# Neural System Identification

Garrett B. Stanley, PhD [1]
Associate Professor of Biomedical Engineering
Division of Engineering and Applied Sciences
Harvard University

October 20, 2003

[1]321 Pierce Hall, 29 Oxford St., Harvard University, Cambridge, MA 02138 USA, Phone: (617) 496-5368, Fax: (617) 496-0601, Email: gstanley@deas.harvard.edu

# 1 Introduction

One could argue that all scientific problems can be described in terms of two fundamental objectives: identification and control. Much of the current body of research in basic neuroscience revolves around the problem of identification, although not formally posed as such. The problem of identification is that of cause and effect. For example, in considering the relationship between two synaptically connected neurons, how does the pre-synaptic action potential cause the post-synaptic potential? At a more macroscopic level, how do the photons of light entering the eye cause the neuronal population activity in the visual pathway of the brain? Due to the overwhelming complexity of the nervous system, it is in fact difficult to think of threads of investigation that are not in some way reliant on identification or modeling at the systems level. The concept of system identification goes beyond simply reporting experimental observations. In many cases, the input can be controlled, and the goal is to identify a functional relationship between stimulus and response that will enable prediction of the response of the system to subsequent arbitrary inputs. Failure in prediction exposes previous misconceptions about the underlying dynamics, often leading to more intelligently designed experiments, and so on. Herein lies the true value of the identification process in this largely empirical field of science.

The field of system identification grew out of the statistics and engineering literature in the 1960s, motivated by the need to predict and control the behavior of complex systems (Box and Jenkins, 1976). Subsequently, there have been a number of general references that range from the applied (Jenkins and Watts, 1968; Ljung, 1987) to the theoretical (Söderström and Stoica, 1987; Brillinger, 1981) ends of the spectrum, as well as those which explicitly focus on the identification of physiological systems (Marmarelis and Marmarelis, 1978). Central to the framework of system identification is the idea that complex systems can be represented as a *black box*, as opposed to more simplistic physical systems whose dynamics suggest relationships based on first principles. As more is learned about a complex system, the goal then becomes to reduce the relationship to smaller and smaller black boxes that represent the individual components of the system, until a sufficient level of detail has been achieved. Neuronal processing of information is complex at all levels, from the microscopic interaction between the pre- and post-synaptic cell, to the macroscopic interactions between the large populations of neurons involved in sensory processing and motor response. Since the 1960s, system identification techniques have been formally applied to processing of neuronal information in a number of studies (Perkel et al., 1967; DeBoer and Kuyper, 1968; Marmarelis and Naka, 1973a; Marmarelis and Naka, 1973b;

Brillinger et al., 1976), and has subsequently become a commonly utilized tool in systems neuroscience. The goals of this chapter are to introduce some important perspectives and techniques for system identification, and to present concrete examples of system identification strategies employed in sensory processing in the central nervous system and neural control in the peripheral nervous system.

The remainder of the chapter is presented in the following manner. Section 2 provides an introductory review of techniques in system identification. The basic characteristics of dynamical systems are discussed, as well as means for estimating the fundamental quantities relating input and output, in both time and frequency domains. The background on system identification, however, is provided for continuous signals. Neurons communicate information through discrete pulses, or action potentials. Section 3, therefore, provides the necessary details on the nature of measured neuronal signals and how they can be represented mathematically, as well as some basic correlation measures for these type of data. Section 4 presents examples of neuronal system identification in the context of the processing of visual information. Specific examples will be given for cases in which neuronal activity is well-predicted by quasi-linear models estimated from stimulus-response data and cases in which the properties of the pathway that continuously change in response to a changing environment can be estimated using adaptive estimation strategies. Section 5 provides examples of the identification of neuronal processing mechanisms in the mammalian somatosensory (touch) pathway, specifically highlighting interactions that are not well explained through linear techniques. Finally, in contrast to the time-domain estimation presented in previous sections, section 6 provides an example of the frequency domain identification of the dynamics related to the peripheral nervous control of cardiac function.

## 2 System Identification

Before presenting issues related to the identification of dynamics related to neuronal processing, it is first necessary to discuss preliminaries regarding system identification in general. The following discussion will focus on discrete-time sampling of continuous-valued signals, but will, in subsequent sections, be extended to capture the discrete nature of neuronal spiking.

## 2.1 Dynamical Systems

In general, problems of system identification revolve around a conceptual picture of an input-output dependence, as shown in Figure 1. The output or response $r(t)$ is in some way a function

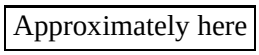


Figure 1: Input-output transformation.

of the input or stimulus $s(t)$ (present and past), but is also influenced by an uncontrolled, and typically unobserved, noise process $v(t)$, which describes all behavior of $r(t)$ not directly explained by $s(t)$. The following questions are then applicable. Is the relationship between $s$ and $r$ linear or nonlinear? Is the relationship between $s$ and $r$ time-invariant, or time-varying? In other words, will an experiment conducted at time $t$ yield the same results as an experiment conducted at $t + \Delta t$? For the rest of this section, we will assume that the properties are time-invariant, but will revisit this issue in the discussion of adaptation in the visual system in section 4.

When encountering a system in which little is known about the underlying input/output relationship, it is often prudent to begin with non-parametric models which minimally constrain the description of the dynamics. A general non-parametric nonlinear relationship between input, $s$, and output, $r$, can be expressed as a Volterra series (Volterra, 1959; Marmarelis and Marmarelis, 1978):

$$r(t) = g_0 + \sum_k g_1(k)s(t-k) + \sum_k \sum_j g_2(k,j)s(t-k)s(t-j) + \text{HOTS} + v(t) \qquad (1)$$

where $g_n$ is the "nth order kernel", HOTS represents *higher order terms*, and $v(t)$ is the noise process (assumed additive). The term $g_0$ is representative of the mean of $r(t)$ not related to the input, $g_1$ is representative of first order dynamics, $g_2$ of second order interactions, and so on. Note that this is analogous to a Taylor series expansion describing arbitrary static nonlinear relationships. The following discussion will focus on non-parametric estimation, where the input/output relationship takes on no more specificity than that of the expression in equation 1, but there are numerous references that can be consulted for parametric estimation for more specific functional forms (Cramér, 1946; Rao, 1973).

## 2.2 Estimation

For the given model structure, the problem of system identification reduces to the minimization of a cost function that depends on observed data and the set of parameters, θ:

$$\hat{\theta} \quad = \quad \arg\min_{\theta} J(\text{data}, \theta) \tag{2}$$

where

$$\theta \triangleq [\ g_0 \quad g_1(0) \quad g_1(1) \quad \ldots \quad g_2(0,0) \quad g_2(1,0) \ldots\ ]^T$$

where $J(\cdot)$ is a cost function, $T$ represents matrix transpose, $\hat{}$ represents estimate, and θ is a vector containing elements of the model. The goal then is to find the set of parameters θ which minimize the cost function for a given data set. The following measures are presented in the context of the class of minimization problems based on least-squares cost functions.

### 2.2.1 Time Domain Measures

There are a number of statistical measures that are important to define for the identification process, several of which are presented here.

*Covariance.* The auto-covariance function is a symmetric quantity that refers to how a signal, *s*, co-varies with itself (i.e. relationship between signal at different time points):

$$C_{ss}(k) \quad \triangleq \quad E\{(s(t)-\mu_s)(s(t+k)-\mu_s)\}$$

where $E\{\cdot\}$ denotes statistical expectation, and $\mu_s$ denotes the mean of *s*. This function can be estimated in the following manner:

$$\hat{C}_{ss}(k) \quad = \quad \frac{1}{N-|k|-1}\sum_{t}(s(t)-\hat{\mu}_s)(s(t+k)-\hat{\mu}_s) \qquad -\frac{N}{4} < k < \frac{N}{4}$$

where *N* is the length of the data set, and $\hat{\mu}_s$ is the mean estimated from data. The covariance estimates should not be computed for lags (shifts) greater than a quarter of the data length to avoid significant errors resulting from the limited number of data points involved in the estimate (Chatfield, 1989). The interpretation is that for a given value of *k*, the signal *s* is shifted by *k* points, and the point-by-point product of the shifted observations is formed, revealing periodicity or temporal structure in the signal. Suppose that $s(t)$ is an independent, identically distributed (IID) process (and therefore uncorrelated). The approximate 95% confidence intervals on an

independent process is represented by the band $\pm 2\hat{C}_{ss}(k)/\sqrt{N}$. This band is often superimposed on the plot of the auto-covariance function to evaluate whether the observed correlation structure deviating from zero is statistically significant. A boot-strapping technique can alternately be employed, which involves repeatedly computing covariances from the randomized (temporally shuffled) data set to estimate the confidence bands on an uncorrelated process (Perkel et al., 1967).

When analyzing the relationship between input $s(t)$ and output $r(t)$, it is often useful to consider the cross-covariance function:

$$C_{sr}(k) \quad \triangleq \quad E\{(s(t)-\mu_s)(r(t+k)-\mu_r)\}$$

which can be estimated in a manner similar to the auto-covariance estimate presented previously. For the case where $s(t)$ and $r(t)$ are independent, and individually uncorrelated IID processes, the band $\pm 2\sqrt{\hat{C}_{ss}(0)\hat{C}_{rr}(0)}/\sqrt{N}$ represents the approximate 95% confidence interval. We can thus assess to what extent the two processes are significantly correlated (Chatfield, 1989).

*Kernel estimation.* In some very special cases, the input/output relationship is well explained by a linear system (i.e. the first order kernel). In this case, the cross-covariance between the input and the output is related to the auto-covariance of the input through the first order kernel (Papoulis, 1984):

$$C_{sr}(k) \quad = \quad \sum_{m=0}^{L} g_1(m) C_{ss}(k-m) \tag{3}$$

Using the fact that $C_{ss}(k) = C_{ss}(-k)$, this relationship can be written in a structured matrix form as:

$$\begin{bmatrix} C_{sr}(0) \\ C_{sr}(1) \\ \vdots \\ C_{sr}(L) \end{bmatrix} = \begin{bmatrix} C_{ss}(0) & C_{ss}(1) & \dots & C_{ss}(L) \\ C_{ss}(1) & C_{ss}(0) & \dots & C_{ss}(L-1) \\ \vdots & \vdots & \ddots & \vdots \\ C_{ss}(L) & C_{ss}(L-1) & \dots & C_{ss}(0) \end{bmatrix} \begin{bmatrix} g_1(0) \\ g_1(1) \\ \vdots \\ g_1(L) \end{bmatrix}$$

or $C_{sr} = C_{ss} \cdot g_1$ in shorthand notation. We can then solve for $g_1$:

$$\hat{g}_1 \quad = \quad C_{ss}^{-1} C_{sr}$$

where the first order kernel is simply the auto-covariance between input and output, normalized by the correlation structure of the input. Note that for a linear system, the first order kernel is also known as the *impulse response* of the system, as it represents the response to an impulse

input. The first order kernel, when estimated in this manner, is precisely the solution to the least-squares problem. The cost function of equation 2 is the sum of the squared errors:

$$\hat{g}_1 \quad = \quad \arg\min_{g_1} \sum_k \left( r(k) - \sum_m g_1(m) s(k-m) \right)^2$$

If the input is uncorrelated (white), then $C_{ss}$ will be a diagonal matrix, with $\sigma_s^2$ (the variance of $s$) along the diagonal. We can therefore write $\hat{g}_1(k) = C_{sr}(k)/\sigma_s^2$. In this case, the first order kernel is simply the cross-covariance between input and output, normalized by the variance of the input. The application of this estimation technique will be discussed in the context of processing the visual pathway in section 4.

*Higher-order interactions.* It should be noted that even if the system is not strictly linear, and thus has nonzero kernels beyond the first order, the use of an orthogonal input, such as Gaussian white noise, makes possible the sequential estimation of kernels of different order (Lee and Schetzen, 1965; Marmarelis and Marmarelis, 1978). The higher-order kernels, in this case, can be estimated from higher-order correlations formed from averaging the products of the signals at varying delays. See (Marmarelis and Marmarelis, 1978) for a detailed discussion of this issue. This topic will again be revisited in the discussion of the nonlinear encoding of tactile information in section 5.

### 2.2.2 Frequency Domain Measures

It is often the case that the input-output relationships are better described (or estimated) in the frequency domain, as opposed to the time domain. Several of the related measures are presented here.

*Spectra.* The power spectrum of a signal, $s$, is defined as the Fourier transform of the auto-covariance:

$$S_{ss}(\omega) \quad \triangleq \quad \sum_{k=-\infty}^{\infty} C_{ss}(k) e^{-j\omega k} \in \mathbb{R}$$

which can be estimated from the Fourier transform of the estimated auto-covariance, where $\omega = 2\pi f$ is the frequency in rad/sec. The interpretation of the power-spectrum is that it is a frequency-dependent variance measure, describing how the variance is distributed across different frequency bands. The Fourier transform of the auto-covariance estimate is a rather raw estimate for the spectrum for finite observations. It is therefore prudent to modify this estimate

of the spectrum by multiplying by a smoothing window, $w$ (Ljung, 1987):

$$\hat{S}_{ss}(\omega) \quad = \quad \sum_{k=-\infty}^{\infty} w(k)\hat{C}_{ss}(k)e^{-j\omega k}$$

Note that multiplication by a smoothing window in the time domain is equivalent to convolving with the smooth function in the frequency domain. See Ljung for a detailed discussion of smoothing windows and the corresponding effects on the spectral estimates (Ljung, 1987). Confidence bands can be generated for a white process, and used as a test for whiteness of the observed process. An alternate use is to place bands around the actual estimate $\hat{S}_{ss}(\omega)$ so that the spectrum can be statistically compared to other spectra (Jenkins and Watts, 1968; Brillinger, 1981).

The cross-spectrum between two processes $s(t)$ and $r(t)$ is defined as the Fourier transform of the cross-covariance:

$$S_{sr}(\omega) \quad \triangleq \quad \sum_{k} C_{sr}(k)e^{-j\omega k} \in \mathbb{C}$$

which can be estimated from the Fourier transform of the estimated cross-covariance. Similar confidence bands exist as above for assessing correlations as a function of frequency (Brillinger, 1981).

*Transfer Function.* For linear systems, the transfer function between input and output is defined as the Fourier transform of the impulse response:

$$G(\omega) \quad \triangleq \quad \sum_{k} g_1(k)e^{-j\omega k} \in \mathbb{C}$$

which can be obtained in practice from the impulse response estimate, previously discussed. An alternate method of computing the transfer function estimate is derived from the relationship between the auto-covariance of the input and the cross-covariance between input and output. Since convolution in the time domain is equivalent to multiplication in the frequency domain, the relationship in equation 3 yields:

$$S_{sr}(\omega) \quad = \quad G(\omega)S_{ss}(\omega)$$

Estimates of the quantities $S_{ss}$ and $S_{sr}$ then directly yield an estimate $\hat{G}$. The impulse response, $g_1$, can be obtained by taking the inverse Fourier transform of the resulting transfer function $G$. This has obvious advantages, since the time-domain estimation of $g_1$ involves matrix inversion, as previously described, and the frequency domain estimation simply involves the Fourier

transform, which can be implemented efficiently using a variety of FFT algorithms (Press et al., 1992). Estimation in the frequency domain will again be discussed in the context of neural control of cardiac function in section 6.

# 3 Representations of Neuronal Activity

The neuron is the fundamental building block of our nervous system. Neurons encode information about the outside world through subthreshold membrane potentials and sequences of discrete electrical events. Various measurement strategies thus exist which produce a myriad of representations of the underlying neural activity, from the micro- to the macroscopic spatial scales (Eichenbaum and Davis, 1998). The measures we will focus on here relate to the suprathreshold spiking activity of the neurons, with respect to both the precise timing and rate at which they occur in response to exogenous stimuli.

### 3.0.3 Spike Times

When observing the activity of a single neuron over time, the neuron undergoes a sequence of action potentials that are typically uniform in their magnitude and shape. By representing the times of these events as delta functions, or *spikes*, highly localized in time, we describe this as a *spike train*. For a more extensive discussion regarding this topic, see (Dayan and Abbott, 2001).

Suppose that the events occur at time $t_i$, with $i = 1, 2, \ldots, n$. The spike train can be represented as a continuous signal, with Dirac delta functions at the times of the events:

$$\rho(t) \quad = \quad \sum_{i=1}^{n} \delta(t - t_i)$$

The response, $\rho(t)$, is therefore a sequence of impulses at the event times. Such a train of events is shown in Figure 2a.

Approximately here

Figure 2: The spike train and various transformations. (**a**) A neuronal spike train, over 1 second. (**b**) The firing rate obtained from the single spike train in (a), by convolving with a rectangular window of width 10ms. (**c**) Raster plot of repeated trials, showing the variability of the response. (**d**) Firing rate obtained from repeated trials, by summing across trials and convolving with a rectangular window of width 10ms.

*Intensity measures.* For discrete processes, such as the firing of a neuron, often termed *point*

*processes*, it is natural to describe the process through intensity measures. We now turn our attention to correlation measures for point processes (Brillinger et al., 1976; Brillinger, 1992). Let the mean intensity of the process $x$ be defined as:

$$m_x \triangleq \lim_{h\to 0} \frac{\Pr\{x \text{ event in } [t,t+h)\}}{h}$$

where $\Pr\{\cdot\}$ denotes probability, and $h$ is a binwidth, which goes to 0 in the limit. For an observation over $[0,T)$, we can estimate this as the number of events in $[0,T)$, divided by the total time interval $T$. The auto-intensity function, much like the auto-covariance function for continuous processes, is defined as:

$$m_{xx}(u) \triangleq \lim_{h\to 0} \frac{\Pr\{x \text{ event in } [t,t+h) | x \text{ event at } t-u\}}{h} \tag{4}$$

Note that in practice, the binwidth $h$ is finite, making the measures somewhat resolution dependent. The heuristic description of the above measure is that given a spike at time $t-u$, it is the probability of observing a spike in a small window of width $h$ at time $t$, normalized by the width of the window. Now, in addition to the process $x$, consider another point process $y$. The cross-intensity function, much like the cross-covariance function for continuous processes, is defined as:

$$m_{xy}(u) \triangleq \lim_{h\to 0} \frac{\Pr\{y \text{ event in } [t,t+h) | x \text{ event at } t-u\}}{h} \tag{5}$$

For all of the above estimators, there are corresponding error bounds that are presented in a number of references (Brillinger et al., 1976; Brillinger, 1992). These measures will be revisited in section 5 in the discussion of nonlinear encoding of tactile information.

### 3.0.4 Firing rate

Suppose that we are interested in the rate, $r$, at which spikes occur over some experimental trial interval $T$, which is equivalent to the mean intensity. This can be written as:

$$r = \frac{1}{T}\int_0^T \rho(\tau)d\tau$$

For example, in Figure 2a, we observe 54 spikes in 1 second, resulting in a mean rate of 54Hz. We can extend this to a more general expression which represents the firing rate in small fixed intervals of width $\Delta t$:

$$r(k\Delta t) = \frac{1}{\Delta t}\int_{k\Delta t}^{(k+1)\Delta t} \rho(\tau)d\tau \qquad k = 0,1,2,\ldots \tag{6}$$

This is essentially just binning and counting spikes within each bin. We will refer to this hereafter as simply $r(k)$, where the sampling interval is implicit. Note that this measure depends strongly on the size and placement of the bins. More generally we can express the rate as a convolution of the original spike train with a smoothing window $w(t)$. A rectangular function $w(t)$ of width $\Delta t$ and height $1/\Delta t$ gives the number of events that occurs in the interval $[t, t+\Delta t]$, normalized by the interval length $\Delta t$, which we will denote the *firing rate*, shown in Figure 2b for $\Delta t = 10$ms. Although this approach alleviates the problem of bin placement, it introduces correlation into the process $r(t)$, since adjacent bins overlap and the same spikes are counted in both.

### 3.0.5 Neuronal Variability

The discussion thus far might lead one to believe that an individual neuron has very complex dynamics that simply need to be fully characterized. However, when presented with the same inputs over repeated trials, most neurons exhibit a significant degree of variability in their response. To illustrate this, Figure 2c shows a *raster* plot of the activity across different trials. For each trial, the spike times are plotted as a dot at the time of occurrence, whereas each row of the plot represents a different trial. This is an effective way of plotting neuronal data, in order to see repeatable structures in the activity and the variability across trials. The firing rate in this case can be determined by aligning the responses of each trial relative to the beginning of the individual trial, and averaging across the trials in a temporal bin. More precisely, let $t_{i,k}$ denote the $i$th spike in the $j$th trial, where $j = 1, 2, \ldots, N$ and $i = 1, 2, \ldots, n_j$. Further, let $\rho_j(t) = \sum_i \delta(t - t_{i,j})$ denote the spike train of the $j$th trial, where $t$ is relative to the beginning of the $j$th trial. We then have:

$$r(k\Delta t) \quad = \quad \frac{1}{\Delta t} \int_{k\Delta t}^{(k+1)\Delta t} \rho(\tau) d\tau \quad \text{where} \quad \rho(t) = \frac{1}{N} \sum_{j=1}^{N} \rho_j(t)$$

As before, the firing rate can also be obtained by convolving $\rho(t)$ with a smoothing window, the results of which are shown in Figure 2d, again for $\Delta t = 10$ms

We now turn to a series of examples that utilize the preceding perspectives and techniques.

## 4 Neuronal Encoding in the Visual Pathway

Humans are incredibly visual creatures. As a result, the visual pathway of the mammalian brain has been the focus of a significant amount of research in sensory coding. The visual pathway therefore serves as a good substrate upon which to discuss the systems approaches presented thus far. Photons of light from the outside world enters the eye through the lens, and fall upon

the back of the eye, or retina. Photoreceptors transduce the photons into electrical signals, which are propagated through the layers of the retina, to the retinal ganglion cells, serving as the output layer of the retina. Action potentials that originate in the retina travel down axon bundles (the optic nerve), projecting to the visual region of the thalamus (the lateral geniculate nucleus, or LGN). The LGN then projects to the primary visual cortex (Hubel, 1995). In the early visual pathway, each neuron encodes information about a restricted region of visual space, which is generally referred to as the receptive field (RF) of the cell. More precisely, however, we will refer to the spatiotemporal receptive field (STRF) as the functional manner in which visual information is integrated over space and time to give rise to neuronal activity in the pathway, which is an explicit description of the input/output properties of the stages of processing in the pathway.

The fundamental task of identification problems in the visual pathway lies in the characterization of the relationship between a temporally continuous stimulus and the discrete process of the firing activity of the neuron at various locations in the pathway. One approach is to relate the continuous input (modulated light intensity) to the firing "rate" of the neuron, allowing the quantification of the relationship between the input and its modulatory effects on the rate of action potentials generated. Let the firing rate of the neuron be denoted $r(k)$, which represents the number of events occurring in the interval $(k\Delta t, (k+1)\Delta t]$ normalized by the interval width, $\Delta t$, as previously described in equation 6. Although the transformation to rate does simplify the relationship a great deal, the remaining relationship between the stimulus, which takes on values both positive and negative relative to the mean level, and the strictly non-negative firing rate is non-trivial. Linear models are insufficient to capture such a transformation, and thus generally must be described through a higher-order kernel expansion, as in equation 1.

An alternate functional model to describe the relationship between the stimulus and the firing rate of a neuron is the Linear-Nonlinear (LN) cascade, which incorporates a linear (L) system followed by a static nonlinearity (N) (Movshon et al., 1978; Tolhurst and Dean, 1990; Ringach et al., 1997; Meister and Berry, 1999; Chichilnisky, 2001; Stanley, 2002), as shown in Figure 3. The output of the linear element is expressed as a convolution of the temporal stimulus, $s$, with the first order kernel $g_1$, and an integration over the visual space:

$$x(k) = \sum_i \sum_j \sum_m g_1(i,j,m)s(i,j,k-m)$$

where $m$ ranges from 0 to $L$, and $L\Delta t$ is the filter length, which can be interpreted as the temporal integration window of the cell. The firing rate of the neuron is then a static nonlinear function

Approximately here

Figure 3: Linear-Nonlinear-Poisson (LNP) model of visual encoding (**a**) The linear-nonlinear-Poisson cascade. The first order kernel was computed from the random binary sequence (**b**) and the corresponding firing rate (in 7.7ms bins) of an X cell in the LGN (**c**). The temporal kernel at the center of the receptive field is shown in (**d**); band represents $\pm$ 2SD around estimate. The actual (positive) and predicted (negative) firing rates (in 40msec bins) from the full spatiotemporal receptive field model are shown in (**e**). Adapted with permission from (Stanley, 2002).

of the intermediate signal $x$, so that $r(k) = f(x(k))$, representing rectifying and saturating properties of the encoding. The output of the nonlinearity is then driving an inhomogeneous Poisson process to produce the discrete neuronal events. In contrast to the complex nature of a higher-order kernel representation, the LN system provides a relatively simple means for describing the inherent nonlinearity in neural encoding (Hunter and Korenberg, 1986; Greblicki, 1992; Paulin, 1993; Korenberg and Hunter, 1999; Chichilnisky, 2001).

### 4.1 Estimation of the STRF

Reverse-correlation techniques, which refer to the cross-correlation between stimulus and response for white noise stimuli, have been used extensively to characterize the dynamics of neurons in the retina, LGN, and primary visual cortex (Jones and Palmer, 1987; Reid et al., 1991; Mao et al., 1998), where there is an implicit assumption of the underlying LN cascade. The reverse-correlation technique closely mirrors the linear estimation techniques presented in section 2.2.1. Let the parameter vector be defined as $\theta \triangleq [\ g_1(0) \quad g_1(1) \quad \ldots \quad g_1(L-1)\ ]^T$ to represent the first order kernel. Here we represent a single visual pixel for simplicity, but the representations can easily be reformulated to represent a two-dimensional array, the elements of which can be restructured into the vector notation here. For this simple case, the response can then be written $r(k) = f(x(k))$, where $x(k) = \theta^T \varphi(k)$ and $\varphi(k) \triangleq [\ s(k) \quad s(k-1) \quad \ldots \quad s(k-L+1)\ ]^T$ is the time history of the stimulus. As discussed previously, the parameter vector for a first-order kernel can be estimated from the cross-covariance between the stimulus and the output of the linear element, normalized by the auto-covariance structure of the stimulus, or $\hat{\theta} = C_{ss}^{-1} C_{sx}$, where $C_{ss}$ is the structured auto-covariance matrix of the stimulus $s$, and $C_{sx}$ is the cross-covariance vector between the stimulus and the output of the linear block, $x$. For Gaussian inputs to the static nonlinearity, the only effect of the nonlinearity is a scaling of the cross-covariance, and the estimate of the parameter vector can be expressed as a function of the stimulus, $s$, and response,

$r$:

$$\hat{\theta} \quad = \quad C_{ss}^{-1} C_{sx} = A \cdot C_{ss}^{-1} C_{sr} \tag{7}$$

where $A$ is a constant of proportionality. For half-wave rectification with Gaussian inputs, this scaling will be approximately 2 (Stanley, 2002). Figure 3 shows the estimate of a low-pass biphasic first-order kernel at a single pixel at the center of the RF from cat LGN X cell response to a spatiotemporal binary stimulus (m-sequence). See (Stanley et al., 1999) for details of experimental methods. A segment of the m-sequence stimulus (at 128Hz) at the center pixel is shown in Figure 3b, and the corresponding firing rate (in 7.7ms bins) is shown in Figure 3c. The kernel estimate for the center pixel computed over the entire trial is shown in Figure 3d. The band represents two standard deviations around the estimator (Stanley, 2002). The uncertainty in the estimation (due to noise and unmodeled dynamics) is useful in comparing the encoding properties in different physiological states. Using the complete spatiotemporal kernel, the firing rate of the cell is predicted in Figure 3e; the actual firing rate is shown in the dark shaded (positive) region, whereas the predicted firing rate is shown in the gray shaded region, reflected about the horizontal axis for comparison. The cascade of the linear system with the static nonlinearity is a good predictor of the neuronal response for this relatively linear cell, under rather rigid conditions of stationary input statistics.

### 4.2 Adaptive Estimation

One of the major assumptions of the reverse correlation technique is that the dynamics of the underlying functional mechanisms in the visual pathway are time-invariant, resulting in spatiotemporal receptive field properties that remain unchanged with time. As early as the retina, however, adaptation mechanisms act on a continuum of time scales to adjust encoding dynamics in response to changes in the visual environment (Enroth-Cugell and Shapley, 1973; Shapley and Victor, 1978). Adaptation mechanisms have also been identified in the lateral geniculate nucleus (LGN) (Ohzawa et al., 1985) and cortical area V1 (Albrecht et al., 1984). An adaptive approach for the estimation of encoding properties from *single* trials can be utilized to specifically address time-varying neural dynamics (Brown et al., 2001; Stanley, 2002; Lesica et al., 2003). A recursive estimate of the kernel at time $t$ is posed as the least-squares estimate based on information up to time $t$ (Ljung, 1987). In this context, the recursive STRF estimation can be reformulated

as (Stanley, 2002):

$$\hat{\theta}_t \quad = \quad \arg\min_{\theta} \sum_{k=L}^{t} \lambda^{t-k}(r(k) - f(\theta^T \varphi(k)))^2$$

where $\lambda \in [0,1]$ is a weighting parameter, and the subscript $t$ denotes that the estimate is a function of time. The estimate of the kernel at time $t$ can be written as the estimate at time $t-1$ plus a correction based on new information arriving at time $t$:

$$\hat{\theta}_t \quad = \quad \hat{\theta}_{t-1} + \text{update} \tag{8}$$

where the update depends on the input-output covariance structure (Ljung and Söderström, 1983). For the explicit formulation for this example, see (Stanley, 2002). This estimate down-weights past information in an exponential manner as controlled by the size of $\lambda$, which is often called the "forgetting" factor. As the encoding properties vary over time, the input/output covariance changes properties, which is accounted for by a single weighting parameter, telescoping backwards in time; the result is an exponentially down-weighting of past information.

X cells in the cat LGN were stimulated from rest with the spatiotemporal m-sequence at 128Hz and 100% contrast over several minutes, inducing an adaptation to the sudden increase in contrast (or variance). Figure 4 illustrates typical results obtained using the adaptive estimation approach. For this ON cell, the kernel at the center of the RF exhibits a clear decrease in response magnitude

Approximately here

Figure 4: Evolution of adaptation in the visual pathway. Continuous tracking of adaptation in LGN X cells. (**a**) The temporal kernel between stimulus and firing rate for the center of the receptive field is plotted as a function of the time since stimulus onset (seconds) for an ON cell in the LGN. (**b**) The magnitude of the transfer function of the system at the beginning (thick) and end (thin) of the trial. Note that the magnitudes are normalized to emphasize the change in bandwidth. (**c-f**) Spatial RF of an OFF cell at peak in temporal kernel at 8, 16, 24, and 32 seconds after stimulus onset (region is approximately $1.8^{\circ}$ of visual space). Dark represents minimum values, bright maximum values. The RF map is normalized by the peak of the center pixel at each time slice. Adapted with permission from (Lesica et al., 2003).

over the first 10-30 seconds of the trial. The frequency response at the beginning and end of the trial is shown in 4b. In addition to changes in magnitude, the kernel shows a sharpening in bandwidth over the same time course, where the bandwidth is the range of frequencies for which the transfer function magnitude was greater than 50% of the peak value. Figures 4c through f illustrate the spatial RF map for an OFF cell at the peak in temporal response at 8 second intervals over the stimulus trial. The RF map is normalized by the peak of the center pixel at

each time slice. The extent of the spatial RF in figure 4c-f appears to increase due to the fact that the magnitude of the center pixel shows a greater decrease than that of the off-center pixels over the course the trial. The point to emphasize here is that these changes in encoding dynamics would not be discovered through traditional reverse-correlation techniques that assume a fixed relationship between stimulus and response over the trial.

Time-varying encoding properties are a ubiquitous characteristic of all sensory pathways. Adaptation has been studied for some time and is known to dramatically effect the temporal and spatial dynamics, while anecdotal evidence of modulatory effects from other brain regions on encoding properties has been reported for a number of stages in the visual pathway. However, it is not yet known what these phenomena imply for the coding strategies of the sensory pathway as a whole. The approach presented here is a first step towards capturing the evolution of spatiotemporal receptive field properties over a range of time scales, so that the pathway may be better understood in the context of the continually changing natural environment.

## 5 Nonlinear Encoding in the Somatosensory Pathway

The previous example focused on encoding mechanisms that are quasi-linear in their behavior. In many cases, however, neuronal encoding can be strongly non-linear, and thus not well explained through linear methods. An example of such nonlinear encoding in the rat somatosensory pathway is presented in this section. Rats and other rodents have arrays of facial whiskers (vibrissae) that are vital for survival; neonates deprived of their vibrissa exhibit severely impaired behavioral development (Carvell and Simons, 1996). By actively palpating objects, rats have also been shown to be able to discriminate between very similarly textured surfaces based on vibrissa exploration alone (Guic-Robles et al., 1989; Carvell and Simons, 1990), illustrating the exquisite sensitivity of this sensory modality.

In general, neurons in the vibrissa-related region of somatosensory cortex (historically referred to as the *barrel cortex*) do not have high spontaneous firing rates (Brumberg et al., 1996). Nonetheless, the statistical properties of the spontaneous activity can reveal much about the underlying functional properties of the cell. Suppose that we denote the neuronal spike train as a discrete process $y$. The auto-intensity function $m_{yy}(u)$, as defined in equation 4, for the spontaneous activity of three characteristic cell types is shown in Figure 5. The top panels of a-c represent the raster of spikes surrounding each spike of the trial for each cell, and the bottom panels represent

Approximately here

Figure 5: Correlation structure of spontaneous activity in barrel cortex. (**a-c**) For typical cells from three classes of observed activity, raster of spikes surrounding each spike of the trial (top), and the square-root auto-intensity function (bottom) for the spontaneous activity, with the null-band of an uncorrelated process shown with the gray band ($h$=25ms). The horizontal axis represents time relative to individual spike. Adapted with permission from (Stanley and Webber, 2003).

the corresponding square-root auto-intensity function (Stanley and Webber, 2003). It should be noted that these raster plots are different from traditional raster plots in that each row does not represent a separate trial, but instead represents a collection of the spikes that occur relative to a different individual spike *within* the same trial. The gray bar represents a 95% confidence interval around the mean level on an uncorrelated process. The first cell, shown in a, exhibits an intensity that is statistically different from the mean only at zero lag. This correlation structure is shared by the (memoryless) homogeneous Poisson process, which also exhibits an auto-intensity that is equal to the mean rate for non-zero lags and has a sharp positive peak at zero-lag. The second cell, shown in b, exhibits suppressive lobes for lags between 50 and 150 ms that extend out of the confidence band. The third cell, shown in c, exhibits a suppressive lobe for lags between 50 and 150 ms, an excitatory lobe between 170 and 200 ms, and an additional suppressive lobe from 270 to 310 ms. The post-excitatory suppression revealed by the auto-intensity measure has significant bearing on how the cell will respond to exogenous stimulation at the periphery. In particular, punctate stimulation at the periphery induces post-excitatory suppression that influences the response to a subsequent stimulus in a strongly nonlinear manner.

## 5.1 The Impulse Response and Nonlinear Encoding

When a whisker is mechanically deflected with a step-input to mimic contact with an object, the barrel neurons typically respond with a transient excitatory response after a short latency of 5-15ms, as shown in Figure 6a. It has been shown that the neurons in the barrel cortex are

Approximately here

Figure 6: Higher-order dynamics in tactile encoding. (**a**) Whisker deflection paradigm. A step input in deflection is an impulse in velocity, producing a transient cortical response $r(t)$, which is equivalent to the cross-intensity (CI) function or first-order kernel. (**b**) Cross-intensity functions (spikes/ms) for caudal followed by rostral stimulation at latencies of 100, 70, and 20 ms.

primarily sensitive to velocity, rather than displacement. The stimulus can therefore be thought

of as an impulse in velocity, and the neural activity is then the impulse response:

$$r(t) \quad = \quad \int_{\tau} g_1(\tau)s(t-\tau)d\tau = \int_{\tau} g_1(\tau)\delta(t-\tau)d\tau = g_1(t)$$

In this case, the response is the first-order kernel. Note that the first-order kernel shown in Figure 6a was estimated from multiple repetitions of the velocity impulse, and averaging across the trials. This is equivalent to the cross-intensity measure $m_{sy}$, previously defined in equation 5.

What is not clearly exhibited is the known post-excitatory suppression that was observed in the auto-intensity measures in Figures 5b and c. Temporal interactions between paired tactile stimuli have been previously utilized to infer underlying levels of post-excitatory suppression (Simons, 1985; Kyriazi et al., 1994; Fanselow and Nicolelis, 1999; Stanley and Webber, 2003). The resulting behavior of a typical cortical neuron is shown in Figure 6b. If the time interval $u$ between the first and second deflection is long (100ms), there is still a robust response to the second deflection (left panels). However, when the interval is decreased, the response to the second stimulus is attenuated (middle panels), until eventually disappearing altogether for very short latencies (right panels). The response following the second stimulus is different from what would be predicted by the superposition of the responses to the two stimuli, and thus reflects second-order dynamics:

$$r(t) \quad = \quad g_1(t) + g_2(t, t-u) \qquad t > 0$$

where the first-order effect associated with the stimulus at $t-u$ is negligible due to the transient nature of the response. Alternately, the relationship can be expressed in terms of the first-order conditioned kernel (Klein, 1992):

$$r(t) \quad = \quad g_1(t, u) \tag{9}$$

which depends on both the time $t$ and the time interval between the first and second stimulus $u$, where $g_1(t,u)$ becomes smaller as $u$ is decreased. This is a means by which the second-order kernel can be embedded in the first-order representation.

The strongly nonlinear dynamics are important in behavioral contexts, where the animal is using its vibrissae to palpate object surfaces, inducing spatial and temporal patterns of tactile input. The approach presented here can be extended to more complex patterns or sequences that may be encountered in the animal's natural environment.

# 6 Neural Control of Cardiac Function

The identification problems thus far have been discussed in terms of time domain estimation, and have arisen from phenomenology related to sensory pathways. An example is presented here in which frequency domain techniques are utilized to identify the dynamics of neural influence on the variability of cardiac rate. In healthy humans, the sino-atrial (SA) node acts as the pacemaker for the heart. Through an upward drift in electrical potential, these cells spontaneously depolarize to a threshold potential, at which point they rapidly depolarize, or "fire" as a group. This event is followed by a reset which marks the start of a new cycle. Firing initiates the spread of electrical activity throughout the heart, followed by contraction of cardiac muscle. The R-wave of an electrocardiogram (ECG), which is readily localized in time, provides a convenient marker from which periods between SA node firings can be extracted. The spontaneous depolarization of SA nodal cells has an intrinsic rate that is regulated by direct input from the sympathetic and parasympathetic branches of the peripheral, or autonomic nervous system (ANS). Neural impulses arriving from the sympathetic branch tend to increase the mean HR, while impulses from the parasympathetic branch have the opposite effect. By this means, the ANS regulates HR. Generally speaking, heart beats do not occur with exact regularity, but rather exhibit random variations around a mean rate. As a result, the beat-to-beat intervals measured from the ECG have a stochastic component, which may be termed "heart rate variability" (HRV) (Berger et al., 1986). Interestingly, normal individuals show much greater HRV than those whose ANS function is attenuated due to aging, disease states, or pharmacologic blockade (Akselrod et al., 1985; Akselrod et al., 1981; Appel et al., 1989; Malliani et al., 1991). It has been found that HRV, while random, exhibits a correlation structure in time, which can be associated with various periodicities of modulation of HR. For example, activity of the higher respiratory centers has been shown to modulate HR at the respiratory frequency via the parasympathetic branch of the ANS (Akselrod et al., 1981; Liao et al., 1995; Stanley et al., 1996; Stanley et al., 1997), which will be the focus of discussion here. This phenomenon is known as respiratory sinus arrhythmia (RSA). Continuous measures of lung volume serve as a non-invasive probe of central rhythm activity, and are well-correlated with heart rate variability in the related frequency bands.

### 6.1 Input Driven Threshold Model

In contrast to the non-parametric models previously described, we turn our attention to a mechanistic model that is a combination of parametric and non-parametric elements, as illustrated in Figure 7a. Note that the spectra are normalized to have a maximum of 1, and are therefore

Approximately here

Figure 7: Peripheral nervous system control of heart rate. (**a**) Input driven threshold model. (**b**) Interval spectrum (solid) and lung volume spectrum (dotted), both normalized for comparison. (**c,d**) Estimated transfer function magnitude and phase, respectively. Adapted with permission (©2000 IEEE) from (Stanley et al., 2000).

unitless, as are the transfer function magnitudes. In the model, the occurrence of R-waves is denoted by "spikes" at times $\tau_0, \tau_1, \tau_2, \ldots$, and the time interval between R-waves is $T_k = \tau_k - \tau_{k-1}$. These events arise from an integrate-and-fire mechanism that functionally represents the spatial and temporal summation of inputs at the SA node. When the output of the integrator reaches a threshold of 1, an event (heart beat) occurs, and the integrator is reset to zero. The constant $\beta$ represents the mean rate at which the SA node depolarizes. The additional input to the integrator consists of filtered versions of the respiratory signal *s* and a noise term *e*. For the time interval between two events, $\tau_{k-1}$ and $\tau_k$, the integration reaches the unity threshold:

$$1 = \int_{\tau_{k-1}}^{\tau_k} [v(t) + r(t) + \beta]dt = \int_{\tau_{k-1}}^{\tau_k} [v(t) + r(t)]dt + \beta T_k = \Phi_k + \beta T_k \tag{10}$$

where *r* is a linearly filtered version the zero-mean respiratory input *s*, and *v* is the output of a linear filter whose input is Gaussian white noise. $\Phi_k$ represents the sum of the noise activity and the respiratory influence, integrated over the interval from $\tau_{k-1}$ to $\tau_k$. The input *s* is restricted to be zero mean, and therefore $E\{\Phi_k\}$ will also be zero, producing a mean R-R interval of $\mu_T \approx 1/\beta$. $\beta$ in this case is representative of a combination of the intrinsic firing rate of the SA node and the mean levels of parasympathetic and sympathetic activity. It can be shown (Stanley et al., 2000) that the relationship between the spectrum of the respiratory input *s* and the sequence of R-R intervals can be approximated as:

$$S_{TT}(\omega) \quad \approx \quad \mu_T^3\{|H(\omega)|^2 + |G(\omega)|^2 S_{ss}(\omega)\} \tag{11}$$

Given the observed sequence of R-R intervals, the measured respiratory input *s*, and the previously identified noise model (see (Stanley et al., 2000) for a discussion of noise model estimation), the respiratory-related transfer function magnitude, $|G(\omega)|$ can be estimated. Using additional arguments, the phase angle $\angle G(\omega)$ can also be reconstructed (Stanley et al., 2000).

Figure 7b shows the lung volume and R-R interval spectra for a typical data set from a previously published study (Stanley et al., 1996), in which subject was breathing at randomly spaced intervals over a relatively broad frequency band. The corresponding transfer function magnitude and phase estimates over frequency bands of high coherence are shown in c and d. The transfer function magnitude tends to be rather flat over frequency bands of sufficient excitation. The phase suggests that there are increased delays between central rhythm activity and the influence upon heart rate variability at higher temporal frequencies.

This example provides a hybrid framework that is mechanistic, yet non-parametric, in contrast to the previous examples. The implementation in the frequency domain provides computational efficiency, but is also a relatively natural representation due to the frequency-band specific influence of the respiratory activity.

## 7 Summary

In summary, an overview of the perspectives and techniques related to system identification have been presented and utilized within the context of several well-defined problems in the nervous system. Importantly, several of the problems presented here involved combinations of non-parametric and parametric representations that reflect an important point to emphasize. In the analysis of complex systems, where little is known about the underlying dynamics, non-parametric techniques provide a relatively unconstrained means by which to characterize and categorize the properties of the system. However, as more is learned about the underlying dynamics, particular nonlinear features of the relationships can be exploited to greatly simplify the characterizations, often into linear-like descriptions, for which many computational tools exist. Finally, the motivation for the application of system identification within the nervous system has been posed here as that of basic science, or exploration. However, as in engineering, the problem of system identification in neuroscience eventually will be oriented towards control applications, in the context of engineered prosthetics designed to enhance neural function impaired due to trauma or disease.

## Acknowledgments

The author would like to thank Nicholas A. Lesica and Roxanna M. Webber for helpful comments in the preparation of this manuscript.

## References


Akselrod, S., Gordon, D., Madwed, J., Snidman, N., Shannon, D., and Cohen, R. (1985). Hemodynamic regulation: investigation by spectral analysis. *AJP*, 249:H867–H875.

Akselrod, S., Gordon, D., Ubel, F., Shannon, D., Bargar, A., and Cohen, R. (1981). Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. *Science*, 213(10):220–222.

Albrecht, D. G., Farrar, S. B., and Hamilton, D. B. (1984). Spatial contrast adaptation characteristics of neurones recorded in the cat's visual cortex. *J. Physiol.*, 347:713–739.

Appel, M., Berger, R., Saul, J., Smith, J., and Cohen, R. (1989). Beat to beat variability in cardiovascular variables: noise or music? *J. Am. Coll. Cardiol.*, 14:1139–1148.

Berger, R., Akselrod, S., Gordon, D., and Cohen, R. (1986). An efficient algorithm for spectral analysis of heart rate variability. *IEEE Trans. Biomed. Eng.*, BME-33:900–904.

Box, G. and Jenkins, G. (1976). *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco, California, revised edition.

Brillinger, D. (1981). *Time Series Analysis*. McGraw Hill, New York, expanded edition.

Brillinger, D. R. (1992). Nerve cell spike train data analysis: a progression of technique. *Journal of the American Statistical Society*, 87(418):260–271.

Brillinger, D. R., H. L. Bryant, J., and Segundo, J. P. (1976). Identification of synaptic interactions. *Biol. Cybernetics*, 22:213–228.

Brown, E. N., Nguyen, D. P., Frank, L. M., Wilson, M. A., and Solo, V. (2001). An analysis of neural receptive field plasticity by point process adaptive filtering. *PNAS*, 98:12261–12266.

Brumberg, J. C., Pinto, D. J., and Simons, D. J. (1996). Spatial gradients and inhibitory summation in the rat whisker barrel system. *J. Neurophysiol.*, 76:130–140.

Carvell, G. E. and Simons, D. J. (1990). Biometric analysis of vibrissal tactile discrimination in the rat. *J. Neurosci.*, 10:2638–2648.

Carvell, G. E. and Simons, D. J. (1996). Abnormal tactile experience early in life disrupts active touch. *J. Neurosci.*, 15:2750–2757.

Chatfield, C. (1989). *The Analysis of Time Series: An Introduction*. Chapman and Hall, London.

Chichilnisky, E. J. (2001). A simple white noise analysis of neuronal light responses. *Network*, 12:199–213.

Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton University Press, New Jersey.

Dayan, P. and Abbott, L. F. (2001). *Theoretical Neuroscience*. MIT Press, Cambridge.

DeBoer, E. and Kuyper, P. (1968). Triggered correlation. *IEEE Trans. Biomed. Eng.*, 15:169–179.

Eichenbaum, H. and Davis, J. (1998). *Neuronal Ensembles: Strategies for Recording and Decoding*. John Wiley and Sons, New York.

Enroth-Cugell, C. and Shapley, R. (1973). Adaptation and dynamics of cat retinal ganglion cells. *J. Physiol.*, 233:271 – 309.

Fanselow, E. E. and Nicolelis, M. A. L. (1999). Behavioral modulation of tactile response in the rat somatosensory system. *J. Neurosci.*, 19:7603–7616.

Greblicki, W. (1992). Nonparametric identification of wiener systems. *IEEE Trans. on Inform. Thry.*, 38:1487–1493.

Guic-Robles, E., Valdivieso, C., and Guajardo, G. (1989). Rats can learn a roughness discrimination using only their vibrissal system. *Behavioural Brain Research*, 31:285–289.

Hubel, D. (1995). *Eye, Brain, and Vision*. Scientific American Library, New York.

Hunter, I. W. and Korenberg, M. J. (1986). The identification of nonlinear biological systems: wiener and hammerstein cascade models. *Biological Cybernetics*, 55:135–144.

Jenkins, G. and Watts, D. (1968). *Spectral Analysis and Its Applications*. Holden-day.

Jones, J. P. and Palmer, L. A. (1987). The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *J. Neurophysiol.*, 58:1187–1211.

Klein, S. A. (1992). Optimizing the estimation of nonlinear kernels. In Pinter, R. B. and Nabet, B., editors, *Nonliner Vision*, pages 109–170. CRC Press.

Korenberg, M. J. and Hunter, I. W. (1999). Two methods for identifying wiener cascades having noninvertible static nonlinearities. *Annals of Biomed. Eng.*, 27:793–804.

Kyriazi, H. T., Carvell, G. E., and Simons, D. J. (1994). OFF response transformations in the whisker/barrel system. *J. Neurophysiol.*, 172(1):392–401.

Lee, Y. W. and Schetzen, M. (1965). Measurements of the wiener kernels of nonlinear systems by cross-correlation. *Int. J. Control*, 2:237–254.

Lesica, N. A., Boloori, A. S., and Stanley, G. B. (2003). Adaptive encoding in the visual pathway. *Network: Computation in Neural Systems*, 14:119–135.

Liao, D., Barnes, R., Chambless, L., Simpson, R., Sorlie, P., and Heiss, G. (1995). Age, race, and sex differences in autonomic cardiac function measured by spectral analysis of heart rate variability - the ARIC study. *Amer. J. of Cardiol.*, 76(12):906–912.

Ljung, L. (1987). *System Identification: Theory For the User*. Prentice-Hall, Inc., New Jersey.

Ljung, L. and Söderström, T. (1983). *Theory and Practice of Recursive Identification*. MIT Press, Cambridge.

Malliani, A., Pagani, M., Lombardi, F., and Cerutti, S. (1991). Cardiovascular neural regulation explored in the frequency domain. *Circulation*, 84:482–492.

Mao, B. Q., MacLeish, P. R., and Victor, J. D. (1998). The intrinsic dynamics of retinal bipolar cells isolated from tiger salamander. *Vis. Neurosci.*, 15:425–438.

Marmarelis, P. and Naka, K. I. (1973a). Non-linear analysis and synthesis of receptive field responses in the catfish retina. i. horizontal cell-ganglion chains. *J. Neurophysol.*, 36:605–618.

Marmarelis, P. and Naka, K. I. (1973b). Non-linear analysis and synthesis of receptive feld responses in the catfish retina. ii. one-input white-noise analysis. *J. Neurophysol.*, 36:619–633.

Marmarelis, P. Z. and Marmarelis, V. Z. (1978). *Analysis of Physiological Systems: The White-Noise Approach*. Plenum Press, New York.

Meister, M. and Berry, M. (1999). The neural code of the retina. *Neuron*, 22:435–450.

Movshon, J. A., Thompson, I. D., and Tolhurst, D. J. (1978). Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *J. Physiol.*, 283:53–77.

Ohzawa, I., Sclar, G., and Freeman, R. D. (1985). Contrast gain control in the cat's visual system. *Journal of Neurophysiology*, 54:651 – 667.

Papoulis, A. (1984). *Probability, Random Variables, and Stochastic Processes*. McGraw Hill, New York, second edition.

Paulin, M. G. (1993). A method for constructing data-based models of spiking neurons using a dynamic linear-static nonlinear cascade. *Biol. Cybern.*, 69:67–76.

Perkel, D., Gerstein, G., and Moore, G. (1967). Neuronal spike trains and stochastic point processes. i. the single spike train. *Biophysical Journal*, 7(4):391–418.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, Cambridge.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*. Wiley, New York.

Reid, R. C., Soodak, R. E., and Shapley, R. M. (1991). Directional selectivity and spatiotemporal structure of receptive fields of simple cells in cat striate cortex. *J. Neurophysiol.*, 66:505–529.

Ringach, D., Sapiro, G., and Shapley, R. (1997). A subspace reverse-correlation technique for the study of visual neurons. *Vision Res.*, 37:2455–2464.

Shapley, R. and Victor, J. D. (1978). The effect of contrast on the transfer properties of cat retinal ganglion cells. *J. Physiol*, 285:275–298.

Simons, D. J. (1985). Temporal and spatial integration in the rat SI vibrissa cortex. *J. Neurophysiol.*, 54:615–635.

Söderström, T. and Stoica, P. (1987). *System Identification*. Prentice Hall, New Jersey.

Stanley, G., Li, F., and Dan, Y. (1999). Reconstruction of natural scenes from ensemble responses in the lateral geniculate nucleus. *J. Neurosci.*, 19(18):8036–8042.

Stanley, G., Verotta, D., Craft, N., Siegel, R., and Schwartz, J. (1996). Age and autonomic effects on interrelationships between lung volume and heart rate. *American Journal of Physiology*, 270:H1833–H1840.

Stanley, G., Verotta, D., Craft, N., Siegel, R., and Schwartz, J. (1997). Age effects on inter-relationships between lung volume and heart rate during standing. *American Journal of Physiology*, 273:H2128–H2134.

Stanley, G. B. (2002). Adaptive spatiotemporal receptive field estimation in the visual pathway. *Neural Computation*, 14:2925–2946.

Stanley, G. B., Poolla, K., and Siegel, R. A. (2000). Threshold modeling of autonomic control of heart rate variability. *IEEE Transactions on Biomedical Engineering*, 49(9):1147–1153.

Stanley, G. B. and Webber, R. M. (2003). A point process analysis of sensory encoding. in press, *J. Comput. Neurosci.*

Tolhurst, D. J. and Dean, A. F. (1990). The effects of contrast on the linearity of the spatial summation of simple cells in the cat's striate cortex. *Experimental Brain Research*, 79:582–588.

Volterra, V. (1959). *Theory of Functionals and of Integral and Integro-Differential Equations*. Dover Publications, New York.