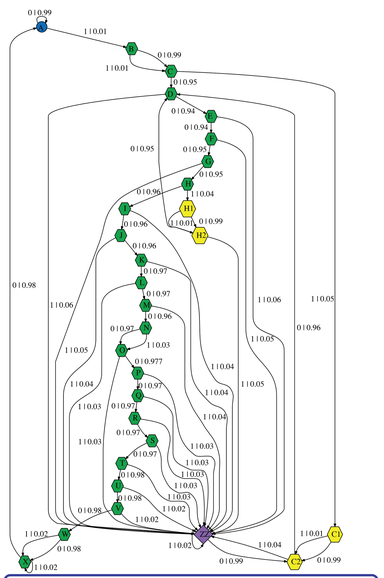
(5) 90 Seconds of Spontaneous Spiking

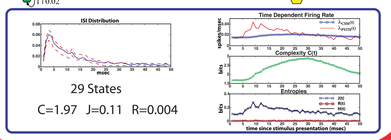




14 States
C=1.02 J=0.10 R=0.005

(6) 300 Seconds of Periodic Vibrissae Deflections





29 States
C=1.97 J=0.11 R=0.004

DISCUSSION

Neurons are complicated biophysical devices, which may not perform computation in a simple manner.

The Causal State Splitting Reconstruction Algorithm can be used to automatically determine the computational structure of spike trains in the form of a minimal optimally predictive hidden Markov model or Causal State Model (CSM).

The states of the CSM are equivalence classes of past spiking histories. Each member of a class has the same distribution of future spiking histories

The CSSR algorithm was used to generate CSMs from both simulated and experimentally recorded neural spike trains without making any assumptions as to the form of the CSM.

A lower bound on the amount of information supplied to the spike generation process by external influences (stimuli) can be derived from the CSM.

Possible applications include the study of the structural complexity of different cortical states, or comparisons of the complexity of neural activity during different sensory, motor or cognitive tasks.

CSMs complement traditional spike train analyses by describing the computational structure of spike trains, information which has previously been unavailable.

THE INFLUENCE OF STIMULI

CSMs are constructed assuming that the past spiking is the only available information for predicting the future spiking. This can lead to inaccurate firing rates when the spiking process is being externally stimulated.

The amount of information supplied to the spike generation process by external influences can be extracted from the CSM using the Kullback Leibler Divergence between the true spike distribution and that predicted by the CSM. This can be calculated using the entropy rate.

If we calculate the entropy over a stimulus ensemble as a function of the time since stimulus presentation:

$$H_{CSM}(t) = -\lambda_{true}(t)\log[\lambda_{CSM}(t)] - (1-\lambda_{true}(t))\log[1-\lambda_{CSM}(t)]$$

We can calculate the true time dependent firing probability

$$\lambda_{true}(t) = \frac{H_{CSM}(t) + \log[1-\lambda_{CSM}(t)]}{\log[1-\lambda_{CSM}(t)] - \log[\lambda_{CSM}(t)]}$$

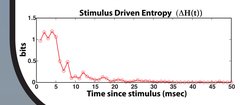
and the true entropy rate as a function of time since stimulus presentation

$$H_{true}(t) = -\lambda_{true}(t)\log[\lambda_{true}(t)] - (1-\lambda_{true}(t))\log[1-\lambda_{true}(t)]$$

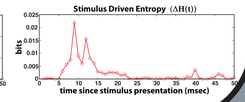
The difference between the CSM's entropy rate and the true entropy rate is the amount of information supplied by external processes

$$\Delta H = H_{CSM}(t) - H_{true}(t) = \lambda_{true}(t)\log\frac{\lambda_{true}(t)}{\lambda_{CSM}(t)} + (1-\lambda_{true}(t))\log\frac{1-\lambda_{true}(t)}{1-\lambda_{CSM}(t)}$$

Model Neuron



Barrel Cortex Neuron



ACKNOWLEDGEMENTS

The Authors acknowledge David Feldman, Kristina Klinkner Rob Kass, Emery Brown and Christopher Moore for thoughtful discussions.

The Authors thank Mark Andermann and Christopher Moore for use of their rat barrel cortex data.

This work was funded in part by NIH K25 NS052422