INSTITUTE OF PHYSICS PUBLISHING NETWORK: COMPUTATION IN NEURAL SYSTEMS

Network: Comput. Neural Syst. **14** (2003) 1–4 www.iop.org/Journals/ne

EDITORIAL

Information and Statistical Structure in Spike Trains

Special Issue featuring selected papers derived from two workshops on Information and Statistical Structure in Spike Trains held during Neural Information Processing Systems 2000 and 2001

One of the most challenging and fascinating problems in science is deciphering how neural systems encode information. Mathematical models are critical in any effort to determine how neural systems represent and transmit information. Such mathematical approaches span a wide range, which may be divided approximately into two kinds of research activities. The first kind of activity uses detailed biophysical models (e.g., Hodgkin–Huxley and its variants) of individual neurons, detailed biophysical models of networks of neurons, or artificial neural network models to study emergent behaviors of neural systems. The second kind of activity develops signal-processing algorithms to analyze the ever-growing volumes of data collected in neuroscience experiments.

In an ideal scientific investigation there is a direct link between experiments and theoretical modeling. The theoretical models make predictions that can be used to guide experiments; the experiments provide data that allow for refinement (or rejection) of theoretical models. The growing complexity of neuroscience experiments makes use of appropriate data analysis methods crucial for establishing how reliably specific system properties can be identified from experimental measurements. Thus, careful data analysis is an essential complement to theoretical modeling. It allows validation of theoretical model predictions and provides biologically relevant constraints and parameter values for further analytic and simulation studies (see figure 1).

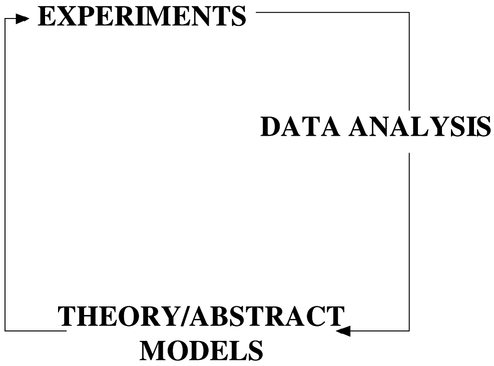

Figure 1. Neuroscience data: dynamic and multivariate.

Neural spike train data have special features that present new, exciting challenges for signal processing research. For this reason in 2000 and 2001 we organized two workshops entitled 'Information and Statistical Structure in Spike Trains' as part of the Neural Information

Processing meetings for those years. The papers in this special issue are a compilation of some of the work presented at those two workshops. All focus in some way on the question of how to decipher the neural codes. In this regard, the broad range of topics addressed is a sample of the breadth of issues required to make the link from experiments to theory, and of the theoretical considerations that can make experimental predictions. The data analysis methods divide into three types: data pre-processing (spike sorting), statistical modeling of neural spike train data, and algorithms for calculating information with explicit and implicit models. The system-dependent or theoretical models use simulations to gain physiologic and mechanistic insight.

The recent advent of the capability to record with multiple electrode arrays the simultaneous spiking activity of many neurons (>100) has made it possible to study information encoding by ensembles rather than by just single neurons. Simultaneous recording of multiple neurons is now a standard tool in neuroscience research. Often, the first data-analysis problem confronting the investigator is that of discerning the discharges of individual neurons within the multichannel, noise-contaminated signals that result from tetrode recordings. Nguyen, Frank and Brown's contribution, 'An application of reversible-jump MCMC to spike classification of multi-unit extracellular recordings', presents a new approach to this problem. The core of the approach is a Bayesian Markov chain Monte Carlo procedure that simultaneously estimates the correlation structure of the noise (the variability of the individual spike waveforms) and the number of distinct waveforms. In this way the authors determine the number of neurons being simultaneously recorded while assigning each action potential to its source neuron.

Kass, Ventura and Cai ('Statistical smoothing of neuronal data') take a careful look at the notion of firing rate. Though seemingly a simple problem, estimation of the firing rate (and, using this estimate to determine quantities such as the time of the peak response) is far from straightforward. As they show, straightforward approaches based on the post-stimulus histogram and raster plot—visualization tools that are ubiquitous in the exploratory analysis of neural data—are rather inefficient, and smoothing techniques based on adaptive splines offer substantial advantages.

Information theory methods are perhaps the most widely used techniques for analyzing neural data. In 'An exact method to quantify the information transmitted by different mechanisms of correlational coding', Pola, Theil, Hoffman and Panzeri extend their previous work on exact measures of information encoded by an individual neuron to exact measures of information encoded by a population of neurons. Their analysis technique allows a decomposition of the information structure in terms of the mean neural response, the correlation among the neurons and stimulus-induced changes in correlation among the neurons.

Dimitrov, Miller and co-workers ('Analysis of neural coding using quantization with an information-based distortion measure') present another approach for the estimation of information transmitted by spike trains, and apply this approach to the cricket cercal (air velocity sensation) system. Rather than calculate information directly by estimation of joint input–output probabilities, their approach applies rate distortion theory to identify a 'codebook' that minimally distorts the information available in the stimulus. One advantage of this approach is that it does not assume that the neural code resembles a rate code. More importantly, along with the estimate of information that the procedure yields, the codebook provides insight into how information is represented.

Understanding the distinction between single spikes and spikes that belong to bursts is crucial for characterizing neural encoding schemes. In 'Information encoding and computation with spikes and bursts', Kepecs and Lisman use simulation studies based on the Hodgkin–Huxley model combined with principal components analysis (PCA) and discriminant analysis to address this question. Specifically, the authors simulate a Hodgkin–Huxley model of a bursting neuron stimulated by stochastic inputs and study the relation between spiking patterns and stimulus features using the PCA discriminant analysis applied to a form of the spike-triggered average covariance matrix. The authors suggest a way to identify the distinct stimulus features that spikes and bursts encode.

Dynamic synapses (synapses whose efficacy increases or decreases in a manner that depends on their recent activity) clearly are differentially affected by isolated spikes and by spikes in bursts. Pantic, Torres and Kappen (‘Coincidence detection with dynamic synapses’) shows that dynamic changes in synaptic efficacy not only affect the way that spike trains from a single input are processed, but can also play an important role in the way that spike trains from convergent inputs interact. Via a computational study of idealized integrate-and-fire and Hodgkin-Huxley neurons, this work demonstrates that synaptic dynamics substantially improve the range over which neurons can act as coincidence detectors.

Although action potentials are the primary way in which neurons communicate, a crucial part of understanding this communication process lies in understanding the intricacies of the subthreshold processes that lead to the generation of action potentials. In ‘Influence of subthreshold nonlinearities on signal-to-noise ratio and timing precision for small signals in neurons: minimal model analysis’, Svirskis and Rinzel use an elementary integrate-and-fire model to examine the different roles that subthreshold voltages and time-dependent conductances play in signal integration and the production of action potentials. The key to their analysis is the role that a non-inactivating low-threshold outward current can play in increasing the precision of small signal integrations. Svirskis and Rinzel provide several examples to illustrate the importance of the subthreshold feedback mechanism.

Defining what a neuron encodes (i.e., its receptive field properties) is a basic question in neuroscience. Many experiments in neuroscience allow investigators to study this question by providing a statistical characterization of the response of the neuron to a given stimulus. In ‘Likelihood approaches to sensory coding in auditory cortex’, Jenison and Reale use likelihood methods to study the problem of sound localization based on the ensemble response recorded from primary motor cortex. They do this by using an inverse Gaussian probability density to model the neural response latency as a function of multiple acoustic parameters. The interesting feature of this approach is the use of likelihood methods based on formal probability model to carry out the analysis. There are several advantages to the likelihood approach. The parameter estimates obtained using the likelihood approach have several optimality properties such as consistency (converging to the true value as the sample size increases), having an asymptotic Gaussian distribution and providing a straightforward way to compute confidence intervals that are as short as possible. This gives a quantitative description of the relative importance of direction, azimuth and sound amplitude in inducing auditory neural responses and hence useful insight into how the neurons use this information for sound localization.

Most sensory systems are faced with the problem of providing useful signals over a wide dynamic range of the input, within the constraints of a relatively narrow range of outputs. In vision, this problem is particularly acute: the retinal output, which consists of spike trains that rarely exceed 200 impulses/s, can signal contrasts as low as one part in 300, over a 10^{10}-fold operating range of intensities—performance that can only be accomplished through gain controls. Lesica, Boloori and Stanley (‘Adaptive encoding in the visual pathway’) present a promising approach to the analysis of such gain controls, through a procedure that adaptively tracks response characteristics. Although described in the context of data from the visual system, the approach is a general one, and is particularly useful in the challenging situation in which the timescale of the adaptive mechanisms and the timescales of the signals being encoded are not well separated.

In sum, the availability of new experimental techniques, analytic and computational strategies, and the hardware with which to implement them has led to a surge of activity at the confluence of experimental, computational and theoretical neuroscience. The papers in this special issue provide windows into the range and vigor of this current research.

Jonathan D Victor *Weill Medical College of Cornell University, New York*
Emery N Brown *Harvard Medical School / Massachusetts Institute of Technology*

Guest Editors