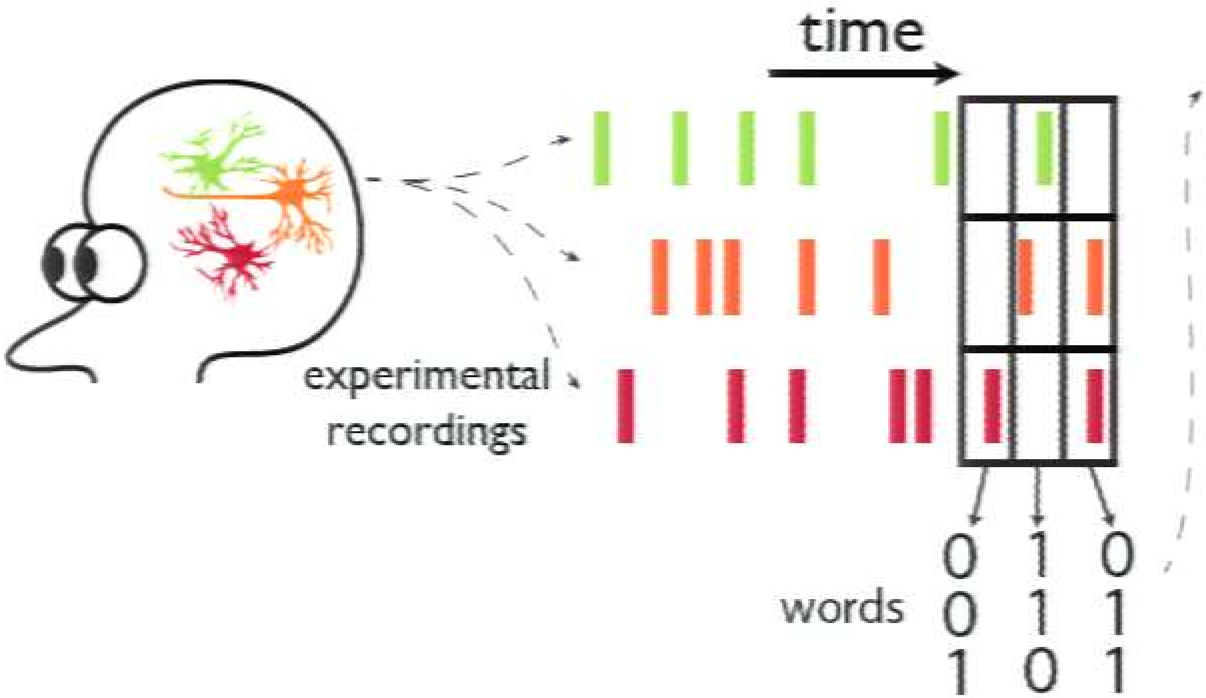
Bayesian entropy estimation for binary spike train data using parametric prior knowledge

Evan Archer   Il Memming Park   Jonathan W. Pillow



spike distribution π

$$\pi_1 = p(0\,0\,1)$$

$$\pi_2 = p(1\,1\,0)$$

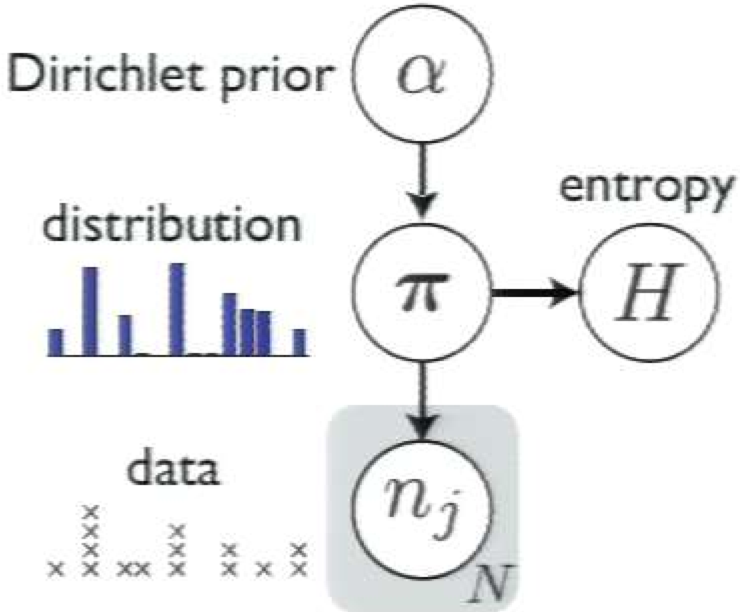
$$\vdots$$

high dimensional!

goal: Estimate entropy of binary spike data from samples

$$H(\boldsymbol{\pi}) = -\sum_{i=1}^{2^m} \pi_i \log(\pi_i)$$

- far more words than samples



Bayesian approach:

$$\boldsymbol{\pi} \sim \mathrm{Dirichlet}(\alpha)$$

Bayes Least Squares Estimator:

$$\begin{aligned}\hat{H} &= \mathbb{E}[H|\mathrm{data}, \alpha] \\ &= \int H(\boldsymbol{\pi}) p(\boldsymbol{\pi}|\mathrm{data}, \alpha) d\boldsymbol{\pi}\end{aligned}$$

problem: Dirichlet priors weight each word equally

- synchronous spikes are unlikely!

$$p(1\,1\,1\,1\,1) \neq p(0\,0\,0\,0\,0)$$

solution: we choose priors that,

- exploit spike train structure
- make computation tractable

our approach: center with simple parametric model

$$\pi \sim \mathrm{Dirichlet}(\alpha G)$$

closed-form moments of entropy

simple parametric model for spike patterns

problem: sum intractable for most base measures

$$E[H|\mathrm{data}] = \psi_0\Big(\sum_j \alpha_j + 1\Big) - \frac{1}{\sum_j \alpha_j} \sum_{i=1}^{2^M} \alpha_i \psi_0(\alpha_i + 1)$$

digamma

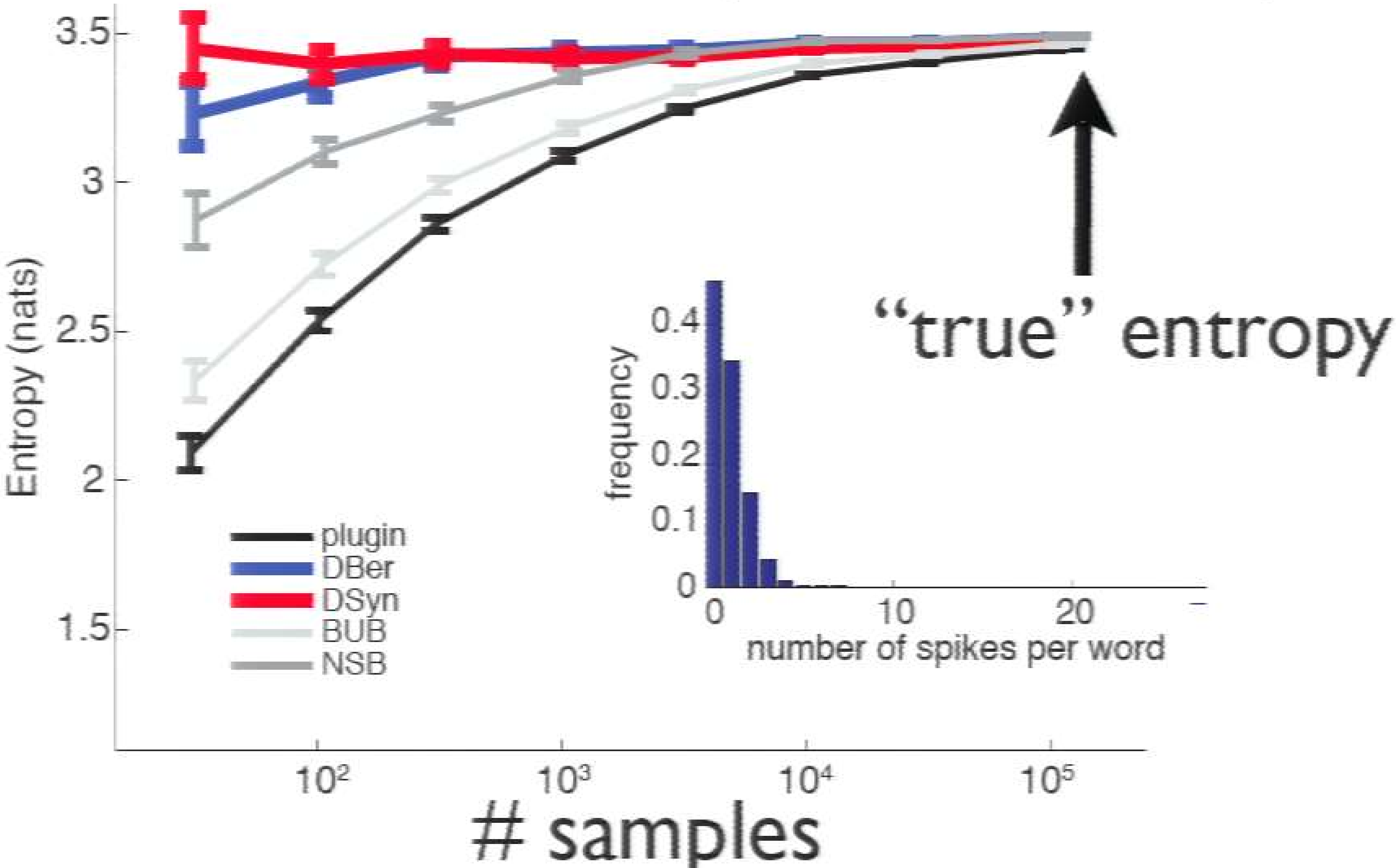
solution: models with equivalence classes of words

- e.g., probability depends on # spikes: $G = \mathrm{Bernoulli}(p)$

m neurons	0	1	0	0	0	1	1	1	0	0	0	0	1	1	1	1
	0	0	1	0	0	1	0	0	1	1	0	1	0	1	1	1
	0	0	0	1	0	0	1	0	1	0	1	1	1	0	1	1
	0	0	0	0	1	0	0	1	0	1	1	1	1	1	0	1
	most likely	less likely				even less likely						less likely still				least likely

Retinal Ganglion Cell data (1ms bins)

27 ON and OFF parasol cells [Chichilnisky lab]



Poster 43 tonight for more details (44 next door)