DISCOVERING SWITCHING AUTOREGRESSIVE DYNAMICS IN NEURAL SPIKE TRAIN RECORDINGS

By Scott W. Linderman,[??] Matthew J. Johnson, Sandeep R. Datta, and Ryan P. Adams[??]

Harvard University, Harvard Medical School

1. Introduction. The generalized linear model (GLM) is a powerful and successful model for identifying dependence in spiking populations of neurons, both over time and across the population. The GLM identifies these dependencies by modeling spiking pattern through a linear regression, with an appropriately-selected link function and likelihood. This linear regression setup is appealing for its simplicity, the wide variety of available regularizers/priors, the potential for interpretability, and the speed with which it can often be fit to data. However, the GLM suffers from at least three notable deficiencies. First, the model is linear up to the link function, which only allows a limited range of response maps from neural spiking histories. Second, the model's parameters are fixed over time; we would expect neural responses to vary due to plasticity and processes that are exogenous to the population. Third, the generalized linear model presupposes a characteristic time scale for all dynamics, when there may be multiple, varying time scales of neural activity in a given population.

cite

In this abstract, ...

Models. The fundamental building block of our spike train models is a hidden Markov model (HMM) consisting of K latent states, an initial state distribution $\boldsymbol{\pi}$, and a stochastic transition matrix $\boldsymbol{A} \in [0,1]^{K\times K}$. The probability of transitioning from state k to state k' is given by entry $A_{k,k'}$. The latent state in the t-th time bin is given by $z_t \in \{1,\ldots,K\}$. We observe a sequence of vectors $\boldsymbol{s}_t \in \mathbb{N}^N$ where $s_{n,t}$ specifies the number of spikes emitted by the n-th neuron during the t-th time bin.

To model autoregressive structure in the spike train, we model the spike count vectors $\boldsymbol{x}_t$ as arising from an autoregressive negative binomial model. The expected number of spikes is a weighted function of recent spike count vectors. Specifically, let

$$\boldsymbol{\psi}_n \sim \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{\psi}}, \sigma^2_{\boldsymbol{\psi}}\boldsymbol{I}), \quad \boldsymbol{\mu}_{\boldsymbol{\psi}} = \boldsymbol{X}\boldsymbol{w}_n^{(k)}, \tag{1}$$

where $\boldsymbol{X} \in \mathbb{R}^{T\times B}$ is a matrix of given regressors and $\boldsymbol{w}_n^{(k)} \in \mathbb{R}^B$ is a vector of weights specific to neuron n and the k-th latent state. Typically the regressors include the filtered spike history and externally applied stimuli, allowing the weights to be interpreted as "functional interaction" strengths and stimulus response functions. Let $\boldsymbol{W}^{(k)} = \{\boldsymbol{w}_1^{(k)},\ldots,\boldsymbol{w}_N^{(k)}\}$ denote the set of all weight vectors for state k. This model allows the weights to vary over time with the latent state of the population.

Together these specify a joint probability distribution,

$$p(\{\boldsymbol{s}_t, z_t\}_{t=1}^T \,|\, \boldsymbol{\pi}, \boldsymbol{A}, \{\boldsymbol{W}^{(k)}\}_{k=1}^K, \boldsymbol{X}, \xi, \sigma^2_{\boldsymbol{\psi}}) = p(z_1 \,|\, \boldsymbol{\pi}) \prod_{t=2}^T p(z_t \,|\, z_{t-1}, \boldsymbol{A}) \prod_{t=1}^T p(\boldsymbol{s}_t \,|\, z_t, \{\boldsymbol{W}^{(k)}\}_{k=1}^K, \boldsymbol{X}, \xi, \sigma^2_{\boldsymbol{\psi}}),$$

$$p(z_1 \,|\, \boldsymbol{\pi}) = \text{Multinomial}(z_1 \,|\, \boldsymbol{\pi}), \qquad p(z_t \,|\, z_{t-1}, \boldsymbol{A}) = \text{Multinomial}(z_t \,|\, \boldsymbol{A}_{z_{t-1},:}),$$

$$p(\boldsymbol{\psi}_t \,|\, z_t, \boldsymbol{X}, \{\boldsymbol{W}^{(k)}\}_{k=1}^K) = \prod_{n=1}^N \mathcal{N}(\psi_{n,t} \,|\, \boldsymbol{X}\boldsymbol{w}_n^{(z_t)}, \sigma^2_{\boldsymbol{\psi}}),$$

$$p(\boldsymbol{s}_t \,|\, \boldsymbol{\psi}_t, \xi) = \prod_{n=1}^N \text{NegBin.}(s_{n,t} \,|\, \xi, \sigma(\psi_{n,t})),$$

where $\sigma(\cdot)$ denotes the logistic function. We use a hierarchical Dirichlet process (HDP) as a prior distribution on the rows of the transition matrix, thereby allowing for nonparametric inference of the number of states. Additionally, we consider is the hidden semi-Markov model (HSMM). In a standard HMM, the amount of

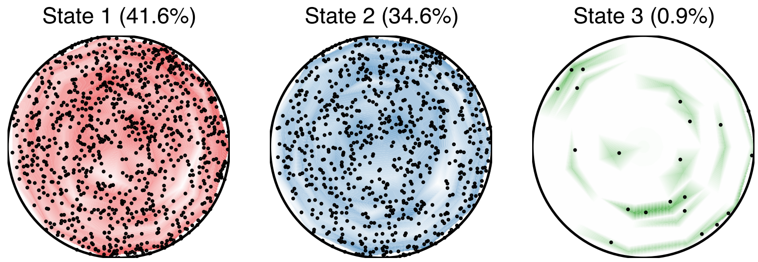

Fig 1

time spent in state k is geometrically distributed as a function of $A_{k,k}$. In many real-world examples, however, the state durations may follow more interesting distributions, such as a negative binomial distribution. As with the HMM, we can also derive a nonparametric extension called the HDP-HSMM with similarly efficient inference algorithms.