Optimal Mutual Information Quantization is NP complete

Brendan Mumey (Computer Science) and
Tomáš Gedeon (Math)

Montana State University – Bozeman

Acknowledgements

Research was supported by NSF BITS grant 0129895.

Quantization Problem

- What is the computational complexity of quantizing a joint **(X,Y)** probability distribution in order to maximize the mutual information of the quantization?

- We show this problem is NP-complete via reductions from various forms of the PARTITION problem.

Neural Coding Application

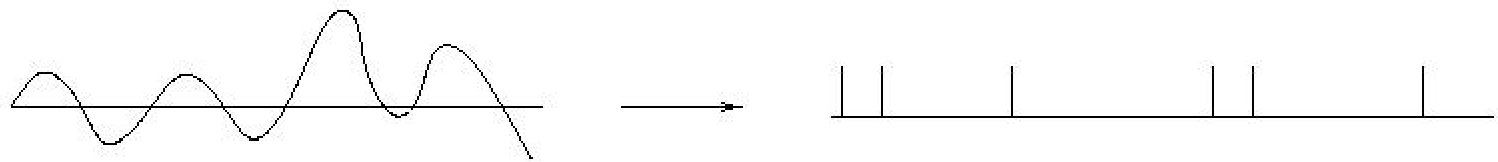
- In the application to neural coding, one models the input to the system as a random variable **X** and the output as a random variable **Y**.

Stochastic Framework

Random Variables:

- X stimulus (waveform)

- **Y** neural response (single channel for now)



stimulus **X=x**

neural response **Y=y**

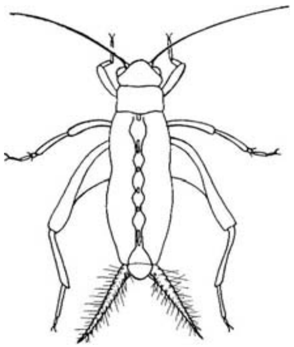
… **(X,Y)** forms a joint distribution

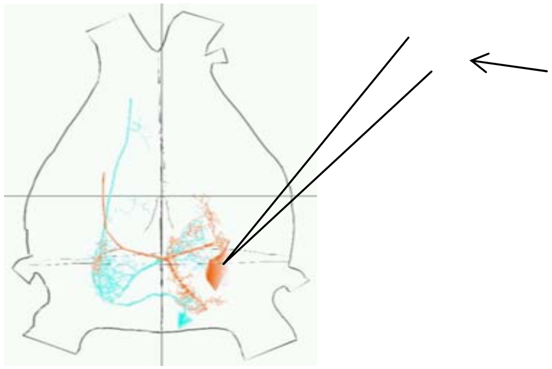
Cricket Data Set

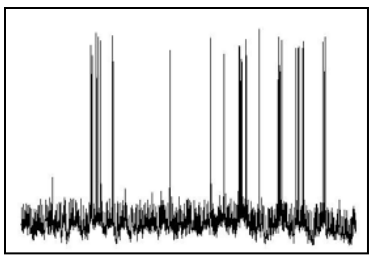
- Target system: 50-70 neurons comprising the terminal ganglion in crickets
- Each probe recording is a stream of inter-spike distances:

stream 1: 3243, 7343, 23433, 322, 8983, 1831, …
stream 2: 17983, 289, 5934, 2893, 42398, 3985, …

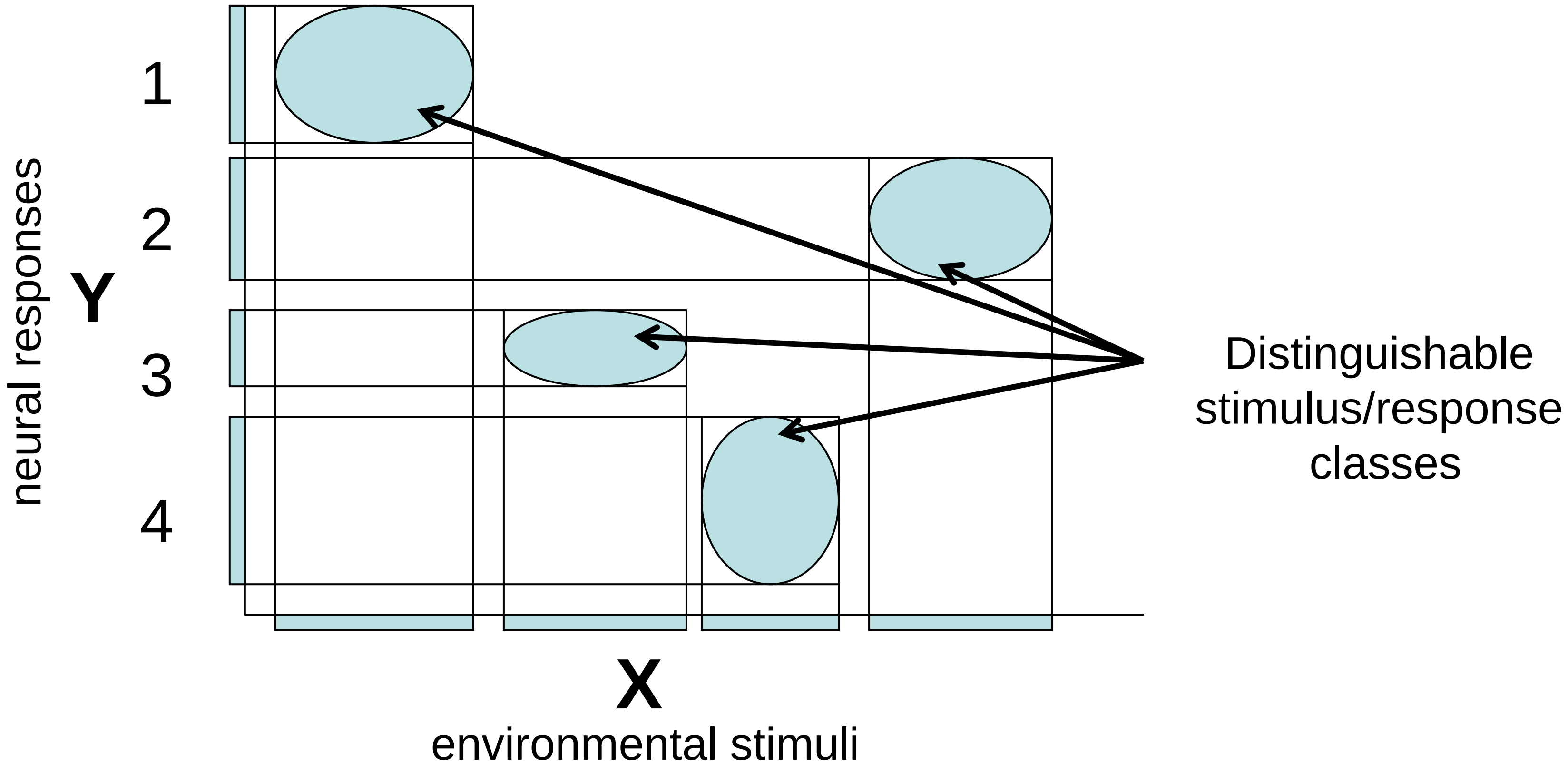
Multiple probes are used to record individual neurons







Example Quantization



Mutual Information

- Given two discrete random variables **X** and **Y** the mutual information **I(X,Y)** is defined by

$$\mathrm{I}(\mathrm{X},\mathrm{Y}) = \sum_{x,y} \Pr(x,y) \log \frac{\Pr(x,y)}{\Pr(x)\Pr(y)}$$

Joint Quantization

- Fix sizes **M** and **N** of the reproduction spaces
- Let $\mathbf{A_M = \{a_1, \ldots, a_m\}}$ for **X** and $\mathbf{B_N = \{b_1, \ldots, b_n\}}$ for **Y** and define quantizers, $\mathbf{q(a_i|x)}$ and $\mathbf{q(b_j|y)}$.
- Quantizers are conditional probabilities so $\mathbf{q(a_i|x)} \in \Delta_M$, and $\mathbf{q(b_j|y)} \in \Delta_N$, where $\Delta_K = \{z \in R^K \mid \Sigma_i z_i = 1, z_i >= 0\}$.
  <u>Fact</u>: optimal when $z_i = 1$ *for some* ***i.***
- The distortion function to be minimized is:

$$\mathbf{I(X,Y) - I(A_M,B_N).}$$

Decision Problem

MxN-QUANT:

Given a joint **(X,Y)** probability distribution, decide if there exists quantizers $\mathbf{q(a_i|x)}$ and $\mathbf{q(b_j|y)}$ of reproduction spaces $\mathbf{A_M}$ and $\mathbf{B_N}$ with mutual information $\mathbf{I(A_M,B_N)}$ at least $\boldsymbol{s > 0}$.

The PARTION Problem

- Given a set of real numbers **R = {r_1, r_2, …, r_n},** decide if there is a set of indices **S** such that

$$\sum_{i \in S} r_i = \sum_{i \notin S} r_i$$

- **Generalized k-PARTITION problem**: Find index sets **{S_1, …, S_k}** such that for any **1 <= u, v <= k**:

$$\sum_{i \in S_u} r_i = \sum_{i \in S_v} r_i$$

Lemma 1

- ***Lemma***: Consider a joint quantization problem with ***M=N=k >= 2***. Then **$I(A_k,B_k) <= \lg k$** bits with equality achieved only when there exists **$q(a_i|x)$** and **$q(b_j|y)$** such that:

$$\Pr(a_i, b_j) = \begin{cases} 1/k & \text{if i = j} \\ 0 & \text{otherwise} \end{cases}$$

Proof of Lemma 1

Observe **$I(A_k,B_k) <= \min(H(A_k), H(B_k))$.**
Since **$I(A_k, B_k) = H(A_k) - H(B_k \mid A_k) = H(B_k) - H(A_k \mid B_k)$,** equality is achieved when **$H(B_k \mid A_k) = H(A_k \mid B_k) = 0$** and **$H(A_k) = H(B_k) = \lg k$**.

A random variable with **k** outcomes achieves a maximum entropy of **lg k** bits only if each outcome equally likely and so **$p(a_i) = p(b_j) = 1/k$.** We have

$$\begin{aligned} I(A_k, B_k) &= \sum_{i,j} p(a_i, b_j) \lg [k^2 p(a_i, b_j)] \\ &= 2 \lg k + \sum_i \sum_j p(a_i, b_j) \lg p(a_i, b_j) \end{aligned}$$

since **$\Sigma_{ij}\, p(a_i, b_j) = 1$.**
Given the constraints **$\Sigma_j\, p(a_i, b_j) = 1/k$** and the fact that **$x \lg x$** is convex, the inner sum is maximized exactly when **$p(a_i, b_j) = 1/k$** for one value of **j** and the remaining probabilities are **0**.
Given the constraint **$\Sigma_i\, p(a_i, b_j) = 1/k$** it follows that if **$p(a_i, b_j) = 1/k$** then **$p(a_u, b_j) = p(a_i, b_v) = 0$** for all **$u \neq i$ and $v \neq j$.**
By permuting the class labels appropriately the lemma is proven.

Reduction

- Let **R = {r_1, r_2, …, r_n}** be an instance of k-PARTITION.
- Let **r =** Σ **r_i**.
- Consider the following **n x n** joint distribution:

$$\Pr(X,Y) = \begin{bmatrix} r_1/r & 0 & \cdots & 0 \\ 0 & r_2/r & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & r_n/r \end{bmatrix}$$

Main Theorem

Theorem: **R** can be **k**-partitioned if and only if there exists a joint quantization (A_k, B_k) of **(X,Y)** with **lg k** bits of mutual information.

Proof:

- **"=>"** : just pick quantization classes corresponding to partitions
- **"<=“** : Lemma 1 implies that the optimal quantizer will partition **(X,Y)** into equally weighted classes. Can easily recover a **k**-partition of **R** from this.

Open Problems

- Constant factor approximation?
- PTAS?