

# 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.

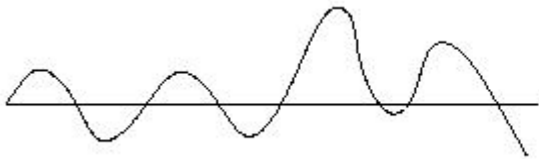
# Neural Coding Application

- In the application to neural coding, one models the input to the system as a random variable  $\mathbf{X}$  and the output as a random variable  $\mathbf{Y}$ .

# 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

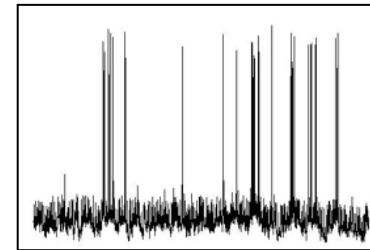
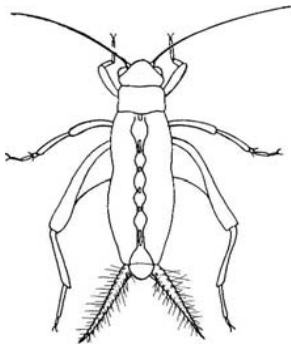
# Cricket Data Set

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

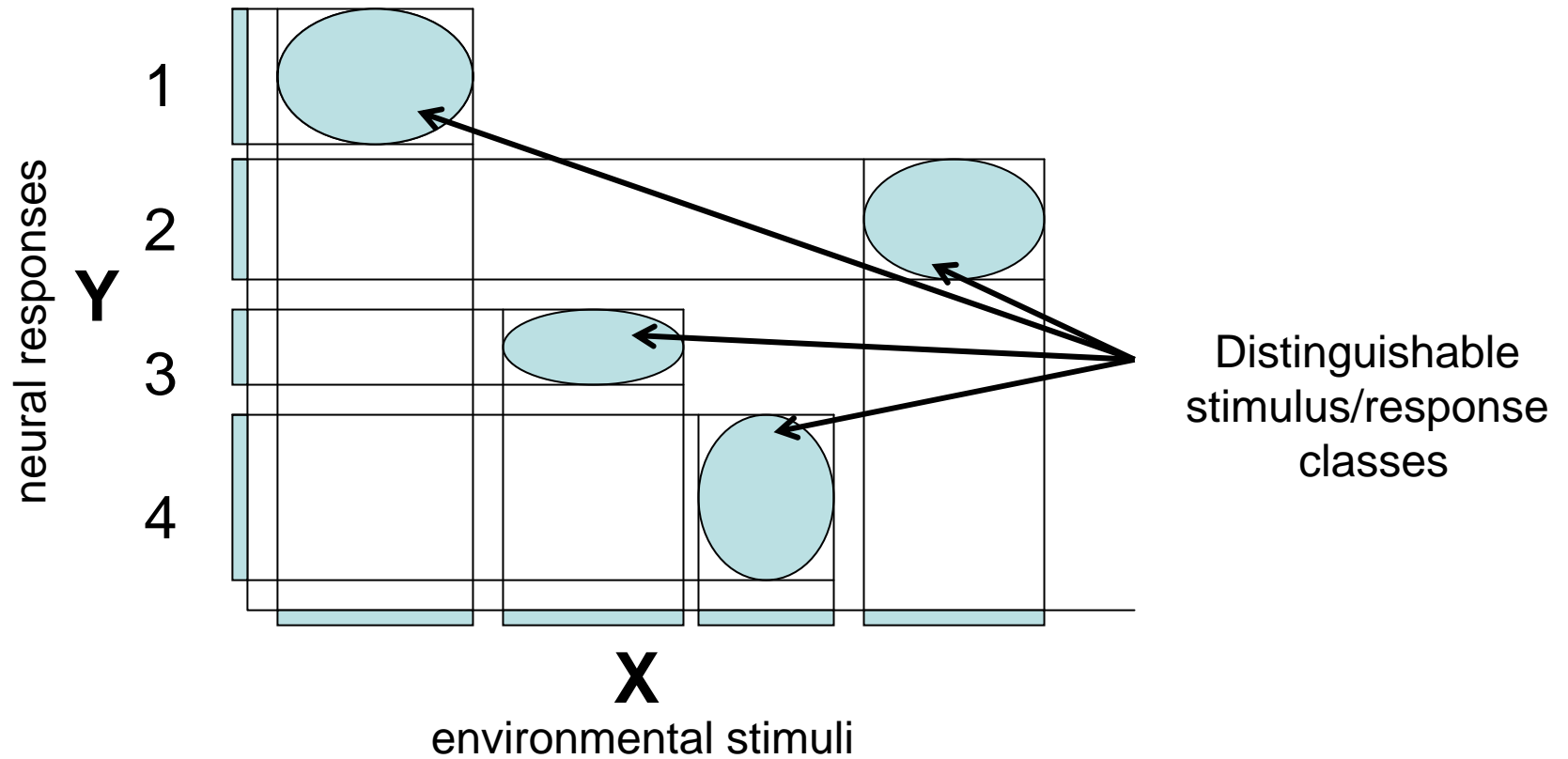
stream 1: 3243, 7343, 23433, 322, 8983, 1831, ...

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

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

# Joint Quantization

- Fix sizes  $\mathbf{M}$  and  $\mathbf{N}$  of the reproduction spaces
- Let  $\mathbf{A}_M = \{\mathbf{a}_1, \dots, \mathbf{a}_m\}$  for  $\mathbf{X}$  and  $\mathbf{B}_N = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$  for  $\mathbf{Y}$  and define quantizers,  $\mathbf{q}(\mathbf{a}_i|\mathbf{x})$  and  $\mathbf{q}(\mathbf{b}_j|\mathbf{y})$ .
- Quantizers are conditional probabilities so  $\mathbf{q}(\mathbf{a}_i|\mathbf{x}) \in \Delta_M$ , and  $\mathbf{q}(\mathbf{b}_j|\mathbf{y}) \in \Delta_N$ , where  $\Delta_K = \{\mathbf{z} \in \mathbf{R}^K \mid \sum_i z_i = 1, z_i \geq 0\}$ .  
*Fact: optimal when  $\mathbf{z}_i = 1$  for some  $i$ .*
- The distortion function to be minimized is:

$$I(\mathbf{X}, \mathbf{Y}) - I(\mathbf{A}_M, \mathbf{B}_N).$$

# Decision Problem

## **MxN-QUANT:**

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

# The PARTITION Problem

- Given a set of real numbers  $\mathbf{R} = \{r_1, r_2, \dots, r_n\}$ , decide if there is a set of indices  $\mathbf{S}$  such that

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

- Generalized k-PARTITION problem:** Find index sets  $\{\mathbf{S}_1, \dots, \mathbf{S}_k\}$  such that for any  $1 \leq u, v \leq k$ :

$$\sum_{i \in \mathbf{S}_u} r_i = \sum_{i \in \mathbf{S}_v} r_i$$

# Lemma 1

- **Lemma:** Consider a joint quantization problem with  $M=N=k \geq 2$ . Then  $I(\mathbf{A}_k, \mathbf{B}_k) \leq \lg k$  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  $I(\mathbf{A}_k, \mathbf{B}_k) \leq \min(H(\mathbf{A}_k), H(\mathbf{B}_k))$ .

Since  $I(\mathbf{A}_k, \mathbf{B}_k) = H(\mathbf{A}_k) - H(\mathbf{B}_k | \mathbf{A}_k) = H(\mathbf{B}_k) - H(\mathbf{A}_k | \mathbf{B}_k)$ , equality is achieved when  $H(\mathbf{B}_k | \mathbf{A}_k) = H(\mathbf{A}_k | \mathbf{B}_k) = 0$  and  $H(\mathbf{A}_k) = H(\mathbf{B}_k) = \lg k$ .

A random variable with  $k$  outcomes achieves a maximum entropy of  $\lg k$  bits only if each outcome equally likely and so  $p(\mathbf{a}_i) = p(\mathbf{b}_j) = 1/k$ . We have

$$\begin{aligned} I(\mathbf{A}_k, \mathbf{B}_k) &= \sum_{i,j} p(\mathbf{a}_i, \mathbf{b}_j) \lg [k^2 p(\mathbf{a}_i, \mathbf{b}_j)] \\ &= 2 \lg k + \sum_i \sum_j p(\mathbf{a}_i, \mathbf{b}_j) \lg p(\mathbf{a}_i, \mathbf{b}_j) \end{aligned}$$

since  $\sum_{ij} p(\mathbf{a}_i, \mathbf{b}_j) = 1$ .

Given the constraints  $\sum_j p(\mathbf{a}_i, \mathbf{b}_j) = 1/k$  and the fact that  $x \lg x$  is convex, the inner sum is maximized exactly when  $p(\mathbf{a}_i, \mathbf{b}_j) = 1/k$  for one value of  $j$  and the remaining probabilities are 0.

Given the constraint  $\sum_i p(\mathbf{a}_i, \mathbf{b}_j) = 1/k$  it follows that if  $p(\mathbf{a}_i, \mathbf{b}_j) = 1/k$  then  $p(\mathbf{a}_u, \mathbf{b}_j) = p(\mathbf{a}_i, \mathbf{b}_v) = 0$  for all  $u \neq i$  and  $v \neq j$ .

By permuting the class labels appropriately the lemma is proven.

# Reduction

- Let  $\mathbf{R} = \{r_1, r_2, \dots, r_n\}$  be an instance of k-PARTITION.
- Let  $r = \sum r_i$ .
- Consider the following  $n \times n$  joint distribution:

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

# Main Theorem

Theorem:  $\mathbf{R}$  can be  $\mathbf{k}$ -partitioned if and only there exists a joint quantization  $(\mathbf{A}_k, \mathbf{B}_k)$  of  $(\mathbf{X}, \mathbf{Y})$  with  $\lg k$  bits of mutual information.

Proof:

“ $\Rightarrow$ ” : just pick quantization classes corresponding to partitions

“ $\Leftarrow$ ” : Lemma 1 implies that the optimal quantizer will partition  $(\mathbf{X}, \mathbf{Y})$  into equally weighted classes. Can easily recover a  $\mathbf{k}$ -partition of  $\mathbf{R}$  from this.

# Open Problems

- Constant factor approximation?
- PTAS?