

# Examining methods for estimating mutual information in spiking neural systems

Christopher J. Rozell\* and Don H. Johnson

*Department of Electrical & Computer Engineering  
Rice University, Houston, TX, 77251*

---

## Abstract

Mutual information enjoys wide use in the computational neuroscience community for analyzing spiking neural systems. Its direct calculation is difficult because estimating the joint stimulus-response distribution requires a prohibitive amount of data. Consequently, several techniques have appeared for bounding mutual information that rely on less data. We examine two upper bound techniques and find that they are unreliable and can introduce strong assumptions about the neural code. We also examine two lower bounds, showing that they can be very loose and possibly bear little relation to the mutual information's actual value.

*Key words:* mutual information, neural coding, information theory

---

## 1 Introduction

Mutual information is in many ways the cornerstone of classic information theory, playing central roles in the analysis of both digital and analog communications systems [1]. One of the primary objectives in studying neural communication is characterizing the amount of stimulus information the output spike train represents. Though it is not clear that mutual information alone answers this question [2], all agree that it is an important measure of the statistical relationship between the stimulus and the response.

Throughout this report we consider the paradigm where a continuous-valued stimulus  $S$  is encoded in a single neuron spike response  $R$  measured over the time window  $[0, T]$ . The stimulus and its probability distribution  $p(S)$  will be

---

\* Corresponding author.

*Email address:* `crozell@rice.edu` (Christopher J. Rozell).

considered in an abstract sense, where  $S$  could represent a time domain signal or a stimulus parameter. The spike response  $R$  consists of both the number of spikes in the interval  $[0, T]$  and the actual spike times. The probability distribution  $p(R)$  is therefore a mixed distribution over one discrete parameter (the count) and a variable number of continuous parameters (the spike times).

Formally, the mutual information [1, 3] between  $S$  and  $R$  is given by

$$I(S; R) = \int p(s) \int p(r|s) \log \frac{p(r|s)}{p(r)} dr ds. \quad (1)$$

It is well known that the mutual information can be written as a difference of differential entropies [3],  $I(S; R) = h(R) - h(R|S)$ , where

$$h(R) = - \int_{-\infty}^{\infty} p(r) \log p(r) dr,$$

$$h(R|S) = - \int_{-\infty}^{\infty} p(r) \int_{-\infty}^{\infty} p(r|s) \log p(r|s) ds dr.$$

The input probability distribution  $p(S)$  is entirely under experimental control, but the input-output relationship of the system,  $p(R|S)$ , needs to be measured. If the spike response contains dependencies (serial or inter-neuron), then a large number of repetitions are required to estimate  $p(R|S)$  completely.

In practice, spike responses are almost always binned with binwidth  $\Delta$  and  $I(S; R)$  is estimated from the discrete approximation  $I_{\Delta}(S; R) = h_{\Delta}(R) - h_{\Delta}(R|S)$ . It is true that  $I_{\Delta}(S; R) \rightarrow I(S; R)$  as  $\Delta \rightarrow 0$ , but smaller values of  $\Delta$  require more data to make reliable estimates. The “direct method” for estimating mutual information [5] estimates  $h_{\Delta}(R)$  by presenting stimuli drawn from  $p(S)$ , and estimates  $h_{\Delta}(R|S)$  from repeated presentations of the same stimulus. Because the amount of data required to estimate  $p(R|S)$  is usually very large, most researchers employ techniques for bounding  $I(S; R)$ .

Neuroscientists use mutual information because they want to characterize how well a spike response conveys the information contained in external stimuli. Information theory gives meaning to mutual information through the channel coding theorem [1] (digital communications) and rate-distortion theory [1, 4] (analog communications). Both cases involve optimizing mutual information over the stimulus probability distribution  $p(S)$ , and calculating  $I(S|R)$  for another choice of  $p(S)$  may not be theoretically interesting [2]. Bounding techniques are especially important as we consider the need to optimize  $I(S; R)$  over  $p(S)$  instead of just calculating it once. If mutual information is to be used in analyzing neural data, we must determine whether our present techniques for bounding it are adequate. We investigate several techniques found in the computational neuroscience literature, to see if the bounds are always valid and useful. We specifically examine cases where spiking neurons violate the underlying assumptions of the method (e.g., Gaussian noise). In our in-

vestigations, we focus on the techniques themselves by assuming that we can make perfect estimates. In reality, estimations will deteriorate with limited sample sizes.

## 2 Simple bounds

One class of bounding methods are all based on simple observations. For example, the simplest upper bound would be to claim that because mutual information can be written as a difference of entropies, it is upper bounded by  $h(R) \geq I(S|R)$ . However, because differential entropy can be negative, this claim is not true in general. After binning, it is true that the discrete entropy is always non-negative  $h_\Delta(R|S) \geq 0$ , meaning that  $h_\Delta(R) \geq I_\Delta(S; R)$ . Unfortunately, any entropy calculated in this way depends heavily on the choice of  $\Delta$ , and actually doesn't converge as  $\Delta \rightarrow 0$ . More importantly, the binned mutual information converges to the true mutual information from below,  $I_\Delta(S; R) \leq I(S; R)$ , and for any particular choice of  $\Delta$ ,  $h_\Delta(R)$  may be above or below  $I(S; R)$ . Therefore,  $h_\Delta(R)$  is not an upper bound for the true mutual information for all choices of  $\Delta$ , and it becomes the trivial upper bound of infinity as  $\Delta \rightarrow 0$ . Consider a simple example with an exponentially distributed stimulus producing a constant-rate spike train. Calculating the binned entropy according to the standard methods [7, p.967] and analytically calculating the true mutual information [10] shows (figure 1(a)) that  $h(R)$  is not an upper bound in general.

The simplest lower bound method involves estimating the stimulus  $\hat{S}$  from the spike response  $R$  [5, 6]. Due to a result known as the data processing inequality [3],  $I(S; \hat{S}) \leq I(S; R)$  regardless of the estimator used. This method will always produce a lower bound, but is estimator specific. Some estimators could produce very tight bounds if they are based on an accurate model of the neural coding process. Other estimators could produce the trivial bound of zero. As [5] points out, this technique could be useful for checking the validity of a coding model, but only *if* we knew  $I(S; R)$  or a correct upper bound.

## 3 Gaussian approximation bounds

Another class of bounding techniques assumes that if the stimulus presented is Gaussian, then the response is also Gaussian and the result of an additive Gaussian noise process that is independent of the stimulus. A Gaussian model makes little sense for a spike train that consists of discrete point events, and Poisson spiking models cannot be statistically independent of the noise (the variance is the same as mean rate determined by the stimulus). However,

because the intent is only to find an upper bound, the reality of the model does not matter if the bound is accurate and effective. With this model, the mutual information rate [1] is

$$I(S; R) = \int_0^W \log(1 + \text{SNR}(f)) df, \quad (2)$$

where  $\text{SNR}(f) = P_{SS}(f)/P_{NN}(f)$ , the ratio of the power spectral densities of the signal and the noise, and  $W$  is the bandwidth.

The upper bound method in this class [5] involves repeating the stimulus many times and estimating the stimulus by averaging the responses at each discrete time step (i.e., bin),  $S(i) = E[R_n(i)]$ . The noise for each trial is estimated from the response by subtracting the mean signal,  $N_n(i) = R_n(i) - S(i)$ . The power spectrum of  $S$  and  $N$  are calculated and equation (2) is claimed as an upper bound because of the Gaussian stimulus distribution. While this would be true when the Gaussian additive noise assumption is true, it is not provably true in general. Borst and Theunissen [5] mention that one must check the correctness of the Gaussian distribution assumption, but also say explicitly that this technique can be used on both “spiking and nonspiking neuronal responses” [5, p.953].

To show the danger of applying this upper bound technique to spiking neural responses, consider an example where the input is a vector of  $M$  independent, identically distributed (iid) Gaussian random variables with  $S(i) \sim N(0, \sigma^2)$ . Each output bin is a Bernoulli random variable with  $p(R(i) = 1|S(i) \geq 0) = q$  and  $p(R(i) = 1|S(i) < 0) = (1 - q)$ . Though this is only a simple example, it could represent a spike response with probability  $q$  of firing in a bin when the input is positive and probability  $(1 - q)$  of firing when the input is negative. We can calculate  $H(R(i))$  and  $H(R(i)|S(i))$  analytically to find the true mutual information rate. Using the technique described above, the mean response vector is  $S(i) = E[R(i)] = \frac{1}{2}$ . The autocorrelation of  $S$  is a triangle function, and the resulting power spectrum is  $P_{SS}(f) = \text{sinc}^2(\pi f M)$ . The noise vectors are  $N_n(i) = R_n(i) - S(i)$ , with a power spectrum of  $P_{NN}(f) = \frac{1}{4}$ . The upper bound on the mutual information rate is calculated according to equation (2), and the results (calculated numerically) are plotted in figure 1(b). Though this quantity is claimed as an upper bound applicable to spike responses, it simply is not always valid for non-Gaussian (spiking) systems.

The lower bound method in this class [5, 7] involves calculating the coherence function  $\gamma^2(f) = |P_{RS}(f)|^2 / (P_{RR}(f)P_{SS}(f))$ , using the cross-power spectrum  $P_{RS}(f)$ . In the additive Gaussian noise case,  $\text{SNR}(f) = \gamma^2(f)/(1 - \gamma^2(f))$ , and the mutual information rate is found from equation (2). This method is claimed as a lower bound to the true mutual information because it only accounts for linear correlations between  $S$  and  $R$  [7, p.967], but no proof is given that it holds for systems beyond those that simply add Gaussian noise.

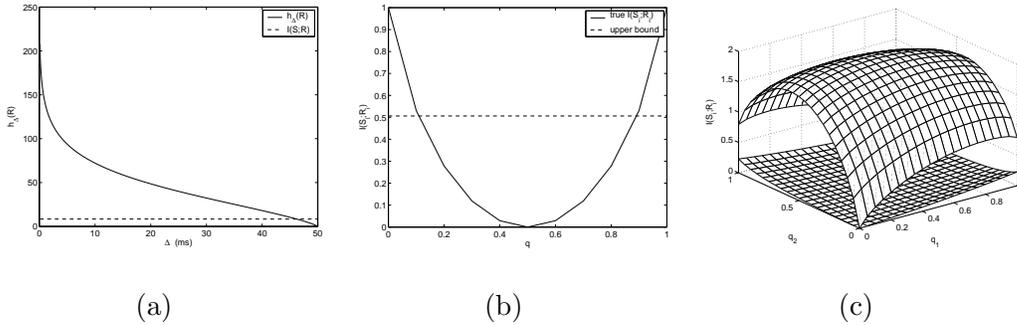


Fig. 1. (a) Binned entropy doesn't converge as  $\Delta \rightarrow 0$  and it is not always an upper bound (b) The Gaussian-based upper bounding technique is violated in this example. (c) The Gaussian-based lower bound can be a trivial lower bound.

At the very least, the best linear fit for the correlation can be zero even though there is significant correlation present. Consider a modified version of the previous example, where the input is again a vector of  $M$  (iid) Gaussian random variables with  $S(i) \sim N(0, \sigma^2)$ . Each output bin is a Bernoulli random variable with  $p(R(i) = 1 | |S(i)| \geq \eta) = q_1$  and  $p(R(i) = 1 | |S(i)| < \eta) = q_2$ , where  $\eta$  is an arbitrary threshold. Again, this is only a simple example, but it could represent a spike response with probability  $q_1$  of firing in a bin when the input magnitude is above a threshold and probability  $q_2$  of firing otherwise.  $H(R(i))$  and  $H(R(i)|S(i))$  are again readily calculable, and so the true mutual information rate  $I(S; R)$  is known. In this case, the coherence function essentially boils down to  $\gamma = \text{Cov}(R(i), S(i)) / \sqrt{\text{Var}(R(i))\text{Var}(S(i))}$ , and the resulting lower bound is plotted in figure 1(c). When  $q_1 = q_2$ , this lower bound method produces the trivial lower bound of zero. An analysis based on this bound appears to show that the mutual information is greatest when  $|q_1 - q_2| \approx 1$  and smallest when  $q_1 = q_2$ . The true mutual information actually has the *opposite* characteristics, being largest when  $q_1 = q_2$  and smallest when  $|q_1 - q_2| \approx 1$ . If only this lower bound were used to analyze the system, the results would be very misleading.

## 4 Conclusions

Given a stimulus set  $S$  and an associated probability distribution  $p(S)$ , mutual information measures the dependence between that specific stimulus source and the spike response  $R$ . However, accurate estimation of  $p(R|S)$  requires many stimulus presentations, making mutual information difficult to calculate directly. This is especially true with responses containing complex temporal and inter-neuron dependencies. If mutual information is directly estimated, techniques have been developed to remove estimation bias and to decompose

the measurement into quantities relevant to neural coding [9].

We examined two upper bound techniques in this report (entropy and the Gaussian noise model) that are in fact not bounds at all. We could not prove whether the coherence-function approach provides a lower bound or not. Estimating the stimulus from the response does indeed provide a lower bound on mutual information, but it is estimator dependent, which makes its approximation accuracy very problem and estimator dependent. Additionally, both lower bounds methods can produce the trivial lower bound of zero. Mutual information plays an essential role in an information theoretic analysis of communication and processing systems. Because it is difficult to estimate it accurately, we need better estimators or bounds than currently exist. Our experience with Gaussian-based estimators indicate that better bounds would probably results if they reflect the nature of spiking responses produced by both single and multiple neurons.

## References

- [1] C. Shannon, W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, IL, 1949.
- [2] D. Johnson, Four top reasons mutual information does not quantify neural information processing, in: *Proceedings of The Annual Computational Neuroscience Meeting, 2002*, Chicago, IL, 2002.
- [3] T. Cover, J. Thomas, *Elements of Information Theory*, John Wiley & Sons, Inc., New York, NY, 1991.
- [4] T. Berger, *Rate Distortion Theory*, Prentice Hall, Englewood Cliffs, NJ, 1971.
- [5] A. Borst, F. Theunissen, Information theory and neural coding, *Nature Neuroscience* 2 (11) (1999) 947–957.
- [6] F. Rieke, D. Warland, R. de Ruyter van Steveninck, W. Bialek, *Spikes: Exploring the neural code*, MIT Press, Cambridge, MA, 1997.
- [7] M. Juusola, A. S. French, The efficiency of sensory information coding by mechanoreceptor neurons, *Neuron* 18 (1997) 959–968.
- [8] P. Dayan, A. Abbott, *Theoretical Neuroscience*, MIT Press, Cambridge, MA, 2001.
- [9] G. Pola, A. Thiele, K. Hoffmann, S. Panzeri, An exact method to quantify the information transmitted by different mechanisms of correlational coding, *Network: Computation in Neural Systems* 14 (2003) 35–60.
- [10] P. Brémaud, *Point Processes and Queues*, Springer Series in Statistics, Springer-Verlag, New York, 1981.