1

Neural coding of a natural stimulus ensemble: Uncovering information at sub-millisecond resolution

Ilya Nemenman (LANL/CCS-3) and William Bialek (Princeton) Rob de Ruyter van Steveninck (Indiana)

http://nsb-entropy.sourceforge.net

Abstract

Our knowledge of the sensory world is encoded by neurons in sequences of discrete, identical pulses termed action potentials or spikes. There is persistent controversy about the extent to which the precise timing of these spikes is relevant to the function of the brain. We revisit this issue, using the motionsensitive neurons of the fly visual system as a test case. New experimental methods allow us to deliver more nearly natural visual stimuli, comparable to those which flies encounter in free, acrobatic flight, and new mathematical methods allow us to draw more reliable conclusions about the information content of neural responses even when the set of possible responses is very large. We find that significant amounts of visual information are represented by details of the spike train at millisecond and sub--millisecond precision, even though the sensory input has a correlation time of about 60 ms; different patterns of spike timing represent distinct motion trajectories, and the absolute timing of spikes points to particular features of these trajectories with high precision. Under these naturalistic conditions, the system continues to transmit more information at higher photon flux, even though individual photoreceptors are counting more than one million photons per second, and removes redundancy in the stimulus to generate a more efficient neural code.

Fly as a model neurocomputing system

Why fly?

- Can record for long times
- Named neurons with known functions
- Nontrivial computation (motion estimation)
- Vision (specifically, motion estimation) is behaviorally important
- Possible to generate natural stimuli

What questions?

- Can we understand the code?
- Which features of it are important?
 - Rate or precise timing (how precise)?
 - Synergy between spikes?
- What/how much does the fly know?
- Is there an evidence for optimality?

Recording from fly's H1



(Lewen et al, 2001)

- Fight the second sec

outdoor

daylight

- Poisson behavior for "boring" stimuli (rotations)
- ~2 ms resolution important for white noise stimuli
- Could "brisk" spikes be due to ~1 ms correlations in stimulus? What if stimulus has natural correlations?

Natural stimuli and responses High repeatability - no rate coding





Is high precision timing for natural stimuli relevant for information transmission, or just anecdotal?

Characterization of coding without an explicit decoding: Experiment design

$$S[x] = -\sum_{x} p(x) \log p(x), \quad x = s, \{t_i\}$$

$$I[s, \{t_i\}] = \sum_{s\{t_i\}} p(s, \{t_i\}) \log \frac{p(s, \{t_i\})}{p(s)p(\{t_i\})}$$



- Captures all dependencies
- s stimulus, $\{t_i\}$ spike train

Problem: undersampling

- Total of about 10-15 min of recordings
- At most 200 repetitions
- Stimulus correlated at 60ms: only 1e4 independent samples
- Need to sample words of duration 30-60 ms at resolutions ~0.2 ms (binary words of *T* ~100) from 100-10000 samples only.

(Strong et al., 1998)

Undersampling and entropy/MI estimation

$$\langle S_{ML} \rangle = \left\langle -\sum_{i} \frac{\langle n_i \rangle}{N} \log \frac{\langle n_i \rangle}{N} \right\rangle \le -\sum_{i} \frac{\langle n_i \rangle}{N} \log \frac{\langle n_i \rangle}{N} = S$$

bias $\propto -\frac{2^S}{N} \gg (\text{variance})^{1/2} \propto \frac{1}{\sqrt{N}}$

- Fluctuations underestimate entropies and overestimate mutual informations.
- Universal bias correction possible only for S << log N
- This condition violated for our data

But there is hope for $S > \log N$ (Ma, 1981):

• For uniform *K*-bin distribution, the first concidence happens at

- Can make estimates for square-root-fewer samples!
- To extended to nonuniform cases need:
 - Assumptions (won't work always)
 - Estimate entropies without estimating distributions

NSB estimator:

Uniformized expectations about S

What is unknown?

Example: binomial distribution $S = -p \log p - p \log p$

$$(1-p)\log(1-p)$$





-2 -1 0 1 2

NSB: one possible S-uniformization strategy

$$P_{\beta}(\{q_i\},\beta) = \frac{1}{Z} \left. \delta \left(1 - \sum_{i=1}^{K} q_i \right) \prod_{i=1}^{K} q_i^{\beta} \left. \frac{dS}{d\beta} \right|_{N=0} P(S|_{N=0}) \right.$$

- Posterior variance scales as N^{-1/2}
- Little bias, except in some known cases.
- Counts coincidences and works in Ma regime (if works).
- Is guaranteed correct for large *N*.
- Allows infinite # of bins.

$$\varepsilon = \left\langle \frac{S_{est} - S_{true}}{\delta S_{est}} \right\rangle$$

- Selection of wrong "unknown" biases the estimation.
- Even worse for large *K*.

(Nemenman et al, 2002-06)



Refractory Poisson, rate 0.26 spikes/ms, refractory period 1.8 ms, *T*=15ms, discretization 0.5ms, true entropy 13.57 bits.



Natural spike trains



(Nemenman et al. 2004)

N = 75

9

Information rate and all that



- Information present up to τ =0.2 ms (comparable to exparimental noise)
- 30% more information at τ<1ms.

- ~1 bit/spike at 150 spikes/s and low-entropy correlated stimulus.
 Design principle?

New bits (optimized code)



- Correlation function at half its value, but fly gets new bits every 25 ms
- Independent info decorrelation

Behaviorally optimized code!

Information about...

Conditional distribution (median \pm 1 quartile) of velocities preceeding a spike pattern.



Precision is limited by physical noise sources



T = 6 ms $\tau = 0.2 \text{ ms}$ 1.49 vs. 1.61 \cdot 10⁶ ph/(s \cdot rec) $I^+ - I^- = 0.0204 \pm 0.0108 \text{ bits (p<6\%)}$

(Lewen, et al 2001)