ABSTRACT

In the last decade-and-a-half, a number of unexpected statistical properties have been found in the pulsatile activity of neurons in many sensory systems. These properties have been variously called "fractal behavior", "self-similarity", "1/f fluctuations", and "long-duration correlation". In this dissertation, we develop a cohesive and robust theory of long-range dependence in point processes that subsumes these statistical properties and connects them to the much broader mathematical and scientific literature on long-range dependence in general stochastic processes. As this theory is developed, we examine ways in which the presence of long-range dependence can be used to discriminate between potential models of neural processing and analyze ways in which it can undermine standard statistical analyses of neural activity.

We first study integrate-and-fire models of cortical processing that were designed to explain the high variability of the activity in cortical neurons to see if they can also produce long-range dependence. Since a large number of these models produce renewal-process outputs, we prove analytically that such models cannot simultaneously possess long-range dependence and physiologically realistic variability. We then consider integrate-and-fire models with renewal point process inputs. By analyzing the outputs of simulations, we show that these models can possess long-range dependence and realistic variability if the interval distribution of their inputs has infinite variance. Since this latter requirement contradicts empirical results, we suggest a new integrate-and-fire model with long-range dependent inputs having finite interval-variance. Through the use of simulations, we show that this model is able to produce both long-range dependence and realistic variability.

We also explore the impact that long-range dependence has on empirical estimates of the mean, standard deviation, and variance of spike counts in neural activity. We derive mathematical formulae for the actual mean and variance and for the statistical behavior of sample mean and sample variance estimators for a model neuron. Using these formulae and simulations of the model, we show that long-range dependence significantly increases the variability of the sample mean and causes the sample standard deviation and sample variance to be negatively biased. The significance of these results for existing empirical measurements is demonstrated.

CONSEQUENCES OF
LONG-RANGE TEMPORAL DEPENDENCE
IN NEURAL SPIKING ACTIVITY
FOR THEORIES OF PROCESSING AND CODING

By

Brian Scott Jackson
B.S. University of Toledo, 1992
M.S. Syracuse University, 1996

DISSERTATION

Submitted in partial fulfillment of the requirements for the
degree of Doctoral of Philosophy in Neuroscience
in the Graduate School of Syracuse University

May 2003

Approved_____
Dr. Laurel H. Carney
Professor of Bioengineering and Neuroscience

Date_____

Committee Approval Page

(To be inserted by Graduate School)

Contents

List of Figures

List of Tables

List of Acronyms

CV_{ISI}	Coefficient of Variation of InterSpike Intervals
DSPP	Doubly Stochastic Poisson Process
EPSP	Excitatory Post-Synaptic Potential
FFC	Fano Factor Curve
FFT	Fast Fourier Transform
fBm	fractional Brownian motion
fGn	fractional Gaussian noise
fGnDP	fractional Gaussian noise Driven Poisson process
IDC	Index of Dispersion (of interspike intervals) Curve
IF	Integrate-and-Fire
IPSP	Inhibitory Post-Synaptic Potential
LRD	Long-Range Dependence (or Dependent)
LRcD	Long-Range Count Dependence
LRiD	Long-Range Interval Dependence
PSP	Post-Synaptic Potential
RM-fGnDP	Refractory-Modified fractional Gaussian noise Driven Poisson Process
RPP	Renewal Point Process
SR	Spontaneous spike Rate

ACKNOWLEDGEMENTS

Through the course of my graduate studies, I have been financially supported by a number of different sources. I would like to thank the Syracuse Graduate School for several years of support through fellowships and the Departments of Bioengineering and Mathematics for offering me teaching assistantships. I am also grateful for the grants from NIH and NSF that allowed me to focus my time and energy on research during summers and several full years. I am indebted to Drs. Robert Barlow and Fred Dodge for accepting this auditory researcher into their "visual fold" for a summer when other research funding had run dry; it was a great experience and a pleasure. Finally, I would like to thank the Gerbers for their endowment to Syracuse University for auditory research, which supported me during the work contained in this dissertation. The support of this endowment was an enormous blessing as I made my final push to the finish line.

I cannot possibly express my gratitude to Dr. Evan Relkin for his many, many years of mentorship and support, even when ill-health limited my abilities for significant portions of time. I am deeply sorry that he was taken from this earth before seeing the fruition of his labors. I am also enormously thankful to Dr. Laurel Carney who so generously "adopted" me and served quite admirably as my self-proclaimed "cheerleader". Although I might not have always responded in an appreciative manner, I am indebted to you for all of the patient prodding and goading. I am also very grateful to the rest of my oral examination committee: Drs. Robert Smith, Sandy Bolanowski, Hyune-Ju Kim, Phillip Griffin, and Brad Motter. Thanks for your perseverance in the face of difficult and unfamiliar material, whether mathematics for the neuroscientists or neuroscience for the mathematicians.

For challenging me, giving me a strong foundation in mathematics and science, and nurturing in me a love of those subjects, I thank Ed Wajda, Larry Sherer, and Jim Piroga at Lakeview High and Dr. Henry C. Wente at the University of Toledo. For encouraging me to pursue a masters degree in mathematics, which greatly facilitated my ability to conceive of and carry out the work of this dissertation, I thank Dr. John L. Troutman. Many thanks are due to the Mathematics Department faculty, staff, and students for embracing a wannabe mathematician, and to the faculty, staff, and students at the Institute for Sensory Research for their help, advise, and camaraderie. Thanks for making ISR such a great place to "live" and work. In particular, I thank Adam Pack for the many computer programming discussions and for making graduate school a much more enjoyable experience. I especially thank Anita Sterns, who skillfully performed the surgical preparations for all of my experiments and helped out during many of them. Although those experiments did not make their way into this document, they certainly undergirded my motivation to pursue this line of research.

I am deeply grateful to my many brothers and sisters in the Graduate Christian Fellowship at Syracuse University and at Trinity Fellowship for helping me to grow spiritually and for helping (or forcing) me to take a break from work and enjoy the more important and fun things in life. So many of you mean more to me than you can imagine; unfortunately I cannot thank each of you separately in this space. I would, however, like

Chapter 1

Introduction

1.1 Background

1.1.1 Neurophysiology and Simple Point Processes

Most of the information processed in the nervous system is conveyed from one neuron to another in a sequence of electrical impulses known as action potentials. A short voltage recording containing three of these impulses is shown in the upper trace of Figure 1.1. These action potentials, or "spikes", are all-or-none events. In other words, there is no deterministic gradation in their shape to transport information. Thus, presumably, the only critical parameter of a spike is the time of its occurrence, and a neurophysiological recording can be reduced to the sequence of times at which spikes occurred. Mathematically, such a process can be represented with a point process, which, in this context, is essentially the designation of a sequence of "points" on the real line. The point process associated with the neurophysiological recording in Figure 1.1 is shown just below the voltage trace, and the times of occurrence of the three spikes are designated by the real numbers t_1, t_2, and t_3.

Whether because of the presence of noisy elements in the nervous system or deficiencies in our knowledge of the relevant conditions and mechanisms (or both!), deterministic descriptions of the precise timing of spikes in particular *in vivo* neurons are scarce. Hence, sequences of spikes, or "spike trains", appear to possess a significant random component. Therefore, they are frequently modeled by stochastic point processes.

The simplest stochastic point process, and probably the most common model of neural spike trains, is the Poisson process. Much of the early modeling work on the stochastic nature of neural spike trains, in many different parts of the nervous system, made use of this model (e.g. Mueller, 1954; Kuffler, FitzHugh, & Barlow, 1957; Bishop, Levick, & Williams, 1964; Siebert, 1965; Smith & Smith, 1965). A Poisson process is solely parameterized by the rate of occurrence of its points, and the probability of a point occurring at any particular time is completely independent of both its past and its future. Neurophysiological recordings, however, are conspicuously devoid of pairs of spikes that occur in close proximity, which is not true of the Poisson process. This absence of very short interspike intervals is termed absolute refractoriness. In addition, somewhat longer

Neurophysiological Recording
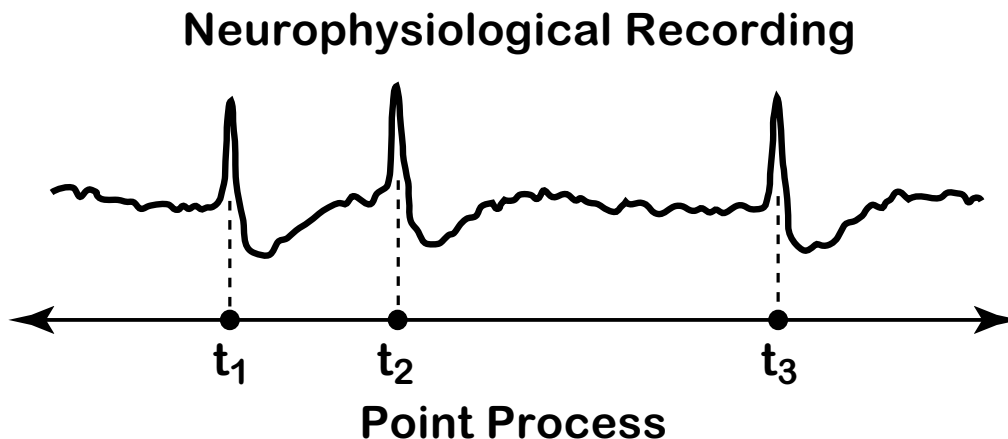
Point Process

Figure 1.1: Reduction of a neurophysiological recording to a point process. The upper trace is a segment of the voltage waveform from an extracellular recording of a neuron. Each of the nearly identical spikes in the voltage is called an "action potential". In the point process depicted below the trace, only the occurrence times (t_1, t_2, and t_3) of the action potentials are preserved.

interspike intervals are more rare than is accounted for by the Poisson process, a condition called relative refractoriness. Refractoriness, absolute and/or relative, can be accounted for in a point process by certain renewal point processes. Unlike the Poisson process, a renewal point process may depend on its past, though only on the period extending back to the last point that occurred. Hence, in a renewal point process, the lengths of the interspike intervals are still independent of each other, as they were in the Poisson process. Renewal processes that incorporate the refractory period(s) of neural spike trains may be described in terms of the application of a thinning procedure, which describes a deterministic, probabilistic, or mixed rule for removing points, to the Poisson process. In the context of neural modeling, such point processes are call refractory-modified Poisson processes.

1.1.2 Long-Range Dependence in Neural Spike Trains

The Poisson process has no temporal autocorrelation, and the autocorrelation of standard renewal point processes has a very limited extent. In contrast, like many other natural processes (see Beran, 1994, for examples), the spike trains of many (or, perhaps, most) neurons have autocorrelations that endure for an exceptionally long period of time. This long-range dependence is evident through the power-law behavior of such statistics as the Fano factor curve and the power spectrum and has been observed in neurons found in the auditory nerve (Teich, 1989; Teich & Lowen, 1994; Lowen & Teich, 1996b), the lateral superior olive (Teich, Johnson, Kumar, & Turcott, 1990; Turcott et al., 1994), the retina and lateral geniculate nucleus (Teich, 1996; Teich, Heneghan, Lowen, Ozaki, & Kaplan, 1997; Lowen, Ozaki, Kaplan, Saleh, & Teich, 2001), the visual system of certain insects (Turcott, Barker, & Teich, 1995), the striate cortex (Teich, Turcott, & Siegel, 1996), the somatosensory cortex (Wise, 1981), and the mesencephalic reticular formation (Gruneis,

Nakao, Yamamoto, Musha, & Nakahama, 1989; Gruneis, Nakao, & Yamamoto, 1990; Gruneis et al., 1993). Since power-laws are self-similar, a vaguer term incorporating the word "fractal" is often used to denote long-range dependence in these neurophysiological studies.

Since the standard point process models of neural spike trains do not possess long-range dependence, new long-range dependent point process models have been developed since the discovery of long-range dependence in neural spike trains. Teich and Lowen (Teich, Turcott, & Lowen, 1990; Lowen & Teich, 1991, 1995, 1996a, 1997) have developed several similar models that incorporate long-range dependent (or "fractal") noises, such as fractional Gaussian noise. They are all based on a similar idea. The rate of occurrence of points in a Poisson process, or in the Poisson process of a refractory-modified Poisson process, does not change with time. Teich and Lowen's models, however, set the instantaneous rate of these processes equal to a long-range dependent noise, thus producing long-range dependence in the output as well. These "fractal-noise-driven Poisson process" models, however, are difficult to handle analytically, especially within the context of a larger model. Furthermore, due to their long-range memory, they are computationally cumbersome to simulate. Hence, although much evidence of long-range dependence in neural spike trains has been produced and models of this property are available, little progress has been made in understanding the significance of long-range dependence to neurophysiology in general.

1.1.3   The Integration-High Variability Paradox

Another approach, different from those mentioned in Section 1.1.1, to modeling neural spike trains employs a threshold or level-crossing mechanism that is applied to a given continuous time series. In this case, the points of the process occur at the instants when the time series crosses the specified threshold or level. In order to represent the randomness apparent in spike trains, the time series may be partially or fully stochastic.

The integrate-and-fire (IF) neuron, which was first investigated by Lapicque (1907, 1926), is a well-established mathematical model of neural spiking that incorporates a threshold mechanism. In addition, it includes an abstraction, the "integration" mechanism, of the way in which a neuron processes and combines its inputs. Thus, the IF neuron has proved particularly useful in modeling neurons, like those in the cortex, that receive large numbers of inputs. In the most basic form of this model, each spike that occurs on an input "neuron" causes the value of the IF neuron's "membrane voltage" to instantaneously increase or decrease by a fixed amount, depending on whether the input is excitatory or inhibitory. When the membrane voltage reaches the threshold value, a spike (or point) occurs at its output. The membrane voltage is then reset to its initial value and the process begins anew. In this form, the IF neuron is a purely deterministic model. However, if the inputs are stochastic then the output of the neuron will be as well.

The appropriateness of the IF neuron model for modeling cortical neurons was called into question by Softky and Koch (1992, 1993). They drew attention to the disparity between the variability of neural spike trains from *in vivo* cortical neurons and the variability that is expected based on IF models of these neurons. Since cortical neurons receive a large number of inputs, these IF models receive spikes from a large number of

Poisson processes. However, in order to produce reasonable spike rates at the output of the IF neuron, the threshold must be large relative to the effect of each individual input. Under these conditions, the integration mechanism of the IF neuron essentially averages the inputs and produces a fairly regular, or nearly periodic, spike train at its output. But, *in vivo* the variability of the time between spikes in cortical neurons is quite large (Noda & Adey, 1970; Burns & Webb, 1976; Dean, 1981; Softky & Koch, 1992, 1993). Two questions raised by these paradoxical observations are: (1) can the IF neuron model reproduce the high variability of *in vivo* neurons in a physiologically realistic way? and (2) how can a neuron that integrates over many inputs produce such high variability? One set of possible answers to these questions are that (1) the IF neuron cannot realistically reproduce *in vivo* -like high variability, and (2) cortical neurons do not integrate over many inputs, but instead act as coincidence detectors. As a coincidence detector, a cortical neuron would respond to input spikes that are highly synchronized in time. Since an output spike could be initiated by only a few synchronized input spikes, the neuron could have many inputs while still preserving a large portion of the variability of each input.

However, a considerable literature of attempts to replicate the variability of *in vivo* neurons with the IF model, in a realistic and meaningful way, developed in response to Softky and Koch (1992, 1993). The high variability in these models is either created by increasing the variability of the input to the IF neuron or by increasing the sensitivity, or gain, of the spiking mechanism to changes in the input. In either case, although the resultant model is integrating a large number of inputs, it is predominately operating near threshold, where the inputs are just sufficient to cause output spikes. This state of the model is what creates the high output variability (Bell, Mainen, Tsodyks, & Sejnowski, 1995; Burkitt & Clark, 2000). Three methods that have been used to increase the variability of the input are (1) balancing the amount of excitation and inhibition (Shadlen & Newsome, 1994; Tsodyks & Sejnowski, 1995; Vreeswijk & Sompolinsky, 1996, 1998; Feng & Brown, 1998a, 1999; Brown & Feng, 1999; Feng, 1999; Christodoulou, Clarkson, Bugmann, & Taylor, 1994; Christodoulou & Bugmann, 2000; Salinas & Sejnowski, 2000; Burkitt, 2000, 2001), (2) using long-tailed interspike interval distributions (Feng, 1997; Feng & Brown, 1998b; Feng, Tirozzi, & Brown, 1998), and (3) using correlated inputs (Stevens & Zador, 1998; Sakai, Funahashi, & Shinomoto, 1999; Shinomoto & Tsubo, 2001; Feng & Brown, 2000a; Feng, 2001; Feng & Zhang, 2001; Destexhe & Pare, 1999, 2000). Methods used to increase the sensitivity of the spiking mechanism include (1) using a post-spike reset potential just below threshold (Lansky & Smith, 1989; Bell et al., 1995; Troyer & Miller, 1997; Bugmann, Christodoulou, & Taylor, 1997), or an equivalent time-varying threshold (Wilbur & Rinzel, 1983), and (2) making the leakage coefficient dependent on the membrane voltage (Feng & Brown, 2000b).

1.1.4 General Statistical Examinations of Neurophysiological Data

If the nervous system transmits and processes information on the basis of a rate code—a common assumption—then the most important aspect of deciphering neural information is distinguishing between different spike counts in fixed-length counting windows. However,

the variability of these spike counts in response to a fixed stimulus will limit the ability of the nervous system to distinguish between different spike counts.

When studying or modeling the activity in neurons, it is frequently necessary to statistically characterize neural spike trains. In particular, especially in the context of rate coding, the mean and/or variance of the number of spikes in a fixed length time interval are often of interest, or are necessary as components of the desired analysis. These values, then, must be estimated from experimental data using the sample mean and/or sample variance. The sample mean and sample variance are so commonly calculated and used, however, that one rarely considers the assumptions which undergird their meaningfulness, as well as that of analyses that are based upon them.

One of these basic assumptions is that the observations, or samples, from which the statistics are calculated are independent. Rarely, in practice, is this requirement met. Usually, however, deviations from these ideals are small, in the sense that they have limited effects on the results. For instance, sample counts from a Poisson process would not violate this assumption, whereas those from a (non-Poissonian) renewal process would. Yet, due to the limited range of the temporal autocorrelation present in a typical renewal process, the detrimental effects on the behavior of the sample statistics would be negligible under most conditions. However, since the autocorrelation present in actual spike trains has a long extent, its effect on the behavior of sample statistics cannot be so easily dismissed. For stochastic time series, negative effects of long-range dependence on the reliability of sample statistics are well-documented (see, for example, Beran, 1994). Therefore, similar effects are likely to occur in statistical analyses of spike trains as well.

1.2   Objectives

The main goals of this thesis are:

(i) to develop a more robust understanding of the properties of long-range dependent point processes, especially with respect to those that are applicable as models of neural spike trains;

(ii) to investigate whether the integrate-and-fire models of cortical processing discussed in Section 1.1.3 are compatible with the presence of long-range dependence;

(iii) and to assess the extent to which statistical estimates of the properties of spike trains are likely to be compromised by long-range dependence in the spike trains.

In Chapter 2, we develop a basic theory of long-range dependence in point processes. We then proceed, in light of this theory, to analyze, as a whole, a large class of "high-variability cortical models" to determine if and to what extent they can simultaneously produce physiologically realistic variability and long-range dependence. Although mathematical analysis of this class of models relies solely on global properties of these models, we also exam a particularly common example in further detail in order to facilitate an intuitive understanding of the interplay between variability of the interspike intervals and long-range dependence.

In Chapter 3, the basic theory developed in Chapter 2 is augmented to better differentiate types of long-range dependence that can occur in point processes. In addition, methods are described that are sensitive to these different forms of long-range dependence in empirical data. Using partial analytical results and analysis of data obtained from simulations, we investigate the ability of two types of high-variability cortical models, which are not subject to the analysis of Chapter 2, to produce high-variability and long-range dependence. Each of these models differs from the basic integrate-and-fire model in that its inputs are not Poisson processes. The first model, which was proposed by Feng and his coworkers (Feng, 1997; Feng & Brown, 1998b; Feng et al., 1998), uses renewal point processes as inputs. The second model, which we propose, uses fractional-Gaussian-noise-driven Poisson processes, which have properties similar to those of many inputs to the cortex, as inputs.

In Chapter 4 we address the subject of the effect that long-range dependence in spike trains has on statistical estimators of spike counts. We do this by studying the bias and variability of the sample mean, sample variance, and sample standard deviation of counts for the fractional-Gaussian-noise-driven Poisson process model. Analytical approximations for most of these properties are derived, and simulations are used to evaluate the remainder. Finally, the significance of these results for published estimates of spike count statistics are discussed.

Chapter 2

Long-Range Dependence and Models of the High Interspike-Interval Variability of Cortical Neurons I: Models Producing Renewal Outputs

ABSTRACT Many different types of integrate-and-fire models have been designed in order to explain how it is possible for a cortical neuron to integrate over many independent inputs while still producing highly variable spike trains. Within this context the variability of spike trains has been almost exclusively measured using the coefficient of variation of interspike intervals. However, another important statistical property that has been found in cortical spike trains and that is closely associated with their high firing variability is long-range dependence. We show that a large class of high-variability integrate-and-fire models are incapable of producing outputs that possess both coefficients of variation within the range that has been empirically measured from cortical spike trains *and* long-range dependence. The argument consists of two parts: first, we prove that no renewal point process can possess both of these properties simultaneously; then, we argue that the outputs of a large number of high-variability integrate-and-fire models are renewal point processes. This class of models includes those that use excitation-inhibition balance, correlated inputs, partial reset, or nonlinear leakage to produce outputs with high variability.

2.1   Introduction

Within computational neuroscience there is a debate over whether cortical neurons act as temporal integrators or coincidence detectors (see, for example, Konig, Engel, & Singer, 1996). Functionally, the contention is whether neurons combine synaptic inputs over time periods on the order of, or longer than, interspike intervals (temporal integration) or over time periods that are much shorter than interspike intervals (coincidence detection). Although the idea of cortical neurons as temporal integrators seems to be more dominant, Softky and Koch (1992, 1993) revealed a simple discrepancy between this theory and

the statistical nature of trains of action potentials (spikes) from real cortical neurons. The statistical discrepancy was found in the variability of neural responses as measured by the coefficient of variation of interspike intervals (CV_{ISI}). The CV_{ISI} is the standard deviation of the intervals divided by the mean of the intervals and, therefore, has a value of zero when all interspike intervals are deterministically equal and has a value of one for a Poisson process.

Softky and Koch (1992, 1993) showed, using both an integrate-and-fire (IF) neuron model and a more detailed compartmental neuron model, that if a neuron integrates a large number of excitatory inputs with small amplitude postsynaptic potentials, then its output should be very regular, with $CV_{ISI} \ll 1$. However, their analysis of recordings from primary visual cortex (V1) and extrastriate cortex (MT) of awake, behaving monkeys revealed that cortical neurons are actually quite irregular, with $CV_{ISI} \approx 1$. Thus, they concluded that cortical neurons do not integrate over many inputs, but instead function as coincidence detectors.

Shadlen and Newsome (1994), however, showed that neural models that integrate over many inputs are able to produce very irregular spike trains if they have both excitatory and inhibitory inputs. In their model highly variable output spike trains were obtained by making the total amount of inhibition equivalent to the total amount of excitation. This idea has been extended (Brown & Feng, 1999; Feng & Brown, 1998a; Feng, 1999; Shadlen & Newsome, 1998; Burkitt, 2000, 2001) to show that the amounts of inhibition and excitation do not need to be perfectly balanced, a condition that seems physiologically unrealistic. In fact, if they are too closely matched the CV_{ISI} of the model becomes much larger than the values measured from real neurons (e.g. Feng & Brown, 1998a, or Section 2.4 below).

Following these two foundational papers, a number of studies have been devoted to finding methods of producing highly variable output from an IF neuron model that requires a large number of input spikes to produce an output spike. In addition to the balancing of excitation and inhibition, other methods that produce realistic CV_{ISI} values include using long-tailed interspike-interval distributions for the input processes (Feng, 1997; Feng & Brown, 1998b, 1998a; Feng et al., 1998), using correlated inputs (Stevens & Zador, 1998; Sakai et al., 1999; Shinomoto & Tsubo, 2001; Feng & Brown, 2000a; Feng, 2001; Feng & Zhang, 2001; Destexhe & Pare, 1999, 2000), and using a post-spike reset potential near threshold (Troyer & Miller, 1997; Bugmann et al., 1997; Lansky & Smith, 1989). Furthermore, several studies have created networks of IF neurons where the network dynamics create highly variable outputs at the level of single neurons (Usher, Stemmler, Koch, & Olami, 1994; Usher, Stemmler, & Olami, 1995; Tsodyks & Sejnowski, 1995; Vreeswijk & Sompolinsky, 1996, 1998).

In addition to highly variable interspike intervals, there is evidence that the spike trains of cortical neurons also exhibit long-range temporal dependence (Teich et al., 1996; Gruneis et al., 1989, 1990, 1993; Wise, 1981).[1] Long-range dependence (LRD), or long memory, denotes that the dependence between random variables in a stochastic process decay "slowly" as the distance between them increases. For processes with finite variance,

---

[1]These papers do not use the term "long-range dependence", but in Section 2.3 we will relate their findings to long-range dependence.

this property is sometimes referred to as long-term correlation, or, since the form of these correlations, and the related power spectral density, often are, or approximate, certain power-law functions, they are sometimes referred to as second-order self-similar, fractal, or having power-law statistics. Whatever they are called, LRD processes with finite variances have correlations that decay so slowly that their sum diverges as the correlations for increasingly distant random variables are included. Therefore, any specification of a finite history of the process will necessarily have ignored a significant portion of these correlations. This leads to behavior in the variance of such a process that is unusual with regard to standard statistical procedures, and, as will become clear in the course of this chapter, is closely related to the variability as measured by the CV_{ISI}. However, its importance to the study of highly variable integrator models of cortical processing has rarely been acknowledged (but see Usher et al., 1994, 1995), and, then only with regard to networks of interconnected IF neurons.

In this chapter, we study the ability of a large class of the simplest and most general high-variability IF neuron models to produce LRD. The observation that the outputs of many of the models mentioned above are renewal processes allows us to deal with this entire subclass of models with a single analytical argument. We show that such models cannot both have a CV_{ISI} value approximating the values that have been measured empirically from cortical spike trains and be LRD within the same parameter regime.

We begin in Section 2.2 with the theory of LRD as it applies to point processes in general and give a specific result for renewal point processes. This section contains an introduction to LRD which motivates the subsequent definitions of LRD for stochastic point processes. This is followed by a statement and proof of a theorem that identifies when a renewal point process is LRD. In Section 2.3, we discuss the evidence for LRD in neural spike trains, including those produced by cortical neurons. This requires relating empirical analyses as presented in the literature with the definition of a particular type of LRD given in Section 2.2. In Section 2.4, we illustrate the main result by considering an example from the class of high-variability cortical models that produce renewal outputs. The model is a simple, but common, high-variability cortical model consisting of an arbitrary combination of mutually-independent, Poisson excitatory and inhibitory inputs converging on a single IF neuron. This model requires an approximate balance between excitation and inhibition in order to produce output spike trains with variability matching CV_{ISI} measurements from *in vivo* cortical neurons. Due to its simplicity, we are able to show analytically that this model, within any single parameter regime, cannot produce spike trains that simultaneously possess both the CV_{ISI} and LRD properties that have been measured in real cortical neurons. Next, in Section 2.5, we generalize this result to a large class of IF models. We first describe conditions on IF models that are sufficient to ensure that they produce renewal spike trains and argue that the failure of the cortical neuron model illustrated in Section 2.4 is generally applicable to this entire class of models.

2.2 Long-Range Dependence: Definitions and Theory

2.2.1 Long-Range Dependence in Stochastic Processes

A stationary stochastic process $\{X_i : i \in \mathbb{Z}\}$ is said to have short-range dependence if it possesses the following essentially equivalent properties (Cox, 1984), where C_k, for any integer k, denotes an arbitrary finite constant:

(i) $\lim\limits_{n \to \infty} \sum\limits_{j=0}^{n} Cov\{X_0, X_j\} = C_1$;

(ii) $S_X(0) = C_2$, where $S_X(\omega)$ is the spectral density (or power spectrum) of $\{X_i\}$;

(iii) $\dfrac{Var\left\{\sum_{j=1}^{n} X_j\right\}}{n} \sim C_3$ as $n \to \infty$.

Note that according to these properties, a short-range dependent process does not necessarily have any dependence at all. For instance, a process where X_i is independent of X_j for all $i \neq j$ is consistent with these three properties. In such a case, we might say that zero dependence is just an extreme example of short-range dependence.

Property (i) above implies that for increasing lags the covariance decays quickly enough that it is (infinitely) summable. On a practical level, this tells us that an arbitrarily large, though perhaps less than total, amount of the influence of the past can be contained in a sufficiently large finite history of the process. Since the spectral density function is just the Fourier transform of the autocovariance function, property (ii) is the frequency domain equivalent of property (i). Finally, property (iii) expresses the fact that the variance of the sample means decays to zero at least as quickly as $1/n$.

Most classical statistical estimators and tests assume that a process conforms to properties (i)-(iii) above, and the "classical" types of stochastic processes fall into this category. However, observations from many "physical" processes do not meet the requirements of these three properties. Such processes are said to exhibit long-range dependence, and this long-range dependence can result in grossly inaccurate results if classical statistical methods are applied to processes that possess it (Beran, 1994; Cox, 1984). In contrast to properties (i)-(iii) above, a long-range dependent stationary process $\{X_i\}$ possesses the following properties (Cox, 1984):

(i') $\lim\limits_{n \to \infty} \sum\limits_{j=0}^{n} Cov\{X_0, X_j\} = \infty$;

(ii') $\lim\limits_{\omega \to 0} S_X(\omega) = \infty$;

(iii') $\lim\limits_{n \to \infty} \dfrac{Var\left\{\sum_{j=1}^{n} X_j\right\}}{n} = \infty$.

Thus, by property (i'), the covariance of a long-range dependent process decays so slowly that it is not (infinitely) summable. Therefore, any finite history, no matter how large, of such a process will necessarily leave out an infinitely large amount of the influence from the past. Equivalently, according to property (ii'), such a process has a singularity in

its spectral density at zero frequency. In addition, the variance of the sample means of a long-range dependent process decays to zero more slowly than $1/n$.

Property (i'), the divergence of the infinite sum of the covariances, or, alternately, the correlations, is perhaps most helpful in forming an intuitive understanding of long-range dependence. It is important to note that this is an asymptotic property (as are the other two properties in the list) and that this asymptotic property is related to the asymptotic form of the covariance, or correlation, function. That it is an asymptotic property means that each individual correlation, by itself, is not critical. In other words, the correlation at any fixed finite lag can be arbitrarily large or arbitrarily small for either a short-range or long-range dependent process. The asymptotic form of the correlations is the crucial element since long-range dependent processes can, and usually do, have correlations that approach zero for arbitrarily large lags. Thus, long-range dependence results from the joint effect of all correlations.

If we are going to define long-range dependence for point processes, then we need to have a definition of long-range dependence for arbitrary stochastic processes. Any of the three properties above would suffice as a definition (see, for example, Beran, 1994), but, following Daley and Vesilo (1997) and due to its ease of use analytically and empirically, we will use property (iii') to define long-range dependence.

Definition 2.1. A stationary stochastic process $\{X_i : i \in \mathbb{Z}\}$, for which X_i has finite variance for all i, exhibits long-range dependence (LRD) when

$$\limsup_{n \to \infty} \frac{Var\left\{\sum_{i=1}^n X_i\right\}}{n} = \infty.$$

2.2.2   Long-Range Dependence in Stochastic Point Processes

A point process on the real line can be described by either its interpoint distances or by the numbers of points in any arbitrary set of intervals on the real line. If the points of a point process are given by the increasing sequence $\{\tau_i : i \in \mathbb{Z}\}$, then the sequence of intervals is given by $\{Y_i = \tau_i - \tau_{i-1} : i \in \mathbb{Z}\}$. On the other hand, the number of points in an arbitrary interval $(a, b]$, $a < b$, is given by $N(a, b] = N\big((a, b]\big) = \#\{i : \tau_i \in (a, b]\}$. Then, for example, the counts in a sequence of adjacent intervals of length $T > 0$ would be given by $\left\{N\big(a + (i-1)T, a + iT\big] : i \in \mathbb{Z}\right\}$, where a is any real number.

Since either the sequence of interpoint intervals or the sequence of counts in adjacent intervals (or both) can be long-range dependent, there are two different ways in which a point process can exhibit long-range dependence. Again following Daley and Vesilo (1997), we will call these two types of long-range dependence "long-range *interval* dependence" and "long-range *count* dependence", respectively. Thus, we have the following two definitions:

Definition 2.2. *(Daley & Vesilo, 1997)*[2] A stationary point process $N(\cdot)$ on the real line exhibits long-range interval dependence (LRiD) when the stationary sequence of interpoint

---

[2]The original definition of Daley and Vesilo (1997) does not contain the phrase "with finite variances", though the finiteness of the variances of the intervals is necessary and may have been implied.

intervals $\{Y_i\}$, with finite variances, is LRD in the sense that

$$\limsup_{n\to\infty} \frac{Var\left\{\sum_{i=1}^{n} Y_i\right\}}{n} = \infty.$$

Definition 2.3. *(Daley & Vesilo, 1997)* A second-order stationary point process $N(\cdot)$ on the real line exhibits long-range count dependence (LRcD) when

$$\limsup_{t\to\infty} \frac{Var\left\{N(0,t]\right\}}{t} = \infty.$$

The definition of LRiD is a direct analog of the general definition for LRD. The definition for LRcD, however, takes a slightly different form since we wish to obviate the use of the parameter T described above. This parameter is superfluous with regard to the definition of LRcD since

$$\sum_{i=1}^{n} N\left((i-1)T, iT\right] = N(0, nT]$$

and nT goes to infinity as n goes to infinity.

2.2.3 Long-Range Dependence in Stochastic Renewal Point Processes

In this chapter, we are considering models that produce renewal point processes, point processes in which all interpoint intervals are mutually independent. Clearly, a renewal point process cannot be LRiD, since by definition there is no dependence between the intervals,[3] but it can be LRcD (e.g. Daley & Vesilo, 1997; Lowen & Teich, 1993c, 1993d) . However, when is a renewal point process LRcD? As it turns out, this question has a straightforward answer: a renewal point process is LRcD when the variance of the interpoint intervals is infinite. This was stated by Daley (1999), but only a terse argument was included there. Due to the abbreviated nature of those comments,[4] we will now state and prove (following Daley's argument) this result.

Theorem 2.1. *A stationary renewal point process with distribution function F of the generic interpoint-interval random variable X, which has $F(0) = 0$ and finite mean $\mu = E\{X\}$, is LRcD if and only if $E\left\{X^2\right\} = \infty$.*

Proof. Since the interpoint interval random variable X has a finite mean, the moments $E\{N(0,1]\} = 1/\mu$ and $E\{N^2(0,1]\}$ are finite, where $N(\cdot)$ is the counting measure of the point process. Therefore, by Theorem 3.5.III of Daley and Vere-Jones (1988),

$$E\{N^2(0,t]\} = \int_0^t \frac{2U(s)-1}{\mu}\, ds,$$

---

[3]This is true even if the variance of the interval distribution is infinite, even though Definition 2.2 is not applicable to that case. If an extension of the concept of LRiD to the infinite interval variance case is to be meaningful, a point process with independent intervals cannot also be LRiD.

[4]Another reason that we have included the proof of this theorem is that there is a mismatch between the text and the equation in the argument of Daley (1999). The text specifies when equation (1.2) in that paper diverges, but equation (1.2) is an equation for $Var\{N(0,t]\}$, not $Var\{N(0,t]\}/t$. The former, which diverges for most interesting stochastic point processes, is only a necessary condition for long-range dependence, whereas the latter is the definition of long-range dependence.

where

$$U(t) = \sum_{j=0}^{\infty} \left(\lim_{h \downarrow 0} \Pr\{N(0,t] \geq j | N(-h,0] > 0\} \right)$$

is called the expectation function. For a general stationary ergodic point processes with finite second moments, $U(t)$ is the analogue of the renewal function (and, for a renewal point process, *is* the renewal function). The variance function can now be written as

$$Var\{N(0,t]\} = E\{N^2(0,t]\} - \left(E\{N(0,t]\} \right)^2$$
$$= \int_0^t \frac{2U(s) - 1}{\mu} \, ds - \left(\frac{t}{\mu} \right)^2$$
$$= \int_0^t \left[\frac{2U(s) - 1}{\mu} - \frac{2s}{\mu^2} \right] ds$$
$$= \frac{2}{\mu} \int_0^t \left[U(s) - \frac{s}{\mu} \right] ds - \frac{t}{\mu}$$

So,

$$\frac{Var\{N(0,t]\}}{t} = \frac{2}{\mu} \left(\frac{1}{t} \int_0^t \left[U(s) - \frac{s}{\mu} \right] ds \right) - \frac{1}{\mu},$$

which goes to infinity as $t \to \infty$ if and only if the integrand goes to infinity. In other words,

$$\lim_{t \to \infty} \frac{Var\{N(0,t]\}}{t} = \infty \qquad \text{if and only if} \qquad \lim_{t \to \infty} \left(U(t) - \frac{t}{\mu} \right) = \infty, \qquad (2.1)$$

where the first limit is the definition (Definition 2.3) of LRcD.

Now, if $\sigma^2 = Var\{X^2\}$, then for a renewal point process (Feller, 1971, Chapter XI, Section 3, Theorem 1),

$$0 \leq U(t) - \frac{t}{\mu} \to \frac{\sigma^2 + \mu^2}{2\mu^2} \qquad \text{as } t \to \infty,$$

where the right side is replaced by ∞ if $Var\{X^2\}$ does not exist. Therefore,

$$\lim_{t \to \infty} \left(U(t) - \frac{t}{\mu} \right) = \infty \qquad \text{if and only if} \qquad E\{X^2\} = \infty. \qquad (2.2)$$

Putting (2.1) and (2.2) together yields the desired result. □

Remark. It may prove instructive to the reader to consider the following alternate proof of the fact that an LRcD renewal point process necessarily has infinitely variable interpoint intervals. This proof requires only a well known application of the central limit theorem to renewal point processes (e.g. Cox, 1967, p. 40; Feller, 1971, p. 372). If the variance of the generic interpoint interval random variable X is finite, then, for large t, $N(0,t]$ is asymptotically normally distributed with mean t/μ and variance $\sigma^2 t/\mu^3$, where μ and σ^2 are the mean and variance of X, respectively. Therefore, if $E\{X^2\} < \infty$, then

$$\lim_{t \to \infty} \frac{Var\{N(0,t]\}}{t} = \frac{\sigma^2}{\mu^3} < \infty,$$

and the process is not LRcD. Thus, if the renewal point process *is* LRcD, then $E\{X^2\} = \infty$.

Remark. Theorem 2.1 is restricted to stationary processes. However, since LRcD is a property of the limiting behavior of the process, the renewal point process need only be asymptotically stationary for the result to apply. This proves useful, for example, since we often start a renewal point process with a point at the origin. Such a process is not stationary, but is asymptotically stationary.

2.3   Long-Range Dependence in Cortical Spike Trains

The spike trains of many neurons possess long-range dependence, although this term is infrequent in the neurophysiological literature. Sometimes this property is called long-term correlation, a designation that is quite similar to long-range dependence, but in other contexts more specialized terms such as second-order self-similarity, fractal behavior, and power-law (second-order) statistics are used. The use of the latter terms derives from the fact that the correlation function (and other equivalent second-order statistical functions) of a long-range dependent process usually assumes the form of a power-law function. Also, since the power-law function of the power spectral density, the Fourier transform of the correlation function, has the form of $1/f^\alpha$ for $0 < \alpha < 2$ under these conditions, long-range dependent spike trains are sometimes designated as having $1/f$ fluctuations.

A large majority of the spike trains from mammalian sensory systems that have been assayed for long-range dependence also exhibit it, and long-range dependence has also been found in the visual system of certain insects (Turcott et al., 1995). Long-range dependence has been found at many different levels of the mammalian visual and auditory systems, including the primary auditory nerve (Teich, 1989; Teich & Lowen, 1994; Lowen & Teich, 1996b), the lateral superior olive (Teich et al., 1990; Turcott et al., 1994), the retina and lateral geniculate nucleus (Teich, 1996; Teich et al., 1997; Lowen et al., 2001), and the visual cortex (Teich et al., 1996). Furthermore, long-range dependence has been found in somatosensory cortex (Wise, 1981) and the mesencephalic reticular formation (Gruneis et al., 1989, 1990, 1993). However, one notable exception to the existence of long-range dependence in mammalian sensory neurons is the peripheral vestibular system (Teich, 1989).

The papers mentioned above primarily analyze the count statistics of spike trains, but none of these papers refer to long-range count dependence as formulated in Definition 2.3. However, they all illustrate statistical characteristics of spike trains that are related to long-range dependence. Often the Fano factor curves,

$$F(t) = \frac{Var\{N(0,t]\}}{E\{N(0,t]\}},$$

for these spike trains are shown to approximate a power-law for large counting times, t. Specifically,

$$F(t) \sim t^\alpha, \quad \text{for } \alpha > 0, \quad \text{as } t \to \infty.$$

More generally, however, the Fano factor curves of these spike trains are shown to have a form that diverges for large counting windows. This, according to the following lemma, implies that the spike trains exhibit LRcD.

Lemma 2.2. *If the Fano factor curve, $F(t)$, of a stationary, orderly point process with finite rate $E\{N(0,1]\} \neq 0$ diverges in the limit of large t, i.e.*

$$\lim_{t \to \infty} F(t) = \infty, \tag{2.3}$$

then the point process is LRcD.

Proof. Let $\rho = E\{N(0,1]\}$ be the rate of the point process. Thus (e.g. Cox & Isham, 1980, p. 31),
$$E\{N(0,t]\} = \rho t.$$
So, by the definition of $F(t)$,

$$F(t) = \frac{Var\{N(0,t]\}}{E\{N(0,t]\}} = \frac{1}{\rho} \cdot \frac{Var\{N(0,t]\}}{t}$$

Hence, using equation (2.3),

$$\frac{1}{\rho} \cdot \lim_{t \to \infty} \frac{Var\{N(0,t]\}}{t} = \lim_{t \to \infty} F(t) = \infty.$$

Therefore, since ρ is a finite, nonzero constant,

$$\lim_{t \to \infty} \frac{Var\{N(0,t]\}}{t} = \infty,$$

and by definition the point process is LRcD. $\square$

Another statistic of spike trains that is often shown to approximate a power-law is the power spectral density. In this case, what is shown is that

$$S(f) \sim 1/f^{\alpha}, \quad \text{for } \alpha > 0, \quad \text{as } f \to 0.$$

In general, however, a power spectral density that diverges in the zero frequency limit is sufficient to imply that the spike train is LRcD.

Lemma 2.3. *If the power spectral density, $S(f)$, of a stationary, orderly[5] point process of finite (non-zero) rate has a singularity at zero frequency, i.e.*

$$\lim_{f \to 0} S(f) = \infty, \tag{2.4}$$

then the point process is LRcD.

---

[5]A point process $N(\cdot)$ on the real line is orderly if $\Pr\{N(t, t+\delta] > 1\} = o(\delta)$, for all $t \in \mathbb{R}$. This implies that the point process has no multiple simultaneous occurrences (see, e.g., Cox & Isham, 1980; Daley & Vere-Jones, 1988).

Proof. Let $V(t) = Var\{N(0, t]\}$ be the variance function, and let $V'(t)$ be the first derivative of the variance function. Then (Cox & Lewis, 1966, pp. 74–75)

$$\lim_{t\to\infty} V'(t) = \pi \cdot \lim_{f\to 0} S(f) = \infty,$$

where the final equality follows from (2.4). Thus,

$$\lim_{t\to\infty} \frac{Var\{N(0, t]\}}{t} = \lim_{t\to\infty} \frac{V(t)}{t} = \lim_{t\to\infty} \frac{1}{t}\int_0^t V'(u)du = \infty.$$

$\square$

Thus, we have shown that the previously mentioned analyses of spike trains, whether using the Fano factor curve or the power spectral density, imply that those spike trains are LRcD. In particular, in this chapter we are concerned with the fact that cortical neurons exhibit LRcD, and we wish to determine whether high-variability cortical models also exhibit LRcD. In the following two sections, we show that a certain class of high-variability cortical models, those that produce renewal point process outputs, only produce LRcD under conditions that produce erroneous CV_{ISI} values, at least according to published empirical measurements. We will not concern ourselves with whether cortical neurons also exhibit LRiD, though they likely do. This would require additional complexity in our arguments that is unnecessary for showing the incompatibility of these models with empirical long-range dependence analyses.

2.4 The Balanced Excitation-Inhibition Model

Gerstein and Mandelbrot (1964) first proposed that the output of a neuron is highly variable if there is a balance between the amounts of excitation and inhibition in its synaptic inputs. Later, Shadlen and Newsome (1994) used this idea as a solution to the apparent incompatibility, noted by Softky and Koch (1992, 1993), of temporal integration and high variability for cortical neurons. Work on this solution was further extended in several subsequent studies (Brown & Feng, 1999; Feng & Brown, 1998a; Feng, 1999; Shadlen & Newsome, 1998; Burkitt, 2000, 2001).

In this section, we will evaluate whether excitation-inhibition balance can produce LRcD, in addition to high variability, in simple IF models. We restrict ourselves to the simplest of these models, since (i) they can be handled analytically, (ii) they include the primary models considered in the literature, and (iii) more complex models are covered by the arguments for generic renewal point processes in the subsequent section. This exercise will serve to illustrate the conflict between LRcD and realistic values of CV_{ISI} as it occurs in a particular renewal model.

The model that we are considering here is the basic IF model, without leakage or reversal potentials, that has both excitatory and inhibitory inputs, all of which are Poisson processes and mutually independent. Each event, or "spike", arriving on an input causes a postsynaptic potential, a depolarizing potential (EPSP) for an excitatory input and a hyperpolarizing potential (IPSP) for an inhibitory input. Both EPSPs and IPSPs are Dirac delta functions, causing instantaneous jumps in the voltage, and, for simplicity and

because exact analytical results are available in this case, we begin with EPSPs and IPSPs that are equal in amplitude. We denote the EPSP amplitudes by $a_E > 0$ and the IPSP amplitudes by $a_I > 0$. Thus, in the present case, we can let $a = a_E = a_I$. Furthermore, we let M_E and M_I denote the number of excitatory and inhibitory inputs, respectively, and λ_E and λ_I denote the input rate for each excitatory and inhibitory fiber, respectively. Since all inputs are independent Poisson processes, they can be combined such that only two effective inputs need be considered: an excitatory input of rate $\Lambda_E = M_E \lambda_E$ and an inhibitory input of rate $\Lambda_I = M_I \lambda_I$.

The integral, with respect to time, of all of the resulting postsynaptic potentials is a random walk, which forms the time-varying potential, $V(t)$, of the IF model. When this potential crosses a predetermined, constant threshold, V_{th}, an output spike is initiated. Following an output spike, the voltage is reset to its resting level, v_0, and the process starts anew.

Since the inputs are Poisson processes, the postsynaptic potentials are delta functions, and the IF neuron completely resets at each occurrence of an output spike, the output of the present model is clearly a renewal process. Hence, the output is completely specified by its interval distribution, or, equivalently, the first passage time of the potential $V(t)$ to level V_{th} from $V(0) = v_0$.

Tuckwell (1988, pp. 128ff) has derived the interval density, $f(t)$, for the output of this model:

$$f(t) = \hat{\theta} \left(\frac{\Lambda_E}{\Lambda_I} \right)^{\hat{\theta}/2} \frac{e^{-(\Lambda_E + \Lambda_I)t}}{t} I_{\hat{\theta}}(2\sqrt{\Lambda_E \cdot \Lambda_I}\, t), \qquad t > 0,$$

where $\hat{\theta} = \left\lceil \frac{V_{th} - v_0}{a} \right\rceil$ is the number of excitatory inputs that are required for $V(t)$ to cross threshold, $\lceil x \rceil$ is the least integer greater than or equal to x, and $I_\rho(x)$ is the modified Bessel function

$$I_\rho(x) = \sum_{k=0}^{\infty} \frac{1}{k!\, \Gamma(k + \rho + 1)} \left(\frac{x}{2} \right)^{2k+\rho}.$$

Now, let X be an arbitrary interspike interval with density function $f(t)$. Then, according to Tuckwell (1988), the first three (central) moments of X are

$$\Pr\{X < \infty\} = \begin{cases} 1, & \text{if } \Lambda_E \geq \Lambda_I, \\ \left(\frac{\Lambda_E}{\Lambda_I} \right)^{\hat{\theta}}, & \text{if } \Lambda_E < \Lambda_I \end{cases},$$

$$E\{X\} = \begin{cases} \frac{\hat{\theta}}{\Lambda_E - \Lambda_I}, & \text{if } \Lambda_E > \Lambda_I, \\ \infty, & \text{if } \Lambda_E \leq \Lambda_I \end{cases},$$

and

$$Var\{X\} = \begin{cases} \frac{\hat{\theta}(\Lambda_E + \Lambda_I)}{(\Lambda_E - \Lambda_I)^3}, & \text{if } \Lambda_E > \Lambda_I, \\ \infty, & \text{if } \Lambda_E \leq \Lambda_I \end{cases}.$$

Thus, in the case when $\Lambda_E \leq \Lambda_I$, the coefficient of variation of the interspike intervals, CV_{ISI}, does not exist. However, if $\Lambda_E > \Lambda_I$, the coefficient of variation is

$$CV\{X\} = \sqrt{\frac{\Lambda_E + \Lambda_I}{\hat{\theta}(\Lambda_E - \Lambda_I)}}.$$

Therefore, by adjusting Λ_I within the interval $[0, \Lambda_E)$, the CV_{ISI} can be set to any value in the interval $[1/\sqrt{\hat{\theta}}, \infty)$, with arbitrarily large values occurring as Λ_I approaches Λ_E. In other words, as the amount of inhibition is increased to bring the model into a perfect balance between excitation and inhibition, the CV_{ISI} goes to infinity.

Now we wish to know when the output of this process exhibits LRcD. Since the output of the model is a renewal point process, by Theorem 2.1, we know that the output will be LRcD if and only if $E\{X^2\} = \infty$, assuming that $E\{X\} < \infty$. Hence, the model can only be LRcD under the condition that $\Lambda_E \leq \Lambda_I$.[6] Thus, in trying to fit this IF model to both the CV_{ISI} and LRcD properties of real neurons, we find ourselves at an impasse. With the amount of inhibition less than the amount of excitation, the model can match the CV_{ISI} values measured from real neurons, but under these conditions it is not LRcD, since both the mean and variance are finite. By making the amount of inhibition equal to or greater than the amount of excitation, we may create LRcD in the model, but the CV_{ISI} becomes infinite. Therefore, this model of highly variable cortical neurons cannot manifest LRcD, while still producing interspike-interval variability consistent with empirical measurements.

Now suppose that we relax the condition that EPSPs and IPSPs have equal magnitude. In other words, consider the case when $a_E \neq a_I$. This is the model that was proposed by Shadlen and Newsome (1994) for matching the variability of real cortical neurons. This model, unlike the case when $a_E = a_I$, cannot be analyzed directly. However, the potential V(t) in this case can be approximated by a Wiener process with drift, an approximation that becomes exact as the input rates Λ_E and Λ_I go to infinity and the PSP amplitudes a_E and a_I go to zero. Thus, we have that (Tuckwell, 1988, p. 135ff)

$$V(t) \sim U(t) = v_0 + \sigma W(t) + \mu t, \qquad t > 0,$$

where $W(t)$ is the standard Wiener process (or Brownian motion) and the mean and variance are given by

$$\mu = a_E \Lambda_E - a_I \Lambda_I \qquad \text{and} \qquad \sigma^2 = a_E^2 \Lambda_E + a_I^2 \Lambda_I,$$

respectively. For this Wiener process approximation, the first-passage time, or interval, density is given by Tuckwell (1988):

$$f(t) = \frac{\theta}{\sqrt{2\pi\sigma^2 t^3}} \exp\left[-\frac{(\theta - \mu t)^2}{2\sigma^2 t}\right], \qquad t > 0,$$

where $\theta = V_{th} - v_0$ does not depend on the PSP amplitudes since we no longer have the simplifying assumption that they are all equal. Furthermore, the first three moments of

---

[6]Although the model is quite likely to exhibit LRcD under these conditions, Theorem 2.1 does not apply since $E\{X\} = \infty$ when $\Lambda_E \leq \Lambda_I$.

the random variable X with density $f(t)$ are (Tuckwell, 1988)

$$
\Pr\{X < \infty\} =
\begin{cases}
1, & \text{if } \mu \geq 0, \\
\exp\left[\frac{2\mu\theta}{\sigma^2}\right], & \text{if } \mu < 0
\end{cases}
$$

$$
=
\begin{cases}
1, & \text{if } a_E\Lambda_E \geq a_I\Lambda_I, \\
\exp\left[-\frac{2(a_I\Lambda_I - a_E\Lambda_E)\theta}{a_E^2\Lambda_E + a_I^2\Lambda_I}\right], & \text{if } a_E\Lambda_E < a_I\Lambda_I
\end{cases},
$$

$$
E\{X\} =
\left\{
\begin{array}{ll}
\frac{\theta}{\mu}, & \text{if } \mu > 0, \\
\infty, & \text{if } \mu \leq 0
\end{array}
\right\}
=
\begin{cases}
\frac{\theta}{a_E\Lambda_E - a_I\Lambda_I}, & \text{if } a_E\Lambda_E > a_I\Lambda_I, \\
\infty, & \text{if } a_E\Lambda_E \leq a_I\Lambda_I
\end{cases},
$$

and

$$
Var\{X\} =
\left\{
\begin{array}{ll}
\frac{\theta\sigma^2}{\mu}, & \text{if } \mu > 0, \\
\infty, & \text{if } \mu \leq 0
\end{array}
\right\}
=
\begin{cases}
\frac{\theta(a_E^2\Lambda_E + a_I^2\Lambda_I)}{a_E\Lambda_E - a_I\Lambda_I}, & \text{if } a_E\Lambda_E > a_I\Lambda_I, \\
\infty, & \text{if } a_E\Lambda_E \leq a_I\Lambda_I
\end{cases}.
$$

Thus, in the case when $\mu \leq 0$, or equivalently $a_E\Lambda_E \leq a_I\Lambda_I$, the coefficient of variation of the interspike intervals, $CV(ISI)$, does not exist. However, if $\mu > 0$, i.e. $a_E\Lambda_E > a_I\Lambda_I$, the coefficient of variation is

$$
CV\{X\} = \sqrt{\frac{\sigma^2}{\theta\mu}} = \sqrt{\frac{a_E^2\Lambda_E + a_I^2\Lambda_I}{\theta(a_E\Lambda_E - a_I\Lambda_I)}}. \tag{2.5}
$$

Therefore, we see that the CV_{ISI} may take any value in the interval $\left[\sqrt{a_E/\theta}, \infty\right)$ when $a_E\Lambda_E > a_I\Lambda_I$, with arbitrarily large values occurring as $a_I\Lambda_I$ approaches $a_E\Lambda_E$.

Thus, we are in the same predicament as before. If $a_E\Lambda_E > a_I\Lambda_I$, then the model can be adjusted to match the CV_{ISI} values as measured from real neurons. However, under this condition, both the mean and variance are finite, and, therefore, the model is not LRcD. On the other hand, if $a_E\Lambda_E \leq a_I\Lambda_I$ the process may be LRcD, but the CV_{ISI} is infinite. Therefore, even when the EPSP and IPSP are unequal, the IF model fails to match both the interspike-interval variability and LRcD as measured in real cortical neurons.

2.5   General Principles for Renewal Models

In the previous section, we showed that the basic high-variability IF model that requires balanced excitation and inhibition cannot produce both a finite CV_{ISI} and LRcD at the same time. However, this result is easily extended to all renewal models, a class that includes a significant portion of the single-neuron high-variability IF models. By using the term "single-neuron" we are excluding consideration of network models, where the statistical nature of the entire set of inputs to each neuron is not explicitly specified, but instead consists, at least partially, of outputs from other similar neurons within an interconnected network.

If each component of an IF model is renewal, i.e. has no memory of the past beyond the last output spike, then output of this model must be renewal. Hence, the following conditions are jointly sufficient to render the output of such a model renewal:

(i) all inputs to the IF neuron are Poisson processes;

(ii) the cross-covariance between any set of inputs is zero for any non-zero lag;

(iii) the PSPs are direct changes in the IF potential, and either the IF potential is reset to a fixed value after the occurrence of each output spike or these reset values form a set of independent and identically distributed random variables;

(iv) all other parameters of the model, e.g. threshold, leakage conductance, or reversal potentials, either are constant, are set to a fixed value upon the occurrence of an output spike, or have post-output-spike values that form a set of independent and identically distributed random variables.

These conditions are not entirely general, but they are practically general for the high-variability IF models that have been studied in the literature. Note that if the above conditions are met, it does not matter whether the model has leakage or reversal potentials or a dynamic threshold, as long as their parameters meet condition (iv). Furthermore, the PSPs may have non-zero duration as long as condition (iii) is met.

Many of the non-network high-variability IF models in the literature meet these four conditions and are therefore renewal. Excitation-inhibition balance models that fit this category were considered by Shadlen and Newsome (1994), Brown and Feng (1999), Feng and Brown (1998a, 1999), Feng (1999), Burkitt (2000, 2001) . Renewal models that produced highly variable outputs due to correlations between inputs were considered by Feng and Brown (2000a), Feng (2001), Feng and Zhang (2001), Salinas and Sejnowski (2000). Other models that are renewal as well are the partial reset models considered in Troyer and Miller (1997), Bugmann et al. (1997), the time-varying threshold model considered in Wilbur and Rinzel (1983), and the nonlinear leakage model considered in Feng and Brown (2000b).

From Theorem 2.1, we see that if any of these models have an interval distribution with a finite mean, then either (i) their interval distribution also has a finite variance and they are not LRcD or (ii) their interval distribution has infinite variance and they are LRcD. In case (i), it might be possible to match empirically measured values of CV_{ISI}, but the LRcD property of real neurons is unattainable. However, in case (ii), the model will be LRcD, but it will be impossible to match empirically measured CV_{ISI} values. Thus, just like the model that was considered in Section 2.4, all renewal models with finite-mean intervals fail to match both the interspike-interval variability and LRcD as measured in real cortical neurons.

The preceding argument assumes that the interval distribution has a finite mean. The example in Section 2.4, where the limit of CV_{ISI} as the model approached the infinite-mean condition could be determined analytically, shows that at least some renewal models with infinite mean intervals fail to produce the required variability and LRcD properties, and fail in a manner similar to that described in the preceding argument for the finite-mean case. This, however, does not prove that all such models fail in this way. Nevertheless, models with infinite mean intervals have, at least, a couple degenerate

properties. First, such a model cannot be stationary. This means, for instance, that if we were to analyze or simulate a renewal point process with an infinite mean interval, we would need to begin with a point at some specified time, for instance at the origin. Second, since an interval distribution with an infinite mean will also have an infinite variance, the CV_{ISI} has no meaning for such a model. Therefore, whether or not the model is LRcD, a direct comparison of the interval variability of the model with that of real cortical neurons is impossible using a single measure.

To test whether such a model is reasonable, we should consider the behavior of both the sample mean and the sample variance of interspike intervals from physiological recordings as an increasing number of interspike intervals are included in the calculation. If both increase without bound for long recordings from cortical neurons, then it is possible that a renewal model with infinite mean interval would be reasonable. Even so, unless the sample CV_{ISI} values (the sample standard deviation divided by sample mean) converge to a finite value as the amount of data included in the calculations is increased, these values are unfit as model constraints. However, to the best of our knowledge, this convergence analysis has not been carried out on spike trains recorded from cortical neurons.

2.6   Discussion

A number of different types of integrate-and-fire models have been created in order to explain how cortical neurons can integrate over large numbers of inputs while still producing highly variable outputs. Although these models can produce values for the coefficient of variation of interspike intervals similar to those calculated from *in vivo* cortical spike trains, we considered whether such models can also produce long-range dependence in their spike counts, a property that also is known to exist in cortical spike trains. Based on the observation that a large class of these models produce outputs that are renewal point processes, we were able to prove analytically that the output of these models cannot simultaneously have both a finite coefficient of variation of interspike intervals and long-range dependence in its spike counts. Therefore, assuming that the spike trains of *in vivo* cortical neurons are long-range count dependent, none of these renewal models produce highly variable outputs in a way that is consistent with the properties of cortical spike trains. Thus, we suggest that their success in representing the cortical processing that leads to highly variable spike trains is doubtful.

The approach that we have taken in this chapter has allowed us to handle a large number of models analytically without having to deal with the details of each specific model. In fact, most of these models are analytically intractable in detail. Furthermore, the consideration of a general statistical description that is applicable to many models adds to our intuition of cortical variability and narrows the range of feasible models for future studies. The arguments presented in this chapter highlight the need for more thorough study of the variability of cortical spike trains before judging the validity of models of this variability and drawing general conclusions from them. In particular, the insufficiency of the CV_{ISI} estimator alone as a measure of this variability and the necessity of also using a measure that is sensitive to the correlational structure of the spike trains is

demonstrated.

We have assumed, for the sake of argument, that empirical measurements of the CV_{ISI} of cortical spike trains, which have resulted in values typically in the range from 0.5 to 1.2, are valid. However, another possible scenario exists. If the "real" CV_{ISI} of cortical spike trains was actually very large, such that for all practical purposes it could be assumed infinite, estimates of CV_{ISI} from short-duration recordings of cortical spike trains might still be small. Under such conditions, it is theoretically possible that cortical spike trains could be long-range count dependent, but that empirical estimates of CV_{ISI} could be on the order of one. However, since CV_{ISI} estimates in this case would increase with increasing recording duration, and since such estimates have been measured under numerous recording conditions by many different researchers, this potentiality seems improbable. More likely, under such conditions, empirical estimates of CV_{ISI} would be widely variable and span a much larger range than 0.5 to 1.2. However, to rule out this possibility, the behavior of the CV_{ISI} estimate as an increasing amount of the recording duration is included in the calculation should be analyzed in future studies in order to determine whether it converges or diverges.

By considering long-range count dependence, in addition to the CV_{ISI}, we were able to show that a large portion of the single neuron models of cortical high variability are incompatible with cortical spike trains. The incompatibility of one of these models, the standard leaky integrate-and-fire model with Poisson inputs, with cortical spike trains was also found by Shinomoto and his colleagues (Shinomoto & Sakai, 1998; Shinomoto, Sakai, & Funahashi, 1999). They argued that the leaky integrate-and-fire model could not match both the coefficient of variation and the skewness coefficient of the interspike-interval density of cortical spike trains. The skewness coefficient, or just skewness, is the third central moment divided by the standard deviation cubed and is a measure of the symmetry of a distribution. Our result, however, has several advantages over the arguments in these papers. First, our result is more generally applicable. Second, their arguments were based on an approximation to the leaky integrate-and-fire model, whereas ours are directly applicable to the models themselves. Third, in order to discount the models, they placed restrictions on the parameter ranges of the models, something that was unnecessary in our arguments.

Our results apply only to renewal models, but other models of the high variability present in cortical spike trains exist that do not necessarily produce renewal outputs. In particular, one type of model is almost identical to the models considered in Section 2.4, except that the inputs are not necessarily Poisson processes. Instead, in these models, the inputs are allowed to be any renewal point process (Feng, 1997; Feng & Brown, 1998b; Feng et al., 1998). Therefore, since a Poisson process is the only type of renewal point process with no memory, or temporal correlation, non-Poisson renewal inputs render the entire model non-renewal and the arguments of this chapter do not apply. Furthermore, even this slight increase in complexity renders analytical treatment of the basic IF model, like that in Section 2.4, intractable. In Chapter 3, we will, therefore, investigate these models using a combination of simulations and partial analysis to determine their validity as explanations of high cortical variability.

Another set of high-variability models consist of interconnected networks of integrate-and-fire neurons. In these models, the output of each individual neuron is highly variable

due to the complex network dynamics. The output of each neuron in such a network is unlikely to be a renewal point process, since temporal correlations are probably present in the network dynamics. However, only one of these models has been tested for long-range dependence. In simulations of their network, Usher and his colleagues (Usher et al., 1994, 1995) demonstrated that the outputs of individual units had power spectrums that behaved like power-laws at low frequencies while the coefficient of variations of their interspike intervals were still approximately in the range of those measured from cortical spike trains. Thus, since a power spectrum that follows a power-law demonstrates the presence of long-range count dependence (see Section 2.3), their model accomplishes exactly what we have shown to be impossible for renewal models. However, a further test of their model, which was unnecessary when considering renewal models, would be to compare both the long-range count dependence and long-range interval dependence of its individual outputs with measurements from cortical spike trains. This is currently the topic of ongoing work in our lab.

As reviewed in Section 2.3, many non-cortical sensory neurons are known to exhibit long-range count dependence, including neurons in the lateral geniculate nucleus, which are inputs for visual cortex. This is suggestive of a solution to the problem raised in this chapter. We speculate that if the inputs to an integrate-and-fire model had more realistic long-range dependent properties, similar to those found in the neurons that project into the cortex, then the output of the model would also have variability and long-range count dependence in accord with cortical spike trains. The answer to this proposal is beyond the scope of this chapter, as it requires consideration of deeper issues related to the long-range dependence of spike trains. In Chapter 3, we will develop the theory necessary to evaluate this proposal and will then study its merits. Nevertheless, Sakai et al. (1999) and Shinomoto and Tsubo (2001) have shown that temporally correlated inputs to leaky integrate-and-fire models can produce not only coefficients of variation and skewness coefficients of interspike intervals, but also correlation between consecutive interspike intervals, that match those estimated from cortical spike trains. Although their study of the correlation between consecutive interspike intervals would not reveal long-term correlations, their result lends support to the idea that long-term correlations in the inputs will result in long-term correlations in the output of integrate-and-fire models.

Chapter 3

Long-Range Dependence and Models of the High Interspike Interval Variability of Cortical Neurons II: Models with Non-Poissonian Inputs

ABSTRACT The empirical observation that the variability of the interspike intervals of cortical spike trains is higher than expected from the classical integrate-and-fire model has led to a number of proposals for modifications to this model that increase its output variability. However, we have suggested previously that long-range dependence should also be considered when matching models to the variability of cortical spike trains, and have shown that an entire class of integrate-and-fire models are incapable of producing long-range dependence while simultaneously exhibiting interspike interval variability that matches empirical measurements. Here we study two integrate-and-fire models that are not contained within this class of models, one previously studied by Feng and Brown and one that is new, in order to evaluate whether they are capable of producing both long-range dependence and realistic interspike interval variability simultaneously. Whereas these "high-variability" integrate-and-fire models usually have Poisson process inputs, the first model has non-Poissonian renewal point process inputs, and the second model has fractional-Gaussian-noise-driven Poisson process inputs. Non-Poissonian renewal point processes are common generalizations of the interval independence property of the Poisson process, and thus the first model is a natural extension of standard high-variability integrate-and-fire models. The fractional-Gaussian-noise-driven Poisson process was chosen for the second model because it shares many statistical attributes with sub-cortical neurons in sensory pathways, including inputs to the cortex. The confluence of our analytical and simulation results implies that the renewal-input model is capable of producing high variability and long-range dependence comparable to that seen in spike trains recorded from cortical neurons, but only if the interspike intervals of the inputs have infinite variance. On the other hand, the second model is able to produce similar results with inputs that have finite-variance interspike intervals. Hence, we suggest that the second model is to be preferred, since its inputs are better

justified based on physiological recordings.

3.1 Introduction

The integrate-and-fire (IF) model is a common model of general neuronal processing when simplicity is of the essence. Its simplicity is particularly beneficial when one wishes to model a large network of interconnected neurons, such as portions of the cortex. The essential features of neuronal dynamics that the IF model retains are a potential that is the result of the integration of a number of inputs and the production of a stereotypical output event when this potential crosses a threshold. Furthermore, this basic IF model can be modified to incorporate additional features that one may also deem essential, the most common being leakage that causes the potential to decay to its resting state when no input is applied.

Each neuron in the cortex receives a very large number of inputs, but produces a relatively low output firing rate, which is on the order of that of each individual input. In order for the standard, nonleaky IF model to have a large number of inputs and an output spike rate like its inputs, it must require a large number of input spikes to produce each output spike. Thus, although the inputs may contain a high degree of variability, the output will typically be very regular due to the averaging mechanism of the integrator. However, it has been known for some time that the spike trains of cortical neurons are quite variable. This discrepancy can be easily explained with a leaky IF model if the rate of decay of the potential is short relative to the time between output spikes. This explanation seems reasonable for the conditions under which the variability of cortical neurons has typically been measured, when the firing rates have been low. However, Softky and Koch (1992, 1993) showed that this high variability, as measured by the coefficient of variation (the standard deviation divided by the mean) of the interspike intervals (CV_{ISI}), persists even when cortical neurons produce high firing rates with interspike intervals that are short relative to reasonable decay times due to membrane leakage. This apparent contradiction between the IF model and physiological measurements suggested that the IF model does not represent the essential nature of neuronal dynamics.

Since this finding, there have been many responses to this paradox. Softky and Koch (1992, 1993) suggested that cortical neurons act as coincidence detectors instead of as integrators, where the coincidence detection mechanism is likely to be located in the dendrites (see, also, Softky, 1994, 1995). Others (Shadlen & Newsome, 1994; Brown & Feng, 1999; Feng & Brown, 1998a; Feng, 1999; Shadlen & Newsome, 1998; Burkitt, 2000, 2001) have shown that the notion of the cortical neuron as integrator may be retained if the amounts of excitation and inhibition at the input are approximately balanced. Furthermore, a number of modifications to the IF model have been postulated to reconcile it with the high variability of cortical neurons (reviewed in Section 2.1).

Almost without exception, the comparisons with respect to variability that have been made between IF models and cortical neurons have used the CV_{ISI}. However, measures of the variability of interspike intervals do not exhaustively describe even the second-order statistical structure of these intervals, much less of the spike train itself.

In Chapter 2, we suggested that another important statistical aspect of cortical spike trains that should be considered when building models is long-range temporal dependence. Long-range dependence (LRD) denotes that the dependence between distant random variables in a stochastic process decreases "slowly" as the distance is increased, and is sometimes referred to as long memory. In particular, finite variance processes with long-term correlation or second-order self-similarity are LRD. Evidence of LRD has been found in the neurons of many sub-cortical sensory pathways (Teich, 1989; Teich & Lowen, 1994; Lowen & Teich, 1996b; Teich et al., 1990; Turcott et al., 1994; Teich, 1996; Teich et al., 1997; Lowen et al., 2001), as well as in the activity of cortical neurons (Teich et al., 1996; Gruneis et al., 1989, 1990, 1993; Wise, 1981).

In Chapter 2, we showed analytically that an entire class of IF models, which includes a large portion of those that have been used to explain high cortical variability, are incapable of producing long-range dependence while simultaneously exhibiting interspike interval variability that matches empirical measurements. The argument is based on the observation that many of the high-variability IF models produce outputs which are renewal point processes, since Poisson process inputs do not have any memory and the remainder of the model resets after each output spike. Furthermore, we showed, based on the arguments of Daley (1999), that a renewal point process is LRD if and only if the interpoint intervals have infinite variance, and thus infinite CV_{ISI}. Hence, such IF models are incapable of producing outputs that have values of CV_{ISI} in the range of those measured from cortical spike trains (usually within the range of 0.5 to 1.5) while still exhibiting LRD.

There are two general types of high-variability IF models in the literature to which the arguments of Chapter 2 do not pertain: the models of Feng and his coworkers (Feng, 1997; Feng & Brown, 1998b, 1998a; Feng et al., 1998) which use renewal point processes other than the Poisson process as inputs and models which consist of large networks of interconnected IF neurons that produce high variability through network dynamics (Usher et al., 1994, 1995; Tsodyks & Sejnowski, 1995; Vreeswijk & Sompolinsky, 1996, 1998). In this chapter, we will study the ability of the former set of models to produce physiological values of CV_{ISI} and LRD; the latter set is the subject of future work.

Feng and his colleagues used two types of non-Poissonian renewal point processes as inputs to their IF model. One had interspike intervals distributed according to the positive Gaussian distribution, and the other had interspike intervals distributed according to a Pareto distribution. The significance of these choices is that the positive Gaussian distribution has a shorter tail than the exponential distribution, the distribution of intervals for a Poisson process, and the Pareto distribution has a longer tail than the exponential distribution. Since the superposition of non-Poissonian renewal processes is not a renewal process, the cumulative inputs to the integrate-and-fire neuron in these model is not renewal, having a dependence structure that is longer than the interspike intervals. Hence, the arguments of Chapter 2 do not apply.

Feng and his coworkers found that both of these models, the one with positive-Gaussian-renewal inputs and the one with Pareto-renewal inputs, can produce values of CV_{ISI} similar to those measured in actual cortical neurons, if the amounts of excitation and inhibition are appropriately balanced. With positive-Gaussian-renewal inputs, the IF model requires a larger amount of inhibition, closer to the amount of excitation, than

does the standard IF model with Poisson inputs. On the other hand, with Pareto-renewal inputs, the IF model requires less inhibition, farther from the amount of excitation, than does the Poisson-input model. For any of these models, if the amount of inhibition is too large, the CV_{ISI} of the output of the model will be much larger than the values that have been measured *in vivo*. Thus, a specific goal will be to determine whether the outputs of these models exhibit LRD when the excitation and inhibition levels are such that the model produces physiological values of CV_{ISI}.

Furthermore, we will describe and study a new model for the high variability of cortical neurons and test its ability to produce both LRD and physiological values of CV_{ISI}. In this model, instead of changing the interspike interval distribution of the inputs directly, we will create dependence between the intervals within each input. In fact, the dependence will be strong enough that each input will be LRD. This model is motivated by the fact that LRD of this type has been found in a majority of the sub-cortical sensory neurons that have been assayed for it (Teich, 1989; Teich & Lowen, 1994; Lowen & Teich, 1996b; Teich et al., 1990; Turcott et al., 1994; Teich, 1996; Teich et al., 1997; Lowen et al., 2001). This includes neurons in the lateral geniculate nucleus that serve as inputs to visual cortical areas (Teich, 1996; Teich et al., 1997; Lowen et al., 2001).

3.2   Long-Range Dependence in Point Processes

3.2.1   General Concept of Long-Range Dependence

For a stationary stochastic process, $\{X_i : i \in \mathbb{Z}\}$, with finite variance, LRD may be defined using the asymptotic behavior of any of three functions related to the second-order properties of the process (Cox, 1984; Beran, 1994). The most obvious definition says that such a process is LRD when the sum of the covariances (or correlations) diverges, i.e. when

$$\lim_{n \to \infty} \sum_{j=0}^{n} Cov\{X_0, X_j\} = \infty.$$

Since the power spectral density (or power spectrum) is the Fourier transform of the autocorrelation of a process, an equivalent definition is that the power spectral density of an LRD process has a pole at the origin. Finally, LRD may be defined using the asymptotic behavior of the variance of the sum of the random variables:

$$\lim_{n \to \infty} \frac{Var\left\{\sum_{j=1}^{n} X_j\right\}}{n} = \infty. \tag{3.1}$$

Whichever definition above is used, the essential feature of LRD is that the dependence between the random variables decays slowly as the distance between them increases, where the slowness of the decay that is necessary for finite variance processes to be LRD is specified by the definitions above. The absolute magnitude of the dependence for widely separated random variables is not the issue, only the relationship between the dependencies at different lags.

3.2.2 Types of Long-Range Dependence in Point Processes

For point processes two types of LRD are possible. A point process may be specified by either the joint distribution of the intervals between successive points or the joint distribution of the number of points in arbitrary sets. Accordingly, they may exhibit LRD in either the sequence of interpoint intervals or in the sequence of counts.

Daley and Vesilo (1997) have suggested definitions for these two types of LRD. To define LRD in the interpoint intervals, they considered the asymptotic behavior of the sum of the intervals.

Definition 3.1. *(Daley & Vesilo, 1997)* A stationary point process $N(\cdot)$ on the real line exhibits long-range interval dependence (LRiD) when the stationary sequence of interpoint intervals $\{Y_i\}$ with finite variances is LRD in the sense that

$$\mathsf{LRiD:} \qquad \limsup_{n\to\infty} \frac{Var\left\{\sum_{i=1}^{n} Y_i\right\}}{n} = \infty.$$

Since the count in an interval $(0, t]$, denoted by $N(0, t]$, may be written as the sum of the counts in $n \in \mathbb{N}$ equal-length, successive intervals, i.e.

$$N(0, t] = \sum_{i=1}^{n} N\left(\frac{i-1}{n} \cdot t, \frac{i}{n} \cdot t\right],$$

they suggest the following definition for LRD in the counts of a point process.

Definition 3.2. *(Daley & Vesilo, 1997)* A second-order stationary point process $N(\cdot)$ on the real line exhibits long-range count dependence (LRcD) when

$$\mathsf{LRcD:} \qquad \limsup_{t\to\infty} \frac{Var\left\{N(0, t]\right\}}{t} = \infty.$$

3.2.3 The Moment Index and the Hurst Index

The two types of LRD, LRcD and LRiD, are not independent. In order to study their relationship, it is helpful to define two concepts: the moment index and the Hurst index. The moment index, κ, of a non-negative random variable indicates which moments exist.

Definition 3.3. *(Daley, 1999)* The moment index of a random variable X with distribution function F is

$$\kappa \equiv sup\left\{k \geq 0 : \mu_k \equiv E\left(X^k\right) = \int_{-\infty}^{\infty} x^k dF(x) < \infty\right\}.$$

Thus, all moments less than κ exist, while all moments greater than κ are infinite. The κth moment can either exist or not. Of course, we may also say that κ is the moment index of the distribution of X, say. In the sequel, we will particularly be concerned with the moment index of the (marginal) distribution of the interpoint intervals of a point process, and, therefore, κ will always refer to this particular moment index.

The Hurst index, H, is a measure of the self-similarity of a stochastic process.

Definition 3.4. *(e.g. Beran, 1994; Samorodnitsky & Taqqu, 1994)* The real-valued stochastic process $Y(t)$, with continuous time parameter t, is self-similar with Hurst index H if for all time-rescaling factors $c > 0$,

$$c^{-H} Y(ct) \stackrel{d}{=} Y(t),$$

where "$\stackrel{d}{=}$" means "is equal in distribution to".

If the covariances of a self-similar process exist and decay to zero as the lag is increased, then (Beran, 1994, p.53)
$$0 < H < 1.$$

Since a (non-constant) self-similar process is necessarily nonstationary (Beran, 1994, pp. 50f), it is not directly useful in modeling data that appear stationary. However, a self-similar process $Y(t)$ can have stationary increments $X(t) \equiv Y(t) - Y(t - s)$, for some fixed $s > 0$. It is these stationary increment processes that are suitable for modeling long-range dependent processes.

Let $Y(t)$ be a self-similar stochastic process with stationary increments. Then an increment process can be defined by $X_i = Y(i) - Y(i - 1)$, and we will associate the Hurst index H of Y with X. The value of H indicates the type and strength of the dependence present in the stochastic process X. If $H = 0.5$, then X_i and X_j are uncorrelated for all $i \neq j$, i.e. the correlations are zero for all non-zero lags. An example of a self-similar process with $H = 0.5$ is Brownian motion, of which ("white") Gaussian noise is the stationary increment process. If $0 < H < 0.5$, then the correlations of the increment process are negative for large lags, a behavior often referred to as "negative dependence". In this case (Beran, 1994, p. 52),

$$\sum_{i=-\infty}^{\infty} Cov\{X_0, X_j\} = 0.$$

Since the addition of even a minute disturbance to such a process is deleterious to this property, it is very unstable. Thus, for all practical purposes, the increments of self-similar processes with $H < 0.5$ are not useful as models. Finally, if $0.5 < H < 1$, the correlations of the increment process are positive, and they decay to zero so slowly that

$$\sum_{i=-\infty}^{\infty} Cov\{X_0, X_j\} = \infty.$$

Thus, in this case, the increment process is LRD, and a larger value of H signifies stronger dependence.

In general, however, the Hurst index does not need to be associated with a self-similar process or its increments. Instead of being a "self-similarity parameter", it may be thought of as a "long-memory parameter" for stationary processes. In fact, the Hurst index derives its name from the exponent parameter that Hurst (1951) used to demonstrate the existence of long-range dependence in records of the level of the Nile River and records of other geophysical processes.

To see the analogy between the use of the Hurst index as a self-similarity parameter and its use as a long-memory parameter, consider the variance of the sample mean $\bar{X}$ for the increments of a self-similar process. For such a process (Beran, 1994, p. 54),

$$Var\{\bar{X}\} = Var\left\{\frac{1}{n}\sum_{i=1}^{n}X_i\right\} = \sigma^2 n^{2H-2},$$

where $\sigma^2 = Var\{X_i\}$, or

$$\frac{Var\left\{\sum_{i=1}^{n}X_i\right\}}{n} = \sigma^2 n^{2H-1}, \tag{3.2}$$

for any integer $n > 0$. If $H = 0.5$, then this equation becomes

$$\frac{Var\left\{\sum_{i=1}^{n}X_i\right\}}{n} = \sigma^2 < \infty,$$

which implies that

$$\lim_{n\to\infty}\frac{Var\left\{\sum_{i=1}^{n}X_i\right\}}{n} < \infty.$$

This inequality is the complete negation of the property (3.1) for a long-range dependent process. On the other hand, if $H > 0.5$, then the exponent of n in (3.2) is positive, and the limit as $n \to \infty$ is infinite, implying long-range dependence. Now, (3.2) may be written as

$$Var\left\{\sum_{i=1}^{n}X_i\right\} = \sigma^2 n^{2H}.$$

Therefore, an alternate definition of the Hurst index, which is more closely related to the concept of long-range dependence, is applicable to all stationary stochastic processes, but is still consistent with its prior interpretation in the context of self-similar processes, is

$$H = \sup\left\{h : \limsup_{n\to\infty}\frac{Var\left\{\sum_{i=1}^{n}X_i\right\}}{n^{2h}} = \infty\right\}.$$

For point processes, Daley (1999) has defined the Hurst index in an analogous manner with respect to its counting process.

Definition 3.5. *(Daley, 1999)* The Hurst index of a stationary, ergodic,[1] orderly[2] point process $N(\cdot)$ with finite second moment is

$$H = \sup\left\{h : \limsup_{t\to\infty}\frac{Var\{N(0,t]\}}{t^{2h}} = \infty\right\}.$$

---

[1] A stationary point process $N(\cdot)$ with finite mean density $m = E\{N(0,1]\}$ is ergodic if $\Pr\{\lim_{x\to\infty}N(0,x]/x = m\} = 1$ (Daley, Rolski, & Vesilo, 2000).

[2] A point process $N(\cdot)$ on the real line is orderly if $\Pr\{N(t, t + \delta] > 1\} = o(\delta)$, for all $t \in \mathbb{R}$. This implies that the point process has no multiple simultaneous occurrences (see, e.g., Cox & Isham, 1980; Daley & Vere-Jones, 1988).

Since, for the process assumed in this definition, $Var\{N(0,t]\} = o(t^2)$ for $t \to \infty$, the Hurst index must be no larger than one. On the other hand, as long as it is not the case that the point process is zero with probability one,

$$\limsup_{t \to \infty} \frac{Var\{N(0,t]\}}{t^{2\epsilon}} = \infty, \qquad \text{for any } \epsilon \leq 0.$$

So, the Hurst index for this point process must be positive. Thus, $0 < H \leq 1$, which is the same range that we had for self-similar processes, save the possible inclusion of one. Furthermore, according to Definition 3.2, if a point process is LRcD, then $H \geq 0.5$. In practice, however, most non-LRcD stochastic point processes have $H = 0.5$, while LRcD point processes have $H > 0.5$.

The moment index of the marginal interval distribution and the Hurst index for a point process are related in the following manner:

Lemma 3.1. *Let $N(\cdot)$ be a stationary, ergodic, orderly point process with finite variance (i.e. $E\{[N(A)]^2\} < \infty$ for bounded A), and let F be the marginal distribution of its stationary sequence of intervals. If N is LRcD with Hurst index $0.5 < H < 1$ and κ is the moment index of F, then $2H + \kappa \geq 3$, with equality if N is a renewal point process.*

Proof. See Daley (1999) and Daley et al. (2000). □

Furthermore, the moment index of the interpoint intervals is sufficient in some instances to show that a point process is LRcD, as described in the following lemma.

Lemma 3.2. *Let $N(\cdot)$, a point process, and κ, its interval moment index, be defined as in Lemma 3.1. Then:*

(i) *If $\kappa < 2$, then N is LRcD.*

(ii) *If N is a renewal point process that is LRcD, then $\kappa < 2$.*

Proof. For (i), see Daley et al. (2000). For (ii), see Daley (1999) and Theorem 2.1. □

These results have been extended somewhat by Kulik and Szekli (2001). Their results are based on stochastic orderings of point processes, which is beyond the scope of this study. Essentially, they showed that if the correlation between the interpoint intervals of a point process $N(\cdot)$ is increased, then the variability of N (i.e. the corresponding counting process) is increased as well. Thus, combining this with (ii) in Lemma 3.2, they showed that a finite-intensity point process is LRcD if its interpoint intervals have infinite variance and are positively dependent.

Finally, to complete our representation of LRD in point processes, we need an index of the dependence between the interpoint intervals. *The* Hurst index H of a point process was defined on the basis of the counts of the process, but we may also define the Hurst index of the intervals of the process.

Definition 3.6. Let $N(\cdot)$ be a stationary, ergodic, orderly point process with finite second moment, and let $\{Y_i\}$ be the stationary sequence of its interpoint intervals, also with finite second moment. Then the interval Hurst index of this point process is

$$H_I = \sup\left\{h : \limsup_{n \to \infty} \frac{Var\left\{\sum_{i=1}^n Y_i\right\}}{n^{2h}} = \infty\right\}.$$

Since the sequence of intervals is stationary,

$$Var\Big\{\sum_{i=1}^{n} Y_i\Big\} = \sum_{i=1}^{n} Var\left\{Y_i\right\} + 2\sum_{j=2}^{n}\sum_{i=1}^{j-1} Cov\left\{Y_i, Y_j\right\}$$

$$= nVar\left\{Y_1\right\} + 2\sum_{j=2}^{n}\sum_{i=1}^{j-1} Cov\left\{Y_i, Y_j\right\}$$

$$\leq nVar\left\{Y_1\right\} + 2\sum_{j=2}^{n}\sum_{i=1}^{j-1} Var\left\{Y_1\right\}$$

$$\leq nVar\left\{Y_1\right\} + 2n^2 Var\left\{Y_1\right\}.$$

Thus,

$$\limsup_{n\to\infty} \frac{Var\Big\{\sum_{i=1}^{n} Y_i\Big\}}{n^2} < \infty,$$

and so the interval Hurst index can be no greater than one. Furthermore, if the intervals of the point process are not all identical, then

$$\limsup_{n\to\infty} \frac{Var\Big\{\sum_{i=1}^{n} Y_i\Big\}}{n^{2\epsilon}} = \infty, \qquad \text{for any } \epsilon < 0.$$

So, the interval Hurst index can be no smaller than zero. Hence, as was the case for the (count) Hurst index, $0 \leq H_I \leq 1$. Also, according to Definition 3.1, if the point process is LRiD, then $H_I \geq 0.5$. So the Hurst index of a point process describes the amount of LRcD, while the interval Hurst index describes the amount of LRiD.

3.2.4   Relationship between the Different Types of Long-Range Dependence and the Variability of Intervals

The results reviewed in Section 3.2.3 illustrate part of the interrelationship of the variability of interpoint intervals and the two types of LRD in point processes: LRiD and LRcD. A complete theory of this relationship is not currently available, but, by combining general analytical results with results for specific examples of point processes, we can paint a rough picture of the interplay between LRiD, infinite-variance intervals, and LRcD.

The LRcD renewal point process serves as an example of an LRcD point process which does not exhibit LRiD, but has interpoint intervals with infinite variance. Conversely, Daley et al. (2000) described and analyzed a Wold process that is LRcD and has interpoint intervals with finite variance, but does exhibit LRiD. Thus, LRcD can exist when the interpoint intervals either have infinite variance or are LRD, and, presumably, cannot exist in a point process without the presence of at least one of these interval properties.

Reasonably broad conditions under which the infinite variance of the interpoint intervals produces LRcD are known, although the production of LRcD by LRiD has not been studied as carefully. Kulik and Szekli (2001) have shown that infinite variance of the interpoint intervals produces LRcD in a stationary point process with finite intensity if the interpoint intervals are also positively dependent. On the other hand, Daley et al. (2000)

have shown that if the infinite variance condition is strengthened somewhat that a point process is LRcD regardless of the structure of the interpoint interval dependence. They proved that any stationary point process for which the moment index of the interpoint intervals satisfies $\kappa < 2$, i.e. some moment of order less than two is infinite, is LRcD.

Given the previous discussion, it is reasonable to conjecture that a point process that is LRiD *and* has infinite-variance interpoint intervals should also be LRcD. Although this may be true, it has not been explicitly proven. However, from the previously described results of Kulik and Szekli (2001) and Daley et al. (2000), clear examples of point processes that are LRiD, have infinite interval-variance, and are LRcD are readily obtained. For instance, the dependence between the interpoint intervals of a point process may be both long-range and positive at *all* lags. Thus, according to the results of Kulik and Szekli (2001), such a process would be LRcD if its intervals had infinite variance. On the other hand, according to Daley et al. (2000), a point process with interpoint intervals having infinite variance and being LRiD is certainly LRcD if there is also some infinite moment of the interpoint intervals that is less than order two. Finally, Resnick and Samorodnitsky (1997) have described a model with random variables having an exponential marginal distribution and a "very flexible correlation function". If the sequence of random variables in this model is made to be LRD and is used as the interpoint intervals of a point process, then the resultant point process is, by definition, LRiD. Kulik and Szekli (2001) suggested that if the marginal distribution of these random variables were made to have infinite variance using the inverse transform method, then a point process that both exhibits LRiD and has infinite-variance interpoint intervals would result.

3.2.5 Long-Range Interval Dependence and Infinitely-Variable Intervals

We have so far only discussed the definition of LRD for processes with finite variance. Likewise, the definition of LRiD (Definition 3.1) assumes that the intervals of the point process have finite variance. The lack of a general definition of LRD for infinite-variance processes, and the consequent lack of a definition of LRiD for point processes with infinite-variance intervals, is a significant hindrance to the formation of a complete theory of LRD in point processes. However, the problem with these definitions is that they are based upon the variance. Furthermore, for an infinite-variance process, none of the second-order statistics, the statistics upon which all common definitions of LRD are based, exist. Thus, other conceptions of LRD must be developed that apply when these statistics do not exist. First, however, in order to properly handle infinite-variance processes, a theory of infinite-variance distributions is necessary. For our purposes, the theory of stable distributions will suffice.

The stable distributions are the only limiting distributions for normalized sums of independent, identically distributed random variables. Hence, the central limit theorem implies that the Gaussian (or "normal") distribution is a stable distribution, but this is the only stable distribution with finite variance. On the other hand, there are numerous stable distributions with infinite variance. Often stable distributions, or stable random variables, are said to be α-stable. This derives from the use of the variable α to designate

a key parameter of the stable distributions, the index of stability. For the Gaussian distribution, $\alpha = 2$. But for all other stable distributions, $0 < \alpha < 2$, and the probability tails, $\Pr\{X > x\}$ and $\Pr\{X < -x\}$, of these distributions decay as power functions of the form $|x|^{-\alpha}$. The index of stability is also related to the moments of α-stable random variables. If X is an α-stable random variable with $0 < \alpha < 2$, then (Samorodnitsky & Taqqu, 1994, Property 1.2.16)

$$E|X|^p < \infty \text{ if } 0 < p < \alpha, \text{ and } E|X|^p = \infty \text{ if } p \geq \alpha.$$

One way to define LRD for infinite-variance processes is by analogy. For instance, the fractional Gaussian noises (fGns) are the increment processes of the fractional Brownian motions (fBms), the only Gaussian self-similar processes with stationary increments. As discussed above for stationary increments of self-similar processes in general, each fGn has an associated Hurst index, H, that is the self-similarity parameter of the corresponding fBm. This Hurst index is a measure of the strength of the dependence in fGn, and fGn has LRD when $H > 1/2$ and negative dependence when $H < 1/2$. Now, several different types of α-stable random processes, i.e. those with α-stable marginal distributions, may be created by extending different representations of fBm (see, e.g., ch. 7 of Samorodnitsky & Taqqu, 1994). By analogy, the increments of these α-stable marginal distributions are said to have LRD when $H > 1/\alpha$ and negative dependence when $H < 1/\alpha$ (Samorodnitsky & Taqqu, 1994, pp. 345, 366f, 382f). This, of course, is consistent with fGn, since $\alpha = 2$ for the Gaussian distribution.

Although defining LRD by analogy is profitable for certain infinite-variance processes that are derived from finite-variance processes, this method is obviously limited and cumbersome since a generally applicable definition cannot be produced. Furthermore, it is not useful for data analysis unless the source of the data is known to be well-modeled by a process that is analogous (or identical) to a finite-variance process. This method can, however, lead to some general intuitions about LRD in infinite-variance process. For instance, the discussion of α-stable random processes derived from fBm suggests that the boundary between LRD and negative dependence may be $1/\alpha$, which is never smaller than $1/2$, for all α-stable random processes. In addition, since this boundary is $1/2$ for all finite-variance processes, it also seems likely that any process within the domain of attraction of an α-stable random process has a boundary of $H = 1/\alpha$ between LRD and negative dependence. Since the Hurst index, H, must lie in the interval $(0, 1)$, we have the further implication that an α-stable random process with $0 < \alpha \leq 1$, and potentially any random process in the domain of attraction of this α-stable random process, cannot be LRD.

Another strategy for defining LRD in infinite-variance processes is by making use of a measure that generalizes the covariance, in the sense that it reduces to the covariance for finite-variance processes. Two such measures in the case of stable random variables are the covariation and the codifference (Samorodnitsky & Taqqu, 1994, Ch. 2), which both reduce to the covariance in the case of Gaussian-distributed random variables. The usefulness of the covariation, however, is limited, since it is not defined for α-stable random variables with $0 < \alpha \leq 1$, is not (in general) symmetric in its arguments, and is not (in general) additive in its second argument. Thus, the codifference seems more

promising, and, for instance, Kokoszka and Taqqu (1993, 1994, 1995, 1996) have shown that the summability of the codifference is useful for distinguishing between short-range and long-range dependence in certain autoregressive-moving average processes with stable innovations. But both of these measures have only been defined and analyzed for stable random variables, so their usefulness for other random variables is open to question.

The final method for distinguishing LRD in infinite-variance processes that we will mention makes use of the sample correlation. Although the correlation is not defined for infinite-variance random variables, the sample correlation can obviously be calculated for a sample from an infinite-variance process. Assuming that the sample correlation retains pertinent properties when applied to an infinite-variance process, its asymptotic properties should be useful for distinguishing short-range and long-range dependence. The study of Davis and Resnick (1986) supports this argument. They found that in the case of the moving average process

$$X_n = \sum_{j=-\infty}^{\infty} c_j Z_{n-j},$$

where $\{Z_i\}$ is a sequence of independent and identically distributed random variables called the innovations, the sample correlation converges to the same function of the coefficients $\{c_j\}$ whether the innovations $\{Z_i\}$ have finite or infinite variance.

For the types of point processes that are useful for modeling neural data, the variance of the counts is finite. Thus, in this case, infinite-variance problems are not encountered with respect to LRcD. However, it is often necessary to consider point processes that have interpoint intervals with infinite variance, which subverts our definition of LRiD. In this chapter, we will use, essentially, the last of the three strategies above for handling infinite-variance processes when we consider LRiD. Thus, we will assume that the sample statistics will asymptotically behave in a similar manner whether the intervals have infinite variance or not. Furthermore, we will see that results based upon this assumption are coherent.

3.3 Statistical Procedures for Recognizing Long-Range Dependence in Point Processes

3.3.1 Shuffled Surrogate Data

Shuffled surrogate data can prove useful in determining the relative contributions of infinite-variance intervals and LRiD to LRcD in a point process. Teich, Lowen, and their coworkers (Teich et al., 1990; Lowen & Teich, 1992; Teich & Lowen, 1994; Turcott et al., 1995; Lowen & Teich, 1996b; Teich et al., 1996; Turcott & Teich, 1996; Teich et al., 1997) have made extensive use of this method to "distinguish those properties of the data that arise from correlation among intervals from those properties inherent in the form of the [distribution of interval lengths]" (Turcott & Teich, 1996). The recent work of Daley and his coworkers (Daley & Vesilo, 1997; Daley, 1999; Daley et al., 2000), and Kulik and Szekli (2001), as well as that in Chapter 2, however, offer a more precise understanding of the information that is available from the shuffled surrogate data.

A set of shuffled surrogate data is formed by shuffling the interpoint intervals of a
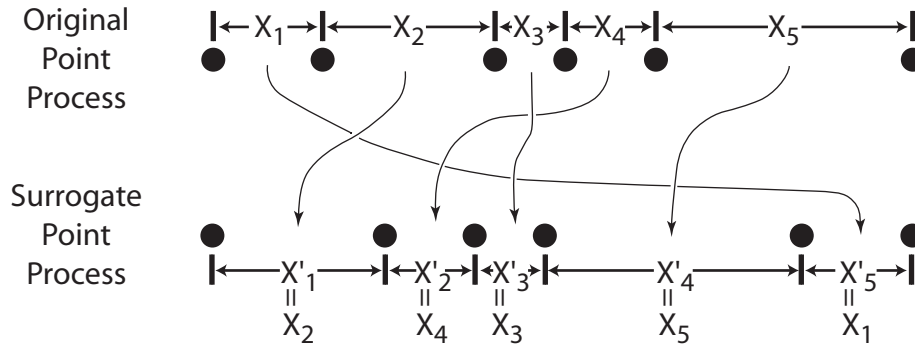
Figure 3.1: The shuffling procedure used to produce surrogate data.

finite sample from a stochastic point process or of related empirical data. This process
is depicted in Figure 3.1 for a short sample from a point process. In general, if the
sequence of interpoint interval lengths in the original data set is denoted by the sequence
$X_1, X_2, \ldots, X_n$, then the sequence of interpoint interval lengths in the surrogate data set
is a random permutation of this sequence $X'_1, X'_2, \ldots, X'_n$ such that the *sets* $\{X_i : i =
1, 2, \ldots, n\}$ and $\{X'_i : i = 1, 2, \ldots, n\}$ are identical. The surrogate point process formed
from the shuffled intervals has the same distribution of interpoint intervals as the original,
but the dependency structure of the interpoint intervals has been disrupted. Thus, the
surrogate point process is essentially equivalent to a sample from a renewal point process
that has an interpoint interval distribution equivalent to the marginal distribution of
the intervals in the original point process. Therefore, the surrogate point process cannot
be LRiD, and any LRcD present is due to the high variability of the interspike interval
distribution (Daley, 1999; Theorem 2.1). Consequently, we expect that for a point process
with no LRiD the "amount" of LRcD in the surrogate data would be equivalent to
that in the original data, since no LRcD would be destroyed by the shuffling procedure.
Furthermore, for a point process with finite-variance interpoint intervals, we expect that
the surrogate data would have no LRcD, since only infinite-variance intervals can create
LRcD in the shuffled data. The results of these arguments are presented in Table 3.1
as the four possible combinations of finite- or infinite-variance intervals and LRiD or no
LRiD.

3.3.2   Statistical Functions for the Variance of Aggregations

In order to investigate the presence of LRD in point processes, we need to examine the
behavior of the variance of aggregations of random variables as the aggregation size
increases. The variance of a time series for different amounts of aggregation can be plotted
in numerous ways. Consider a discrete, stochastic time series $X_1, X_2, X_3, \ldots$, and let

$$X^{(M)} = \sum_{i=1}^{M} X_i$$

be the aggregation of M adjacent random variables in the time series. One could simply
plot the variances of the aggregations, $Var\left\{X^{(M)}\right\}$, versus the amount of aggregation, M.

Variance of Intervals	LRiD in Original Data?	LRcD in Original Data?	LRcD in Surrogate Data?	"Amounts" of LRcD
Finite	No	No	No	Original = Surrogate
Finite	Yes	Yes	No	Original > Surrogate
Infinite	No	Yes	Yes	Original = Surrogate
Infinite	Yes	Yes	Yes	Original > Surrogate

Table 3.1: The four different scenarios for the presence or absence of long-range dependence in a point process and their effect on shuffled surrogate data. The four possibilities are determined by whether the interpoint intervals are finite and whether the point process is LRiD.

This would yield a curve of non-negative slope, where the slope on a double logarithmic plot would equal one if the random variables are independent. A more common method, often called the variance-time curve, is to plot $Var\{\bar{X}^{(M)}\}$ versus M, where

$$\bar{X}^{(M)} = \frac{1}{M}\sum_{i=1}^{M} X_i$$

is the mean value of the aggregated variables. Since

$$Var\{\bar{X}^{(M)}\} = \frac{1}{M^2}Var\left\{X^{(M)}\right\},$$

the variance-time curve has a non-positive slope, with a double logarithmic slope of negative one implying that the random variables are independent. For our variance of counts and variance of aggregated intervals, we have chosen a third method, which is analogous to plotting

$$\frac{1}{M}Var\left\{X^{(M)}\right\} \tag{3.3}$$

versus M for the time series $X_1, X_2, X_3, \ldots$. This results in a curve that will have a slope of zero if there is no dependence between the random variables, a positive slope if there is positive dependence, and a negative slope if there is negative dependence. This facilitates simple qualitative interpretation of these graphs. Furthermore, (3.3) yields expressions that differ by only a constant from the limit arguments in the definitions of LRcD and LRiD. We will make this statement more explicit as we introduce the precise functions that we have chosen to graph.

3.3.3 The Fano Factor Curve

By and large, Teich, Lowen, and their coworkers (see references given above) have used the Fano factor curve (FFC) to investigate the differences between original neural recordings and shuffled surrogate data. The Fano factor (Fano, 1947; Teich, 1989; Lowen & Teich, 1991; Thurner et al., 1997), which is also known as the index of dispersion of counts (Cox & Isham, 1980, pp. 12, 32), of a stationary point process for a given counting interval

length t is the variance of $N(0,t]$ divided by the mean of $N(0,t]$, i.e. ratio of the count variance to the count mean. The FFC is the graph of the function $\mathcal{F}(t)$ that maps each value of t to the value of the Fano factor for a counting interval length of t.

For a Poisson process, the Fano factor is equal to one for all values of t. Hence, the FFC of a Poisson process is a horizontal line at one. Furthermore, for any stationary, orderly point process, (e.g. Cox & Isham, 1980, pp. 32f),

$$E\{N(0,\delta)\} = \lambda\delta \qquad \text{and} \qquad Var\{N(0,\delta)\} = \lambda\delta + o(\delta),$$

where $\lambda = E\{N(0,1)\}$ is the rate of the process. Therefore, as the counting interval length decreases to zero, the Fano factor approaches one, i.e.

$$\lim_{\delta\downarrow 0} \mathcal{F}(\delta) = \lim_{\delta\downarrow 0} \frac{Var\{N(0,\delta)\}}{E\{N(0,\delta)\}} = \lim_{\delta\downarrow 0} 1 + \frac{o(\delta)}{\lambda\delta} = 1. \tag{3.4}$$

Decreases in the FFC are caused either by marginal interval distributions with variance that is lower than that of the exponential distribution (the interval distribution for a Poisson process) with an equivalent mean or by negative correlation between intervals. On the other hand, increases in the FFC are caused either by interval distributions with higher variance than that of the exponential distribution or positive correlation between intervals. Moreover, these increases or decreases in the FFC of the original data will only be present in the FFC of surrogate data if they were created by the shape of the interval distribution; deviations of the original-data FFC from a horizontal line that are due to correlations in the sequence of intervals will disappear in the surrogate-data FFC.

In particular, the FFC of an LRcD point process will, as the counting interval length increases, increase without bound. This fact is evident in the similarity between the Fano factor and the limit argument in the definition for LRcD (Definition 3.2). More explicitly, for a stationary, orderly point process,

$$\mathcal{F}(t) = \frac{Var\{N(0,t]\}}{E\{N(0,t]\}} = \frac{Var\{N(0,t]\}}{t\,E\{N(0,1]\}} = \left(\frac{Var\{N(0,t]\}}{t}\right)\left(\frac{1}{E\{N(0,1]\}}\right), \tag{3.5}$$

where the first expression is the limit argument in the definition of LRcD, and the second is a constant.

The top row of Figure 3.2 shows some illustrative pairs of Fano factor curves that exemplify the basic types of differences that commonly occur between the FFCs of original and surrogate data in neural recordings and models that are LRD. Figure 3.2a depicts the FFCs for samples from an LRcD renewal point process. In this case, the interpoint intervals are completely independent, and LRcD is present only if the interval distribution has infinite variance (Daley, 1999; Theorem 2.1). Thus, shuffling the intervals does not change the statistical properties of the point process, and the FFCs for the original and surrogate data are nearly the same. The outputs of the models dealt with in Chapter 2 would produce this effect in their FFCs.

Figure 3.2b illustrates representative FFCs for point processes with LRiD and finite-variance interpoint intervals. In this case, although the FFC for the original data increases without bound, the FFCs for surrogate data approach a constant limiting value. If the standard deviation of the intervals is equal to their mean, as in the case of exponentially
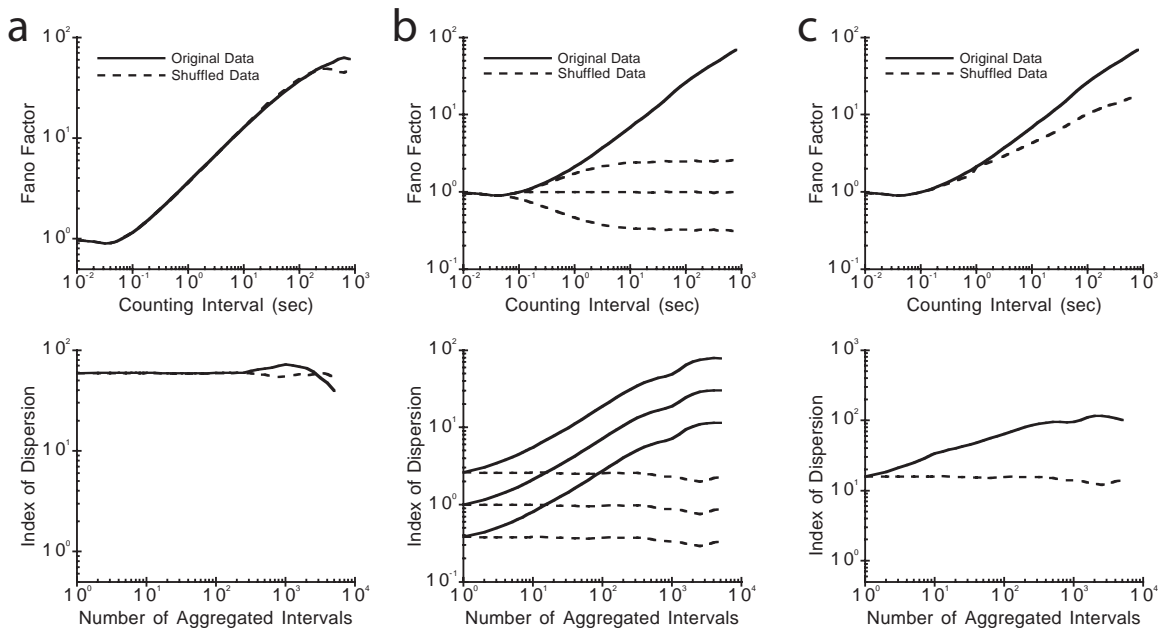
Figure 3.2: Fano factor curves and curves of the index of dispersion of intervals for finite-length samples from three point process types related to neural spike trains and models. Shuffled surrogate data (dashed lines) is formed by randomly shuffling the interpoint intervals of the original sample data from the point process. (a) The curves for a renewal process with an interval distribution possessing infinite variance. (b) The curves for a point process with intervals that are long-range dependent and have a distribution with finite variance. The three Fano factor curves for shuffled surrogate data and three sets of index-of-dispersion curves are from processes with interval variability greater than a Poisson process (top), equal to a Poisson process (middle), and less than a Poisson process (bottom). (c) The curves for a point process with intervals that are long-range dependent and have a distribution with infinite variance.

distributed intervals, then the asymptotic value of the surrogate-data FFC will be one. If the interval variance is larger than in this case, then the surrogate-data FFC will asymptotically approach a value greater than one, and if it is less, then the FFC will approach a value less than one.

Finally, Figure 3.2c depicts the FFCs for an LRcD point process that both is LRiD and has intervals distributed with infinite variance. In this case, the FFC for the shuffled surrogate data increases without bound due to the infinite variance of the interval distribution. However, since the intervals are also LRD, this FFC is not the same as the FFC for the original data. Instead, the surrogate-data FFC has a shallower slope than the original-data FFC due to the loss of LRiD caused by the shuffling procedure.

3.3.4   The Index-of-Dispersion Curve for Intervals

The FFC is the index of dispersion of the counts for different size counting windows, or "aggregations of time", and is closely related to the definition of LRcD. An analogous curve that is closely related to the definition of LRiD may be defined for the intervals of a point process. We will call this curve the index-of-dispersion curve, where it should be understood that we are referring to the index of dispersion *of the interpoint intervals* since the index of dispersion of the counts has another name (i.e. the Fano factor).

The index of dispersion of the intervals of a point process is defined as the variance of the intervals divided by the square of the mean of the intervals (see, e.g., Cox & Isham, 1980, p. 12), i.e.

$$\mathcal{I}_X = \frac{Var\{X\}}{\big(E\{X\}\big)^2},$$

where X is the interval random variable. We may therefore define the index of dispersion for aggregations of k consecutive intervals as

$$\mathcal{I}(k) = \frac{Var\left\{\sum_{i=1}^{k} X_i\right\}}{k\big(E\{X\}\big)^2}, \qquad k \in \mathbb{N}$$

(see, e.g., Cox & Isham, 1980, p. 35). The index-of-dispersion curve (IDC) of the intervals is the graph of the function $\mathcal{I}(k)$.

The denominator of the Fano factor is the mean of the counts, whereas the denominator of the index of dispersion of the intervals is the *square* of the mean interval length. The purpose for this difference is so that, in each case, the statistic enables a straightforward comparison to be made between the point process being analyzed and the Poisson process. The variance of the counts for some specified window size is equal to the mean count for a Poisson process, whereas the *standard deviation* of its interpoint interval lengths is equal to the mean interval length. Thus, the index of dispersion of interval aggregations for a Poisson process is equal to one for all aggregation sizes k, and its IDC is a horizontal line at one. More generally, if the intervals of a point process are independent, then, since in this case

$$Var\left\{\sum_{i=1}^{k} X_i\right\} = k\,Var\{X\},$$

the IDC is a horizontal line at the value of $Var\{X\}/(E\{X\})^2$.

When the intervals of a point process are not independent, the leftmost value of the IDC, i.e. $\mathcal{I}(1)$, is still equal to $Var\{X\}/(E\{X\})^2$. However, the curve will eventually increase or decrease. If intervals at a certain lag are negatively correlated, then the IDC will decrease at the aggregation level equal to this lag. Increases in the IDC, on the other hand, indicate lags at which the intervals are positively correlated. In particular, the IDC of an LRiD point process will, as the aggregation level increases, increase without bound. This can be proven by relating the index of dispersion of aggregated intervals to the limit argument in the definition for LRiD (Definition 3.1). Specifically,

$$\mathcal{I}(k) = \frac{Var\left\{\sum_{i=1}^{k} X_i\right\}}{k\left(E\{X\}\right)^2} = \left(\frac{Var\left\{\sum_{i=1}^{k} X_i\right\}}{k}\right)\left(\frac{1}{\left(E\{X\}\right)^2}\right), \tag{3.6}$$

where the first expression is the limit argument in the definition of LRiD, and the second is a constant. Furthermore, when the intervals of a stationary point process have finite variance, the IDC is asymptotically related to the FFC by the equation (Cox & Isham, 1980, p. 36)

$$\lim_{k\to\infty} \mathcal{I}(k) = \lim_{t\to\infty} \mathcal{F}(t). \tag{3.7}$$

From this we can prove the following partial complement to Lemma 3.2:

Proposition 3.3. *Let $N(\cdot)$ be a stationary, ergodic, orderly point process with finite variance (i.e. $E\{[N(A)]^2\} < \infty$ for bounded A) and a stationary sequence of intervals $\{X_i : i \in \mathbb{N}\}$ that has finite variance. Then N is LRcD if and only if it is LRiD.*

Proof. This result follows directly from (3.7), (3.6), (3.5), and the definitions of LRcD and LRiD. □

The bottom row of Figure 3.2 shows some illustrative pairs of IDCs that correspond to the FFCs above them. For the renewal process of Figure 3.2a, the IDCs for the original and shuffled data are practically identical, as were the FFCs. Furthermore, in accordance with the preceding discussion, the IDCs are essentially horizontal lines, owing to the independence of the intervals in both sets of data. Theoretically, the values of the index of dispersion of aggregated intervals should be infinite, since the variance of the interpoint interval distribution is infinite. But since they are derived from estimated values from a limited sample of the process, the IDCs are, of course, at finite values. These estimated values, however, are high, being approximately equivalent to the highest values of the corresponding FFCs. If the length of the sample from the renewal process is gradually increased, then the trend should be for these calculated IDCs to move upwards without bound.

In Figure 3.2b, the LRiD of the point process is evident in the ever-increasing original-data IDCs. The IDCs for the shuffled surrogate data, on the other hand, are horizontal lines, since the shuffling procedure destroys the serial dependence between the intervals and the intervals have finite variance. The three sets of IDCs (top, middle, and bottom) correspond to the three surrogate-data FFCs (top, middle, and bottom, respectively). So if the standard deviation of the intervals is equal to their mean, as in the case of

exponentially distributed intervals, then the IDCs at one "aggregated" interval are equal to one. If the interval variance is larger than the mean, then the IDCs have a value greater than one at one interval, and if it is less than the mean, then the IDCs have a value less than one at one interval.

Lastly, the IDCs in Figure 3.2c do not differ significantly in shape from those in Figure 3.2b, although the point processes are statistically different in these two cases. What can be seen from these IDCs is that the point process is LRiD, as evident from the ever-increasing IDC for the original data. Furthermore, whereas the slope of the original-data IDCs in Figure 3.2b are approximately the same as those in the corresponding FFCs, the slope of the original-data IDC in Figure 3.2c is shallower than the original-data FFC. In terms of the Hurst indices, this means that in Figure 3.2b $H_I \approx H$, while in Figure 3.2c $H_I < H$. This implies that only some of the LRcD is due to LRiD. The rest, of course, is due to the infinite variance of the intervals. However, as in Figure 3.2a, the infinite variance of the interval distribution is not evident in the IDCs alone due to the finite sample time. But, as the length of the sample increases, the vertical position of these curves should, on average, move upwards without bound.

3.3.5   Analysis of Long-Range Dependence in Point Process Data

In the previous sections we have defined a statistical procedure, the FFC, for recognizing LRD in the counts of point process data, and another, the IDC, for recognizing LRD in the intervals of point process data. Together these two statistical curves can often detect the presence of infinite variance in the interpoint intervals as well. Furthermore, by comparing these statistical curves for an original set of data with those for a set of data obtained by randomly shuffling the original interpoint intervals, a more robust and sensitive indicator of the presence of LRcD, LRiD, and infinite interval-variance is produced. Thus, using the original-data FFC and IDC and the surrogate-data FFC and IDC in combination is a good strategy for discerning among the four potential scenarios of Table 3.1 for LRD in point process data. In addition, this process provides information on the strength of LRcD in the data and the relative contributions of LRiD and infinite-interval variance to its presence.

3.4   Long-Range Dependence in Neural Spike Trains

Statistical properties related to LRD have been found in the spike trains of many different neural systems. These properties include second-order self-similarity and power-law (second-order) statistics, and are often studied within the context of fractal theory. Section 2.3 contains further description of these statistical properties and their relationship to LRD. In the following, we will not discriminate between the precise terminology used by other researchers but will refer to all of these properties as LRD. LRD has been investigated primarily in the sub-cortical and cortical sensory systems of mammals, although it has also been found in the visual system of certain insects (Turcott et al., 1995). Furthermore, almost all of the systems where LRD has been investigated have

provided evidence of its existence. The main exception is the peripheral vestibular system (Teich, 1989). This difference should not be surprising, however, since the pattern of firing in these neurons is clearly qualitatively different than the firing in, for instance, visual and auditory neurons. Whereas the spike trains of the latter seem to be very irregular, the former tend to be quite periodic.

3.4.1 Long-Range Dependence in Sub-Cortical Neurons

Long-range dependence has been found at several different sub-cortical levels of the mammalian visual and auditory systems. Primarily, the studies relevant to LRD in the visual and auditory systems have been undertaken by Teich, Lowen, and their coworkers. They have found LRD in the primary auditory nerve (Teich, 1989; Teich & Lowen, 1994; Lowen & Teich, 1996b), the lateral superior olive (Teich et al., 1990; Turcott et al., 1994), and the retina and lateral geniculate nucleus (Teich, 1996; Teich et al., 1997; Lowen et al., 2001). In all of these sub-cortical auditory and visual neurons, though LRD is evident in the power-law growth of the FFC, the FFC for shuffled surrogate data asymptotes to a constant below one. Thus, the LRcD in these systems is a result of LRiD, and the intervals have sub-Poissonian variability.

In addition to the lateral geniculate nucleus, LRD has also been found in ventrobasal neurons of the thalamus (Kodama, Mushiake, Shima, Nakahama, & Yamamoto, 1989). However, the study of these neurons did not produce data that can be used to draw conclusions about the presence of LRiD and the variability of the interspike intervals.

3.4.2 Long-Range Dependence in Cortical Neurons

In the cortex, LRD has been found in somatosensory cortex (Wise, 1981), the mesencephalic reticular formation (Yamamoto, Nakahama, Shima, Kodama, & Mushiake, 1986; Gruneis et al., 1989, 1990, 1993), and the hippocampus (Mushiake, Kodama, Shima, Yamamoto, & Nakahama, 1988; Kodama et al., 1989), but these studies did not produce data that can be used to distinguish high interval variability and LRiD. The only helpful study in this regard is that of Teich et al. (1996), who found LRD in primary visual cortex. In contrast to the results in sub-cortical systems, in the cortex, the FFCs for shuffled surrogate data asymptote to a value larger than one. Thus, although the FFCs for the original spike trains from both sub-cortical and cortical neurons are similar to that in Figure 3.2b, the FFCs for shuffled surrogate data are different between these two sections of the nervous system. The surrogate-data FFCs for sub-cortical neurons are similar to the bottom dashed curve in Figure 3.2b, whereas those for cortical neurons are similar to the top dashed curve in this figure.

The results for cortical neurons signify that, as in sub-cortical neurons, the LRcD in the spike trains is due to LRiD. However, unlike the interspike intervals in sub-cortical neurons, those in cortical neurons are more highly variable than the intervals of a Poisson process, but this variability is still finite. Thus, the data of Teich et al. (1996) lends further support to the conclusion of Chapter 2 that high-variability IF models with renewal outputs, which would result in curves like those in Figure 3.2a, do not match the statistical properties of real cortical neurons.

3.5 Integrate-and-Fire Models with Renewal Point Process Inputs

3.5.1 The Integrate-and-Fire Models of Feng and His Coworkers

The arguments in Chapter 2 showed that any model of cortical neurons that produces a renewal output cannot exhibit both CV_{ISI} values and LRcD properties that are similar to those seen in real cortical neurons. The (renewal) output of such models either is LRcD and has infinite CV_{ISI} or is not LRcD and has finite CV_{ISI}.

However, there are a number of high-variability models that produce output spike trains that are not renewal point processes (RPPs). One type of non-renewal, high-variability model includes the models analyzed by Feng and his coworkers (Feng, 1997; Feng & Brown, 1998a, 1998b; Feng et al., 1998; Feng, 1999). These models are identical to the class of models considered in Chapter 2, except that the inputs are no longer Poisson point processes. Instead, the inputs can be any other type of RPP, and hence the cumulative inputs consisting of the superposition of all excitatory inputs and the superposition of all inhibitory inputs are no longer memoryless.

In particular, Feng and his coworkers considered RPP inputs with positive Gaussian and Pareto interval distributions. The former has a tail that decreases to zero faster than the exponential distribution (i.e. that of a Poisson process), and the latter has a tail that decreases to zero slower than the exponential distribution. In general, they found that both longer tailed input distributions and more closely balanced amounts of excitation and inhibition increased the CV_{ISI} of the output. This suggests that the CV_{ISI} of an IF model with RPP inputs can be within the physiologically realistic range, regardless of the interval distributions of the RPPs, if the ratio between the amounts of excitation and inhibition is properly adjusted. The necessary range of inhibition-excitation ratios, however, does depend on the interval distribution of the inputs. In particular, for long-tailed input distributions, less inhibition is required to produce realistic CV_{ISI} values, whereas for short-tailed input distributions, the amount of inhibition needs to be much closer to the amount of excitation (Feng & Brown, 1998a).

The IF model with Poisson inputs raises enough analytical difficulties that we do not expect that complete analytical results can be obtained for the case of general RPP inputs. However, a few general observations can be made. First, due to the integration mechanism, the likelihood of the occurrence of an output spike immediately following another output spike is low, but will increase as the time since the last output spike increases. Thus, at small counting windows, we expect the counts to be negatively correlated. This will not, however, affect the correlation structure of the intervals, since the IF mechanism completely resets at the occurrence of each output spike. Therefore, the IF mechanism cannot, by itself, create memory in the model that is longer than the interspike intervals of the output. Hence, the medium- and long-term memory properties of the output must be governed by the inputs and the mechanisms by which they are combined.

The combination of the inputs may be considered as two separate component processes: superposition and excitation-inhibition interaction. For Poisson inputs, the

excitation-inhibition interaction can effect long-term memory by producing high interval variability (Tuckwell, 1988; Section 2.4), but with concomitant increases in the mean interval length. Balancing the amounts of excitation and inhibition will presumably have a similar type of effect when the inputs are RPPs. Any additional memory properties of the output, in particular those that are longer than the interspike intervals, must therefore originate in the superpositions of the input point processes. Thus, we should be able to gain some further intuition about the memory properties of the IF model with RPP inputs by considering the dependency structures of superpositions of RPPs.

3.5.2 Analytical Results for the Superposition of Renewal Point Processes

The Superposition of Generic Renewal Point Processes

Before focusing successively on the superposition of RPPs with positive Gaussian and Pareto distributions, we collect here, or develop when necessary, some useful general results. It is well known that as the number of component RPPs increases, under proper normalization, their superposition approaches a Poisson process (Cox & Smith, 1954; Khintchine, 1960, Chapter 5; Cox, 1967, Section 6.6). Thus, as the number of inputs to the IF model increases, the model becomes more similar to that considered in Chapter 2, where the output was LRD if and only if the variance of the intervals of the output was infinite. Clearly, then, as the number of RPP inputs increases, the model will become less and less likely to possess both a finite CV_{ISI} and LRD.

The forms of the interpoint interval distribution and the serial dependence between the intervals in the superposition, for any fixed, finite number of inputs, will indicate the direction from which it approaches the Poisson process as more component processes are added. Let $G(t)$ be the marginal (cumulative) distribution function of the intervals of the superposition of p independent RPPs, each with intervals distributed according to $F(t)$ with a mean of μ. Then, the distribution functions of the components and the superposition are related by (Cox & Smith, 1954; Lawrance, 1973)

$$1 - G(t) = \left(1 - F(t)\right) \left[\int_t^\infty \frac{1 - F(s)}{\mu} \, ds\right]^{p-1}. \tag{3.8}$$

Thus, for any given interval distribution for the component RPPs, the form of the interval distribution of their superposition can be calculated.

Next, we consider the dependency structure of the intervals of the superposition process. Gath (1974) found, in a computer simulation study, that the intervals in the superposition of independent RPPs are negatively correlated. Although typically true for those RPPs that are used to model physical processes, this is not a general result for all RPPs (Enns, 1970; Linebarger & Johnson, 1986).

However, Enns (1970) established some general analytic results regarding the covariance between two adjacent intervals in the superposition of an arbitrary number of RPPs. Let $H(t)$ be the distribution of the forward recurrence time, i.e. the time interval from an arbitrary time to the next event, of each of the component processes, which has the

density

$$h(t) = \frac{1 - F(t)}{\mu}.$$

Enns (1970) showed that the covariance between the lengths of two adjacent intervals, τ_n and τ_{n+1}, in the superposition process is

$$\lim_{n \to \infty} Cov(\tau_n, \tau_{n+1}) = \mu^2 I(p), \tag{3.9a}$$

where

$$I(p) = \int_0^\infty \Big(h(t) * h(t)\Big)\Big(1 - H(t)\Big)^{p-1} dt - \frac{1}{p^2}, \tag{3.9b}$$

and the symbol $*$ denotes convolution. If there is only one component process, then $I(1) = 0$, which is consistent with the superposition process being an RPP. But, in addition, $\lim_{p \to \infty} I(p) = 0$. This means that the superposition process approaches an RPP as the number of components is increased, in agreement with the aforementioned result that the superposition approaches a Poisson process in this case.

When the hazard rate of the component RPPs is monotone, then the sign of its derivative can be used to determine the sign of $I(p)$, and thus, by (3.9a), of the covariance of adjacent intervals in the superposition process. The hazard rate of an RPP having an interval distribution $F(t)$, with probability density function $f(t) = \frac{d}{dt}F(t)$, is

$$z(t) = \frac{f(t)}{1 - F(t)}.$$

The hazard rate can be interpreted as

$$z(t) = \lim_{\delta \to 0^+} \frac{\Pr\Big\{N(0, t] = 0 \text{ and } N(t, t + \delta] = 1 \mid N(\{0\}) = 1\Big\}}{\delta},$$

where the numerator is the probability that the next point occurs in the interval $(t, t + \delta]$ given that a point occurred at time zero. When the hazard rate of the inputs is monotone non-decreasing, then $I(p) \leq 0$, and when it is monotone non-increasing, then $I(p) \geq 0$ (Enns, 1970). The exponential distribution, which has a constant hazard rate (i.e. both non-decreasing and non-increasing), is the boundary case, yielding $I(p) \equiv 0$ and uncorrelated intervals in the superposition, which is consistent with the fact that the superposition of a number of Poisson processes is also a Poisson process. Thus, if the hazard rate of the inputs is monotone non-decreasing but also non-constant, then $I(p) < 0$, and adjacent intervals in the superposition process are negatively correlated. On the other hand, if the hazard rate of the inputs is monotone non-increasing and non-constant, then $I(p) > 0$, and adjacent intervals in the superposition process are positively correlated.

We are particularly interested here in the variability and long-term memory properties of the output of the model. Since the integration and threshold mechanisms themselves cannot produce LRD, the only possible sources of LRD are the two superpositions of the inputs, one for the excitatory inputs and one for the inhibitory inputs, and their excitatory-inhibitory interaction.

In order to evaluate the LRD of a superposition of independent RPPs, we can consider the variance function $V(t) \equiv Var\{N(0,t]\}$, which is the variance of the number of events in an interval of length t. For a single RPP with a finite-variance interval distribution, it is known (e.g. Cox & Smith, 1954; Cox, 1967; Feller, 1971) that

$$V(t) \sim \frac{\sigma^2 t}{\mu^3} \quad \text{as} \quad t \to \infty,$$

where μ and σ^2 are, respectively, the mean and variance of the interval distribution. Thus, the variance function of the superposition of p independent, statistically-identical RPPs is

$$V_S(t) \sim \frac{p\sigma^2 t}{\mu^3} \quad \text{as} \quad t \to \infty, \tag{3.10}$$

where μ and σ^2 are the mean and variance of the component processes. Therefore, if the component processes have finite interval variances, and thus are not LRcD, then their superposition has finite variance and is not LRcD.

Actually, the connection between the interval variance of the component processes and the LRcD of their superposition is even stronger than the preceding discussion indicates. Theorem 2.1 states that an RPP is LRcD if and only if its intervals have infinite variance. The proof given there can be readily adapted to obtain a similar result for the superposition of a finite number of independent, statistically-identical RPPs.

Theorem 3.4. *Suppose that $N_1(\cdot), N_2(\cdot), \ldots, N_p(\cdot)$ are p independent stationary RPPs, each with the same distribution function F of their generic interpoint interval random variable X, which has $F(0) = 0$ and finite mean $\mu = E\{X\}$. Then the superposition of these p RPPs, $N_S(t) = \sum_{i=1}^{p} N_i(t)$, is LRcD if and only if $E\{X^2\} = \infty$, i.e. the variance of the intervals in the component processes are infinite.*

Proof. From the proof of Theorem 2.1, we have that for each of the component processes

$$Var\{N_i(0,t]\} = \frac{2}{\mu} \int_0^t \left[U(s) - \frac{s}{\mu} \right] ds - \frac{t}{\mu}, \quad i = 1, 2, \ldots, p,$$

where

$$U(t) = \sum_{j=0}^{\infty} \left(\lim_{h \downarrow 0} \Pr\{N_i(0,t] \geq j | N_i(-h,0] > 0\} \right), \quad i = 1, 2, \ldots, p,$$

is called the expectation function in general, or, in the case of renewal processes, is the renewal function. Thus, since the processes are independent, the variance function of their superposition is

$$Var\{N_S(0,t]\} = \sum_{i=1}^{p} Var\{N_i(0,t]\} = \frac{2p}{\mu} \int_0^t \left[U(s) - \frac{s}{\mu} \right] ds - \frac{pt}{\mu}.$$

So,

$$\frac{Var\{N_S(0,t]\}}{t} = \frac{2p}{\mu} \left(\frac{1}{t} \int_0^t \left[U(s) - \frac{s}{\mu} \right] ds \right) - \frac{p}{\mu},$$

which goes to infinity as $t \to \infty$ if and only if the integrand goes to infinity. In other words,

$$\lim_{t \to \infty} \frac{Var\{N_S(0,t]\}}{t} = \infty \qquad \text{if and only if} \qquad \lim_{t \to \infty} \left(U(t) - \frac{t}{\mu} \right) = \infty. \qquad (3.11)$$

But, we also have, from the proof of Theorem 2.1, that

$$\lim_{t \to \infty} \left(U(t) - \frac{t}{\mu} \right) = \infty \qquad \text{if and only if} \qquad E\{X^2\} = \infty. \qquad (3.12)$$

Putting (3.11) and (3.12) together yields the desired result. $\qquad \square$

Remark. This theorem can easily be generalized to the case where the component processes have different interval distribution functions. In that case the superposition process is LRcD if and only if one or more of the component processes has intervals with infinite variance. Since this generalization is unnecessary for the results of this chapter, we have chosen the more specialized form above in order to, for the reader's sake, reduce the notational complexity of the proof.

The Superposition of Renewal Point Processes with Positive Gaussian Interval Distributions

We now apply the general theory of Section 3.5.2 to examine the properties of the superposition of inputs in the models of Feng and his coworkers. The RPP inputs of one of their models possessed interpoint intervals distributed according to a positive Gaussian distribution.

The positive Gaussian distribution is an example of a distribution that has a support of $[0, \infty)$ and a tail that is shorter than the exponential distribution. If X is a random variable with a Gaussian, or normal, distribution and a mean of zero, then $Y = |X|$ has a positive Gaussian, or "folded" Gaussian, distribution. The probability density function of the positive Gaussian distribution is

$$f(t) = \begin{cases} \frac{2}{\pi \mu} \exp\left(-\frac{t^2}{\pi \mu^2}\right), & \text{if } t \geq 0; \\ 0, & \text{otherwise,} \end{cases}$$

where $\mu > 0$ is the expected value. Since its tail is shorter than the exponential distribution, which has a variance of μ^2 when its mean is μ, the positive Gaussian distribution must have a variance that is less than μ^2. Specifically, its variance is $\left(\frac{\pi}{2} - 1\right) \mu^2$.

First, we consider the marginal distribution of the intervals in the superposition of p independent RPPs, all with positive Gaussian interval distributions and means of μ. Let $F(t)$ and $G(t)$ be the (cumulative) distribution functions of the intervals of the component and superposition processes, respectively. Then, the tail of the superposition interval

distribution may be calculated using (3.8):

$$1 - G(t) = \text{erfc}\left(\frac{t}{\mu\sqrt{\pi}}\right)\left(\int_t^\infty \frac{\text{erfc}\left(\frac{t}{\mu\sqrt{\pi}}\right)}{\mu}\,dx\right)^{p-1}$$

$$= \text{erfc}\left(\frac{t}{\mu\sqrt{\pi}}\right)\left[e^{-\frac{t^2}{\pi\mu^2}} - \frac{t}{\mu}\text{erfc}\left(\frac{t}{\mu\sqrt{\pi}}\right)\right]^{p-1}, \qquad \text{for } t \geq 0,$$

where

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-s^2}\,ds$$

is the complementary error function. Since (i) the complementary error function is always between zero and one, (ii) the second term in the difference above must be positive, and (iii) p is a positive integer,

$$1 - G(t) \leq e^{-\frac{(p-1)t^2}{\mu^2\sqrt{\pi}}}.$$

Thus, the tail of the interval distribution for the superposition process decreases faster than an exponential function, as is the case for the component processes. In particular, this result implies that the intervals of the superposition process have finite means and variances.

The results of Enns (1970), reviewed in Section 3.5.2, can be used to study the correlation between adjacent intervals in the superposition process. It is shown in Appendix A that the hazard rate of an RPP with positive-Gaussian-distributed intervals is monotone increasing. Hence, the intervals in the superposition of such RPPs are negatively correlated. This is a common result for RPPs that are used to model real phenomena. If the hazard rate is increasing, then, as time passes since the last event, it becomes more likely, per unit time, that the next event will occur. Conversely, if the hazard rate is decreasing, then, as time passes, it becomes less likely, per unit time, that the next event will occur. In the theory of renewal models of component failures, the latter case implies that with use the component becomes more and more reliable, which is typically not true of actual physical components. On the contrary, a component usually degrades with use.

Since any distribution with a tail that is shorter than the exponential distribution has finite variance, Theorem 3.4 asserts that the superposition of RPPs with this kind of interval distribution are not LRD. For component processes with intervals distributed according to a positive Gaussian distribution, (3.10) implies that the asymptotic behavior of the variance function of the superposition process is given by

$$V_S(t) \sim \frac{p[(\frac{\pi}{2} - 1)\mu^2]t}{\mu^3} = \frac{p(\frac{\pi}{2} - 1)t}{\mu} \quad \text{as} \quad t \to \infty.$$

Thus, in accordance with Theorem 3.4, the limit of $V_S(t)/t$, as $t \to \infty$, is a finite constant. More instructive is the asymptotic behavior of the FFC. For the superposition of positive Gaussian RPPs, as the counting interval length increases, the Fano factor approaches

$$\lim_{t \to \infty} \mathcal{F}_S(t) = \lim_{t \to \infty} \frac{V_S(t)}{\left(\frac{pt}{\mu}\right)} = \frac{\pi}{2} - 1 \approx 0.5708 < 1,$$

which is independent of the number of component processes, p, and the parameter μ of their interval distribution. The fact that the Fano factor is less than one is in agreement with the prior finding that (at least) adjacent intervals are negatively correlated in the superposition.

The Superposition of Renewal Point Processes with Pareto Interval Distributions

In addition to the model with positive-Gaussian-distributed inputs, Feng and his coworkers considered another model with RPP inputs that had intervals distributed according to the Pareto distribution. The Pareto distribution is an example of a distribution that has a support of $[0, \infty)$ and a tail that is longer than the exponential distribution. The probability density function of the Pareto distribution is

$$f(t) = \begin{cases} \alpha K^\alpha \, (t + K)^{-\alpha - 1}, & \text{if } t \geq 0; \\ 0, & \text{otherwise,} \end{cases} \tag{3.13}$$

with parameters $K > 0$ and $\alpha > 0$. K is essentially a normalization constant, while the use of α here is analogous to its use in the theory of α-stable random variables. Consequently, the tail probability $\Pr\{X > x\}$ of a Pareto-distributed random variable decays as $x^{-\alpha}$, and only the moments less than α exist. The mean of the Pareto distribution, if $\alpha > 1$, is $K/(\alpha - 1)$, while, if $\alpha > 2$, the variance is $2K^2/[(\alpha - 1)(\alpha - 2)]$.

We again consider the marginal distribution of the intervals in the superposition of p independent RPPs, but now we let the intervals in the component processes be Pareto distributed with parameters K and α. Let $F(t)$ and $G(t)$ be the (cumulative) distribution functions of the intervals of the component and superposition processes, respectively. Then, integrating (3.13) and subtracting the result from one yields

$$1 - F(t) = K^\alpha \, (t + K)^{-\alpha}, \qquad \text{for } t \geq 0. \tag{3.14}$$

Now, (3.8) can be used to calculate the tail of the superposition interval distribution when $\alpha > 1$:

$$\begin{aligned}
1 - G(t) &= \left(K^\alpha \, (t + K)^{-\alpha} \right) \left[\int_t^\infty \frac{K^\alpha \, (s + K)^{-\alpha}}{K/(\alpha - 1)} \, ds \right]^{p-1} \\
&= K^{\alpha + (\alpha - 1)(p-1)} \, (t + K)^{-\alpha} \left[\int_t^\infty (\alpha - 1) \, (s + K)^{-\alpha} \, ds \right]^{p-1} \\
&= K^{(\alpha - 1)p + 1} \, (t + K)^{-\alpha} \left[- (s + K)^{-(\alpha - 1)} \Big|_t^\infty \right]^{p-1} \\
&= K^{(\alpha - 1)p + 1} \, (t + K)^{-\alpha - (\alpha - 1)(p-1)} \\
&= K^{(\alpha - 1)p + 1} \, (t + K)^{-[(\alpha - 1)p + 1]}, \qquad \text{for } t \geq 0. \tag{3.15}
\end{aligned}$$

Thus, by comparing (3.14) and (3.15), we see that the marginal interval distribution of the superposition process is also a Pareto distribution, when $\alpha > 1$, with parameters K and $\alpha' = (\alpha - 1)p + 1$. Thus, the intervals in the superposition have finite variance as long as $(\alpha - 1)p + 1 > 2$, or, equivalently, if $p > 1/(\alpha - 1)$. Thus, as α, the parameter for

the component processes, approaches the value of one, an increasing number of inputs are required if their superposition is to have intervals with finite variance. Furthermore, if the intervals of the component processes have finite variance, i.e. $\alpha > 2$, then the intervals in their superposition do as well.

Even though the marginal interval distribution of the superposition of Pareto RPPs is also Pareto, the superposition process is not an RPP when there are multiple component processes. To show this, we can examine the hazard rate of the Pareto distribution. This hazard rate is straightforward to calculate:

$$z(t) = \frac{f(t)}{1 - F(t)} = \frac{\alpha K^\alpha (t + K)^{-\alpha - 1}}{K^\alpha (t + K)^{-\alpha}} = \frac{\alpha}{t + K}.$$

Hence, the hazard rate of the Pareto distribution is a monotone decreasing function. More precisely, its derivative is

$$\frac{d}{dt} z(t) = -\frac{\alpha}{(t + K)^2},$$

which is negative for all values of t and all allowable (i.e. positive) values of the parameters K and α. Therefore, according to the results of Enns (1970) discussed in Section 3.5.2, adjacent intervals in the superposition of RPPs with Pareto-distributed intervals are positively correlated.

The Pareto distribution only has finite variance if $\alpha > 2$. Thus, according to Theorem 3.4, the superposition of Pareto RPPs will be LRD if and only if $\alpha \leq 2$. For component processes with intervals distributed according to a Pareto distribution and $\alpha > 2$, (3.10) implies that the asymptotic behavior of the variance function of the superposition process is given by

$$V_S(t) \sim \frac{p\left[\frac{2K^2}{(\alpha - 1)(\alpha - 2)}\right] t}{\left[\frac{K}{\alpha - 1)}\right]^3} = \frac{2p(\alpha - 1)^2 t}{K(\alpha - 2)} \quad \text{as} \quad t \to \infty.$$

Thus, as expected, the limit of $V_S(t)/t$, as $t \to \infty$, is a finite constant when $\alpha > 2$. Also, for the superposition of Pareto RPPs, with $\alpha > 2$, the Fano factor approaches

$$\lim_{t \to \infty} \mathcal{F}_S(t) = \lim_{t \to \infty} \frac{V_S(t)}{\left(\frac{pt(\alpha - 1)}{K}\right)} = \frac{2(\alpha - 1)}{\alpha - 2},$$

which, as in the case of positive Gaussian RPPs, is independent of the number of component processes, p. However, in this case the limiting value of the Fano factor is dependent on a parameter, α, of the distribution. However, for all $\alpha > 2$, the limiting value of the Fano factor is greater than one, which is in accordance with the fact that (at least) adjacent intervals in the superposition of Pareto RPPs are positively correlated. In fact, this value is always greater than two for $\alpha > 2$, with it approaching the value of two as $\alpha \to \infty$ and growing without bound as $\alpha \to 2$.

3.5.3 Simulation Results for the IF Model with Renewal Point Process Inputs

The analytical results of Section 3.5.2, by describing the superposition of the inputs, provide some clues to the properties that should be expected of the output of the IF model with RPP inputs. In addition, the arguments in Section 3.5.1 furnish additional insight into characteristics that the output is likely to possess. However, a complete analytical treatment of this model, save when the inputs are Poisson processes, is not available. Therefore, in order to test the ability of IF models with renewal point process inputs to produce realistic CV_{ISI}'s and LRcD, we have run computer simulations of these models.

For each model and set of parameters, we ran ten simulations, each with a duration of 100,000 seconds. Each simulation used 100 excitatory inputs and $100r$ inhibitory inputs, where r was one parameter of the simulations that was varied. Thus, r represents the ratio of the number of inhibitory inputs to the number of excitatory inputs. The IF mechanism had a reset potential of zero and a threshold of unity, and each spike that occurred in either an excitatory or inhibitory input caused the potential to instantaneously increase or decrease, respectively, by $1/40 = 0.025$. When the potential reached the threshold value, an output spike was registered and the potential was reset. The value of the appropriate parameter of the interval distribution of the inputs was chosen to yield a nominal output rate of 2.5 spikes per second according to the following formula:

$$
\begin{aligned}
input\_rate &= \frac{output\_rate}{(\#\_of\_excitatory\_inputs)(postsynaptic\_potential)(1-r)} \\
&= \frac{2.5}{100 \cdot 0.025 \cdot (1-r)},
\end{aligned}
\tag{3.16}
$$

where *postsynaptic_potential* is the amount by which input spikes increased or decreased the potential in the IF mechanism. If possible the interval from time zero to the first spike in each input process was chosen according to the distribution of the forward recurrence time. This ensured that each input was stationary (or at equilibrium). If this was not possible, then a pseudo-stationary state was created by generating a random interval length from the interval distribution and choosing the origin to be a uniformly distributed point within this interval.

We analyzed the output spike trains of these models using estimators of the CV_{ISI}, the FFC, and the IDC. The sample CV_{ISI} is the standard measure of variability that has been used in prior studies to measure the variability in cortical neurons and to compare the variability of these neurons with that of related models. It is equal to the sample standard deviation (i.e. the square root of the sample variance) of the interspike intervals divided by the sample mean of the interspike intervals. The IDC required calculation of the sample variance of the length of, say, M adjacent intervals for many values of M. Since the IDC was plotted on a double logarithmic graph, we began with a value of $M = 1$ and used ten values of M per decade, which were equally spaced in log coordinates. No M-values greater than one-fifth of the total number of interspike intervals in the output of the simulation were used. Similarly, the FFC required calculation of the sample variance of the counts in windows of size T, say, for many values of T. The FFC was also plotted

on a double logarithmic graph, so values of T were chosen that are equally spaced in log coordinates. Again, we used ten values per decade, but the values of T began at 0.01 seconds. No T-values greater than one-fifth of the total simulation length were used.

Finally, the interspike intervals of the output spike train from each of the ten simulations for a given model and set of parameters were randomly shuffled once to produce a set of ten surrogate spike trains. The sample variances of the counts and aggregated intervals were then calculated, in the same manner as before, for these new spike trains in order to produce the surrogate-data FFCs and IDCs.

The IF Model with Poisson Process Inputs

For comparison, we ran simulations of the IF model with Poisson process inputs with the same parameters as our other simulations. For the IF model with Poisson process inputs, it is possible to derive, in closed form, the theoretical value of the CV_{ISI} of its output as a function of the inhibition-excitation ratio, r. According to (2.5), the coefficient of variation of the intervals in the output of this model is given by

$$CV_{ISI} = \sqrt{\frac{M_E\lambda_E + M_I\lambda_I}{\hat{\theta}\left(M_E\lambda_E - M_I\lambda_I\right)}}, \tag{3.17}$$

where M_E and M_I are the number of excitatory and inhibitory inputs, respectively, λ_E and λ_I are the spike rates of the individual excitatory and inhibitory inputs, respectively, and $\hat{\theta}$ is the number of excitatory inputs required to move the "membrane potential" from the resting potential to threshold in the absence of inhibition. Thus, for our simulations,

$$CV_{ISI}(r) = \sqrt{\frac{1+r}{\hat{\theta}\left(1-r\right)}}, \tag{3.18}$$

where

$$\hat{\theta} = \frac{1-0}{1/40} = 40.$$

This function is plotted as a dashed line in each graph of Figure 3.3.

In Figure 3.3a, the values of CV_{ISI} estimated from ten simulations of the Poisson-input model at each value of r are individually plotted. However, since the variability of these estimates across simulations is so low, the ten symbols at each value of r fall almost exactly on top of each other. It is apparent from this graph that the results of our simulations agree well with the theoretical values of CV_{ISI} for this model.

The IF Model with Positive-Gaussian Renewal Inputs

For simulations of the IF model with positive-Gaussian RPP inputs, the sole parameter of the interval distribution, the mean μ, was set to be the inverse of the nominal input rate from the calculation in (3.16). Thus, the only parameter that we varied in these simulations was the inhibition-excitation ratio r. For this model, we ran simulations with r-values of 0.0, 0.5, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 0.97, 0.98, and 0.99.
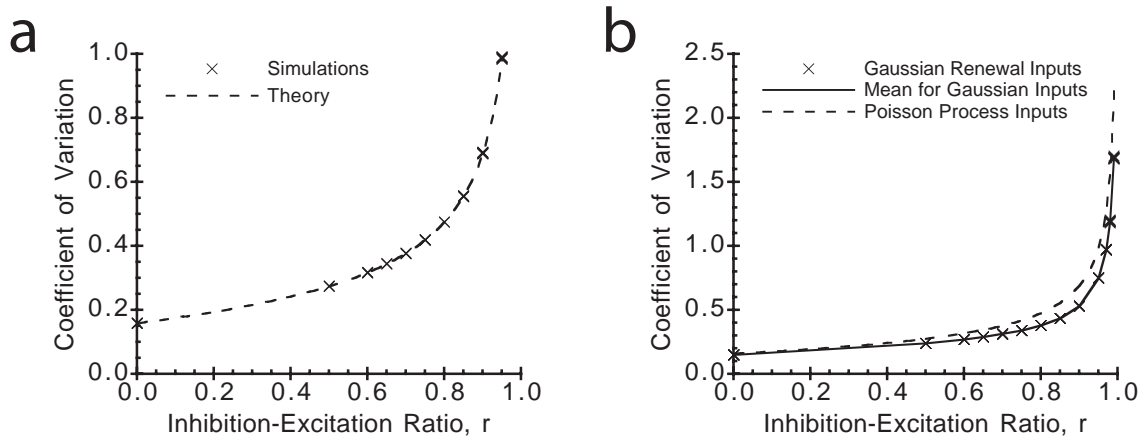
Figure 3.3: The coefficient of variation of interpoint intervals as a function of the inhibition-excitation ratio r for the output of an integrate-and-fire model with inputs that are (a) Poisson point processes and (b) positive-Gaussian-distributed renewal point processes. There were ten simulation runs at each value of r, and the calculated value of the coefficient of variation for each of these runs is plotted with a symbol. The dashed line in each graph is the theoretical curve for the model with Poisson process inputs, and the solid line in (b) connects the means calculated at each value of r.

In Figure 3.3, the calculated value of CV_{ISI} for each simulation is plotted versus the value of r. Again, at each location, there are actually ten symbols, but the low variability of the estimates causes the symbols to fall almost exactly on top of each other at all but the highest values of r. The solid line connects the means of the ten estimates of CV_{ISI} at each value of r, and the dashed line is the theoretical curve for the Poisson-input model. These results are in close agreement with those of Figure 2a of Feng and Brown (1998a). As we expect, due to the negative correlation in the superpositions of the inputs, the CV_{ISI} values for the positive Gaussian model are always less than that for the Poisson model. However, the CV_{ISI} for the positive Gaussian model does increase, apparently without bound, as r approaches one, in the same manner as that for the Poisson model. In particular, (3.17) specifies that the CV_{ISI} function with respect to r for the Poisson-input model is the square root of a rational function with a pole at one, and the increasing, convex shape of the CV_{ISI} curve for the Gaussian-input model is well-approximated by the same type of function. Thus, as observed by Feng and Brown (1998a), the positive Gaussian model requires more balance than the Poisson model in order to achieve any particular CV_{ISI} for its outputs. Only for values of r greater than about 0.9 is the CV_{ISI} greater than 0.5, which is the minimal value needed to match physiological data.

Although the positive Gaussian model may be able to achieve any arbitrary value of CV_{ISI}, albeit perhaps at the cost of a high degree of excitation-inhibition balance, its success as a model of cortical neurons quickly diminishes when the possibility of LRD is considered. Figure 3.4 contains examples of FFCs and IDCs for the output of the IF model with positive Gaussian renewal inputs for several representative values of the inhibition-excitation ratio r; the FFCs and IDCs for all of our simulations are shown in Appendix B.1. The sub-Poissonian variability of the interspike intervals in the output

caused by the integration mechanism and the negative correlation in the superpositions of the inputs is clearly evident in the general downward trend of both the original-data and surrogate-data FFCs and their asymptotic approach to a constant less than one for low r-values. Furthermore, as r is increased, i.e. as the amount of inhibition is brought into balance with the amount of excitation, the effect of the excitation-inhibition interaction is apparent in the rising value of the asymptotes of the FFCs. At a value of about $r = 0.97$ (not shown in Figure 3.4), the excitation-inhibition interaction completely cancels the two variance-decreasing effects, and the intervals of the output of the model have Poisson-like variability. Thus, for larger values of r, the asymptotic values of the FFCs are greater than one. Even at these large r-values, however, the FFCs are still below one for small counting intervals. For instance, the FFCs for $r = 0.99$ do not exceed one until the counting interval is about 0.1 seconds in length.

A small amount of negative correlation between the interspike intervals of the output is also evident at low r-values in the differences between the original-data FFCs and surrogate-data FFCs, as well as in the difference between the original-data IDCs and surrogate-data FFCs. The fact that the original-data FFCs are lower than the surrogate-data FFCs implies that the intervals in the original data must be negatively correlated, since the intervals in the surrogate data are uncorrelated but have the same distribution as the original-data intervals. This argument is supported by the original-data IDCs, which are decreasing at low aggregation levels and are below the corresponding surrogate-data IDCs. Since this negative correlation in the output intervals is apparent, and, in fact, its effect largest, at $r = 0$, where there is no excitation-inhibition interaction, it must be the result of the negative correlation between intervals in the superposition of the inputs. These effects of the negative correlation, however, gradually disappear as r increases. Thus, for r greater than about 0.9, the original-data curves and the surrogate-data curves are nearly equivalent, implying that the output of the model is roughly a renewal process for closely balanced amounts of excitation and inhibition.

Several additional characteristics of the FFCs and IDCs in Figure 3.4 are consistent with theoretical predictions. First, the surrogate-data IDCs are all close to being horizontal lines, and the mean surrogate-data IDC at each r-value is, for all practical purposes, a horizontal line. This is to be expected since surrogate data forms a renewal process. Next, the value of each FFC at the smallest counting intervals is nearly one, in agreement with (3.4), and each FFC has the same asymptotic value as the corresponding IDC, as predicted by (3.7). This also means that the vertical position of the IDCs increases with increasing r.

From this analysis, it is apparent that the positive-Gaussian-input IF model is no better at producing high interval variability and LRcD than the Poisson-input IF model. Not only does this model require perfect excitation-inhibition balance in order to produce LRcD, but, at any given level of balance, the Gaussian-input model is farther from being LRcD than the Poisson-input model. Furthermore, if level of balance is high enough to produce values of CV_{ISI} above 0.5 in the Gaussian-input model, then the output of this model is nearly renewal, which is precisely the characteristic that was used to expose the inadequacy of the Poisson-input model. In sum, the positive-Gaussian-input IF model does not solve the high-variability problem while producing LRD and is, in fact, worse in this regard than the more tractable Poisson-input model.

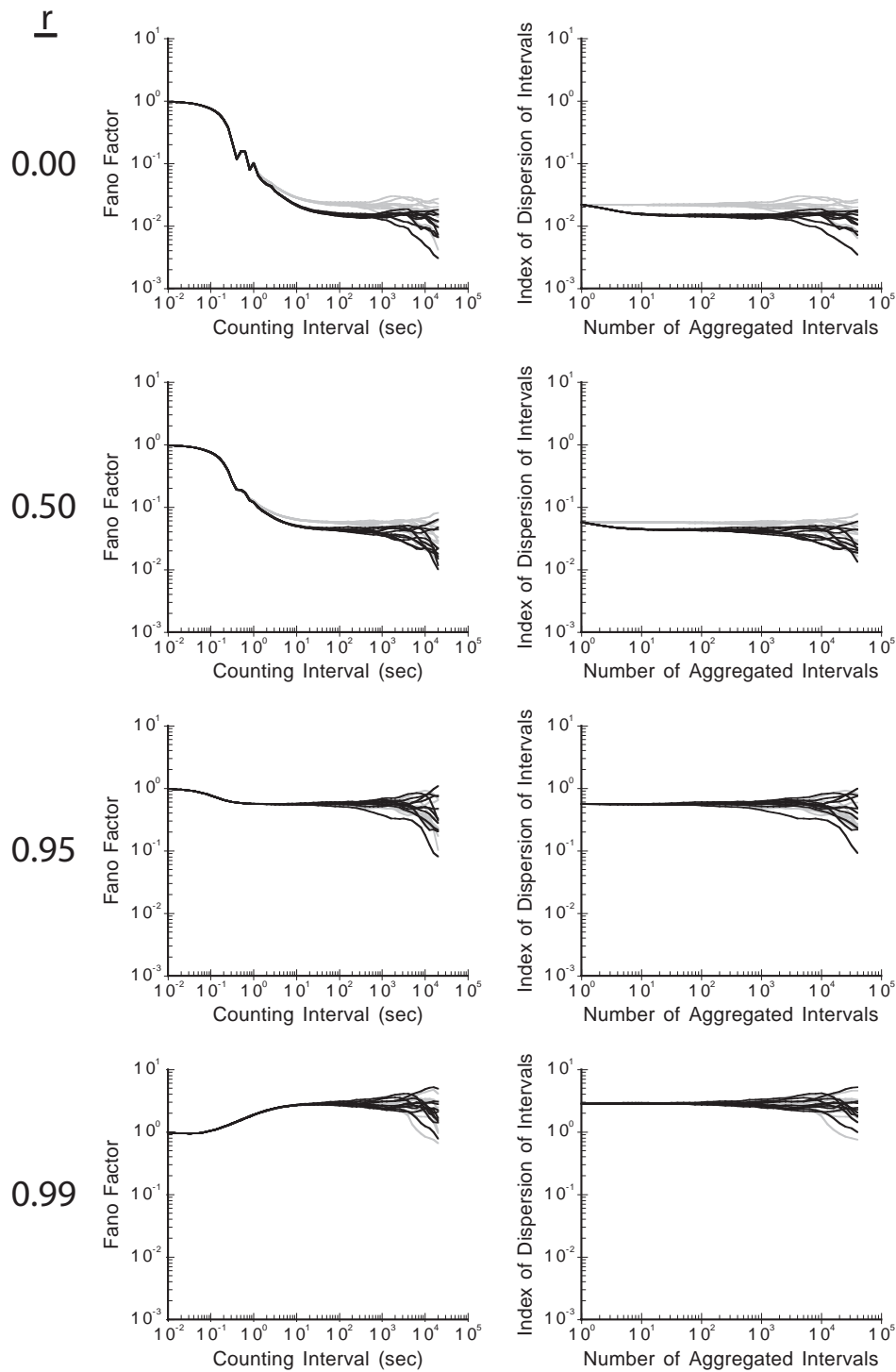Figure 3.4: FFCs and IDCs estimated from simulations of the IF model with positive-Gaussian inputs. Each set of axes contains ten curves calculated from original data (black) and ten curves calculated from the corresponding shuffled surrogate data (gray). For each value of the inhibition-excitation ration r, each individual FFC in the left set of axes was calculated from the same data as one of the IDCs in the right set of axes.

The IF Model with Pareto Renewal Inputs

The IF model with Pareto RPP inputs has three parameters: the inhibition-excitation ratio r, the "tail" parameter α, and the normalization constant K. For any given values for r and $\alpha > 1$, the value of K was determined by setting the mean of the Pareto distribution equal to the inverse of the input rate determined by (3.16). Thus, we set $K = (\alpha - 1)/input\_rate$ if $\alpha > 1$. When $0 < \alpha \leq 1$, the Pareto distribution does not have a mean, and, hence, the previous calculations are not justified. $\alpha = 1$ was the only value in this range that we used, and in this case we set $K = 0.1/input\_rate$, where $input\_rate$ was still determined by (3.16), which is the same value that we would have calculated for K when $\alpha = 1.1$. We found empirically that this value yielded an output spike rate in the vicinity of 2.5 spikes per second for our simulations.

Thus, for the Pareto-input model simulations, we varied two parameters: α and r. For α we used values of 1.0, 1.25, 1.5, 1.75, 1.9, 2.0, 2.1, 2.5, and 3.0, while for r we used values of 0.0, 0.5, 0.7, 0.8, 0.9, and 0.95. Ten simulations were run for each combination of these parameter values. For comparison, Feng and his coworkers (Feng, 1997; Feng & Brown, 1998b, 1998a; Feng et al., 1998) used only values of $\alpha = 1.0$ and $\alpha = 2.1$ in their simulations. We felt that the choice of these two parameter values missed two important ranges. To begin with, they did not use any values of α between 1.0 and 2.0, exclusive, where the Pareto distribution has infinite variance, but still has a finite mean. Also, they did not use any values that were significantly greater than two. Although at the value of $\alpha = 2.1$, the Pareto distribution does have finite mean and variance, this value is close enough to 2.0 that the differences between finite and infinite variance are not going to be very obvious.

Figure 3.5 shows the estimated values of CV_{ISI} from these simulations plotted versus the inhibition-excitation ratio r. The graphs in each column contain the same data, only at different scales. Column (a) contains graphs of the mean values of CV_{ISI} for each combination of parameter values for r and α. Thus, each symbol represents the mean calculated over ten simulations. Means calculated from simulations with the same value of α are connected by solid lines, while the dashed line is the theoretical curve for the Poisson-input model. The curves all have the same increasing, convex shape, which can be well-fit by the square root of a rational function with a pole at $r = 1$. Also, as we expect due to the positive correlations in the superpositions of the inputs, the CV_{ISI} values for the Pareto model are always greater than those for the Poisson model. Thus, in accord with the findings of Feng and Brown (1998a), the Pareto model requires less balance than the Poisson model to achieve any particular CV_{ISI} of its output. Furthermore, as α is decreased, a particular value of CV_{ISI} can be attained with less balance between the excitation and inhibition. For example, at low values of α, the CV_{ISI} is actually significantly larger than physiological measurements in the cortex at even moderate degrees of balance.

Column (b) of Figure 3.5 shows, for a subset of the values of α, the CV_{ISI} estimates for all simulation runs. The solid lines are identical to those in Column (b), connecting the mean for each set of ten simulation runs. Although at low r values and high α values the symbols are tightly grouped, like the CV_{ISI} estimates for the positive Gaussian model, the variance of these estimates of CV_{ISI} increases significantly as r increases and as α
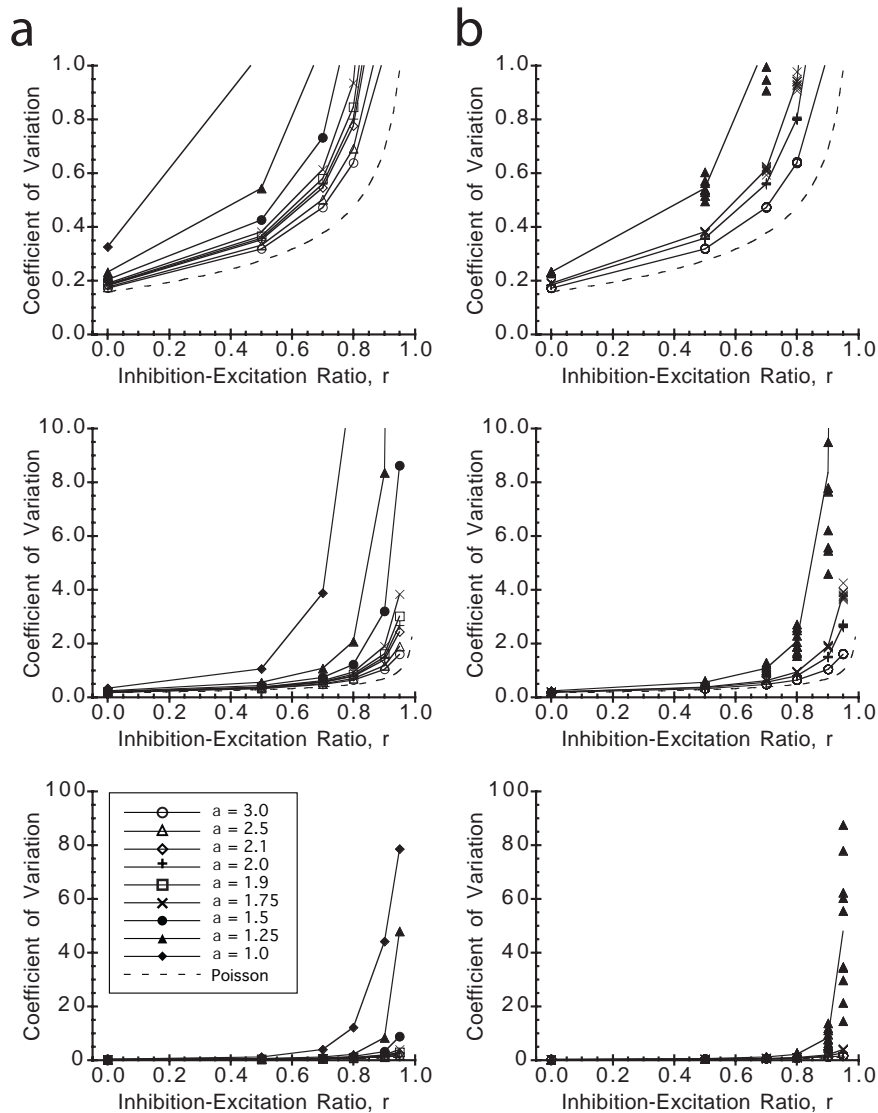
Figure 3.5: The coefficient of variation of interpoint intervals as a function of the inhibition-excitation ratio r for the output of an integrate-and-fire model with inputs that are Pareto-distributed renewal point processes. For each combination of the parameters, r and the "tail" parameter α of the Pareto distribution, ten separate simulation runs were conducted. In column (a), the mean values of the estimates are plotted for each combination of parameters. Each solid line connects the means for a single value of α, and the dashed line is the theoretical curve for the model with Poisson process inputs. All three sets of axes contain the same data, but each displays them on a different vertical scale. In column (b), the estimates calculated from each individual simulation run are plotted for a subset of the values of α. Solid and dashed lines are the same as in column (a), and all three sets of axes contain the same data.

decreases. The most dramatic effect is seen for medium-to-high values of r and low values of α. These increases in estimator variance are almost certainly caused by the progressive movement of the model towards a state of non-stationarity as r or α approach one. We saw in Section 2.4 that, when r is equal to one, the mean of the intervals in the output of the Poisson-input model is infinite, and no point process with infinite mean intervals can be stationary. This is likely to be the case with the present model as well. On the other hand, when α equals one, the input processes have infinite mean intervals, also forcing the model to be non-stationary.

As mentioned above, Feng and Brown (1998a) used only two values of α: 2.1 and 1. Comparing Figure 3.5 with their Figure 2a, we see that our results for $\alpha = 2.1$ are in good agreement with their corresponding results. In contrast, our CV_{ISI} values for $\alpha = 1$ are substantially larger than those in their Figure 2b. For instance, at $r = 0.9$, the most balanced condition that they used, they have $CV_{ISI} \approx 3$, while, at the same value of r, we have $CV_{ISI} \approx 80$. Since Feng and Brown (1998a) do not give many details regarding their simulations, it is difficult to explain this discrepancy. However, we did use the same number of excitatory inputs and an equivalent combination of threshold and postsynaptic potential parameters as they did. One obvious difference between our simulations and theirs is in the parameter K. They did not use this parameter, but the form of their Pareto distribution is equivalent to (3.13) with K set to one. Recall that, for $\alpha = 1$, our value of K was determined by

$$K = \frac{0.1}{input\_rate} = \frac{100 \cdot 0.025 \cdot (1 - r) \cdot 0.1}{2.5}.$$

This means that in our simulations K decreased linearly from 0.1 at $r = 0$ to 0.005 at $r = 0.95$, which is one-tenth to 1/200th of their value of K. Thus, their output (and input) rates would have been much smaller than ours, yielding much larger interspike interval lengths. Since the mean interval length when $\alpha = 1$ does not exist theoretically, it seems plausible that the larger intervals in their simulations, and thus larger sample means, could result in reduced values for the sample CV_{ISI}, though it is difficult to infer the accompanying changes in the sample standard deviation of the intervals. Furthermore, Pareto RPPs with $\alpha = 2.1$ and $\alpha = 1$ are quite different; the first has intervals with finite mean and variance, whereas both the mean and variance of the intervals for the second are, in theory, infinite. Thus, we might expect a large difference between models with these two values of α. However, the CV_{ISI} for their simulations at $r = 0.9$ only changes from about 1.2 to 3, whereas ours changes from 1.4 to 44, the latter seeming to us to be more likely. We did, nevertheless, rerun our simulations with $\alpha = 1$ and $K = 1$ to match those of Feng and his coworkers. For these simulations, the values of CV_{ISI} were significantly reduced from those where we fixed the output spike rate. At $r = 0.9$, the mean CV_{ISI} value for these simulations with $K = 1$ was approximately 8.3, much closer to their value but still larger. Without further information, we cannot expect to determine the cause of this difference, but it is not surprising given the nature of the model with $\alpha = 1$. At this value of α, the inputs cannot be stationary since their mean does not exist. In our simulations, we attempted to produce stationary-like inputs, but it is likely that they began all of their inputs with a point of the point process, whether actual or virtual, at time zero. Given the odd statistical nature of the model at $\alpha = 1$, it would not be

astonishing if this difference in the starting conditions produced the discrepancy between the two sets of results.

Figure 3.6 contains FFCs and IDCs calculated from simulated output of the IF model with Pareto RPP inputs for several different representative combinations of values for the parameters α and r. The curves for $\alpha = 2.5$, $\alpha = 1.75$, and $\alpha = 1.0$ are shown in 3.6a, 3.6b, and 3.6c, respectively. For each of these values of α, FFC/IDC sets are shown for each of four values of r ranging from zero to 0.90 or 0.95. The complete set of FFCs and IDCs for all of our simulations are in Appendix B.2.

These graphs possess some common characteristics across all parameter values. First, as should be the case in all situations, the FFCs approach the value of one at very small counting interval lengths, and the mean surrogate-data IDCs are horizontal lines. Second, like those of the positive Gaussian model, the FFCs have an initially negative slope as a result of the relative unlikelihood of the occurrence of very short intervals, which is caused by the integration mechanism. Third, unlike those of the positive Gaussian model, the original-data IDCs generally have an initial positive slope and remain above the corresponding surrogate-data IDCs. This is a result of the positive correlation between intervals in the superposition of Pareto renewal processes, in contrast to the negative correlation in the superposition of positive-Gaussian renewal processes. These positively correlated intervals also produce positive slopes in the original-data FFCs subsequent to their initial declination from one. This produces original-data FFCs that have an asymptotic value larger than one, except when the value of α is large and the value of r is small. In the latter case, the positive correlation in the superpositions, combined with the variance producing effect of the excitation-inhibition interaction, is not strong enough to overcome the effect of the integration process.

Like those for the positive-Gaussian model, the original-data and corresponding surrogate-data curves for the Pareto model become progressively more similar as the excitation and inhibition are brought into better balance, i.e. as r is increased. This means that the surrogate-data FFCs change more quickly with increases in r than do the original-data FFCs, and that the original-data IDCs better approximate horizontal lines as r increases. In the current model, however, this is the result of a loss of positive, not negative, correlation in the output intervals. Nevertheless, this loss of correlation implies that the output of the model resembles a renewal point process at values of r close to one.

In Section 3.5, we discussed two possible ways of producing LRD in the IF model with RPP inputs. First, LRD may be produced through the balance of excitation and inhibition. In the Poisson-input model this could occur, but, since the output of the model was a renewal process, the inhibition-excitation balance also created output intervals with infinite variance. Our simulations suggest that the balance of excitation and inhibition has a similar effect on the Pareto-input model, the output of which becomes more RPP-like with increasing r, and the CV_{ISI} of which apparently increases without bound as r increases. Thus, the balance of excitation and inhibition is also unable to produce a realistic form of LRD in this model. The second way that LRD can be present in the output of the RPP-input IF model is if the superpositions of the inputs are LRD. According to Theorem 3.4, this will occur when, and only when, the intervals of the input have infinite variance. For Pareto RPP inputs, this condition is met when $\alpha \leq 2$. From our simulations, it is clear that in this case, although the output of the model is LRD, the
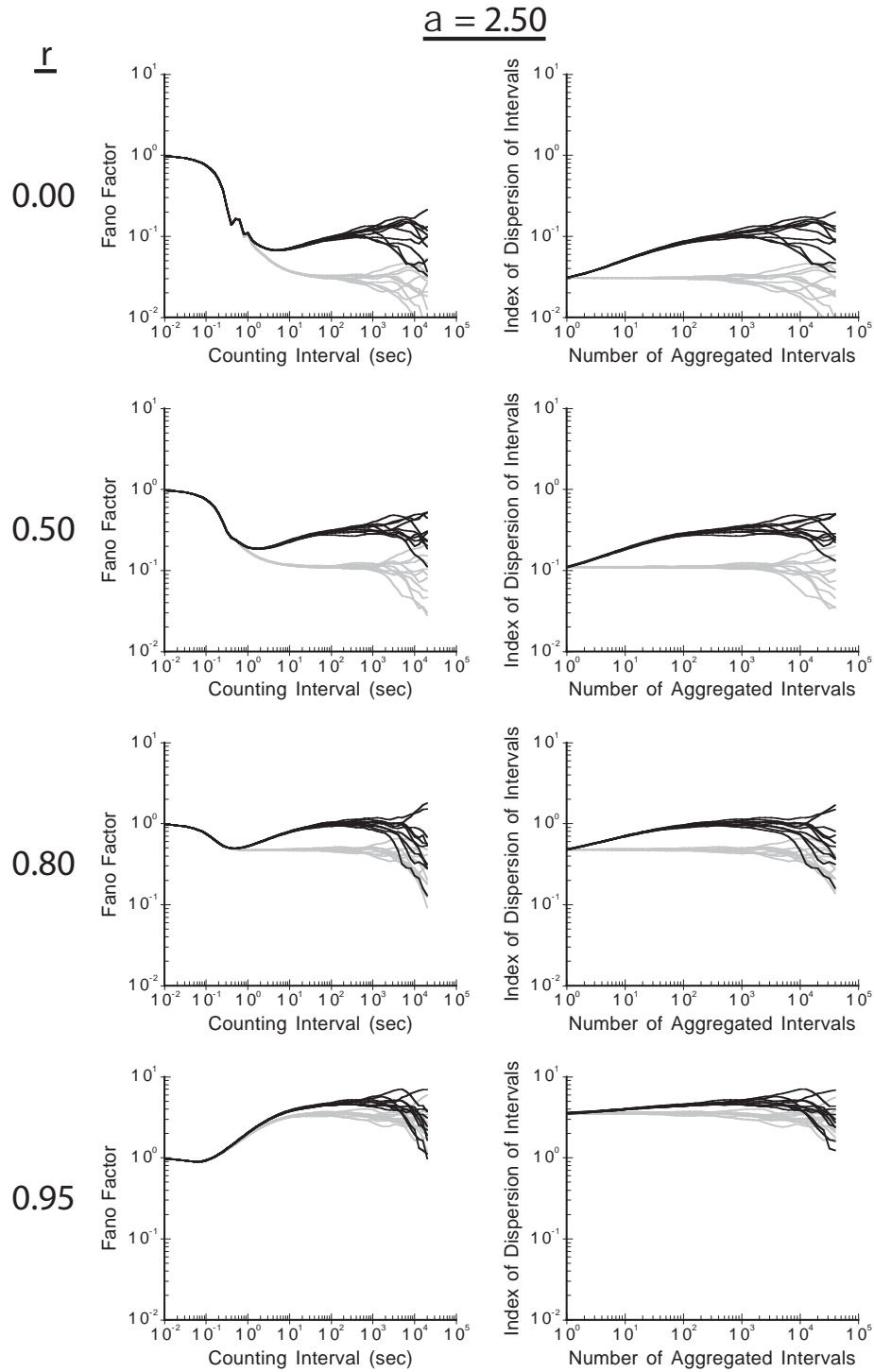
Figure 3.6a: FFCs and IDCs estimated from simulations of the IF model with Pareto inputs and parameter $\alpha = 2.5$. Each set of axes contains ten curves calculated from original data (black) and ten curves calculated from the corresponding shuffled surrogate data (gray). For each value of the inhibition-excitation ration r, each individual FFC in the left set of axes was calculated from the same data as one of the IDCs in the right set of axes.
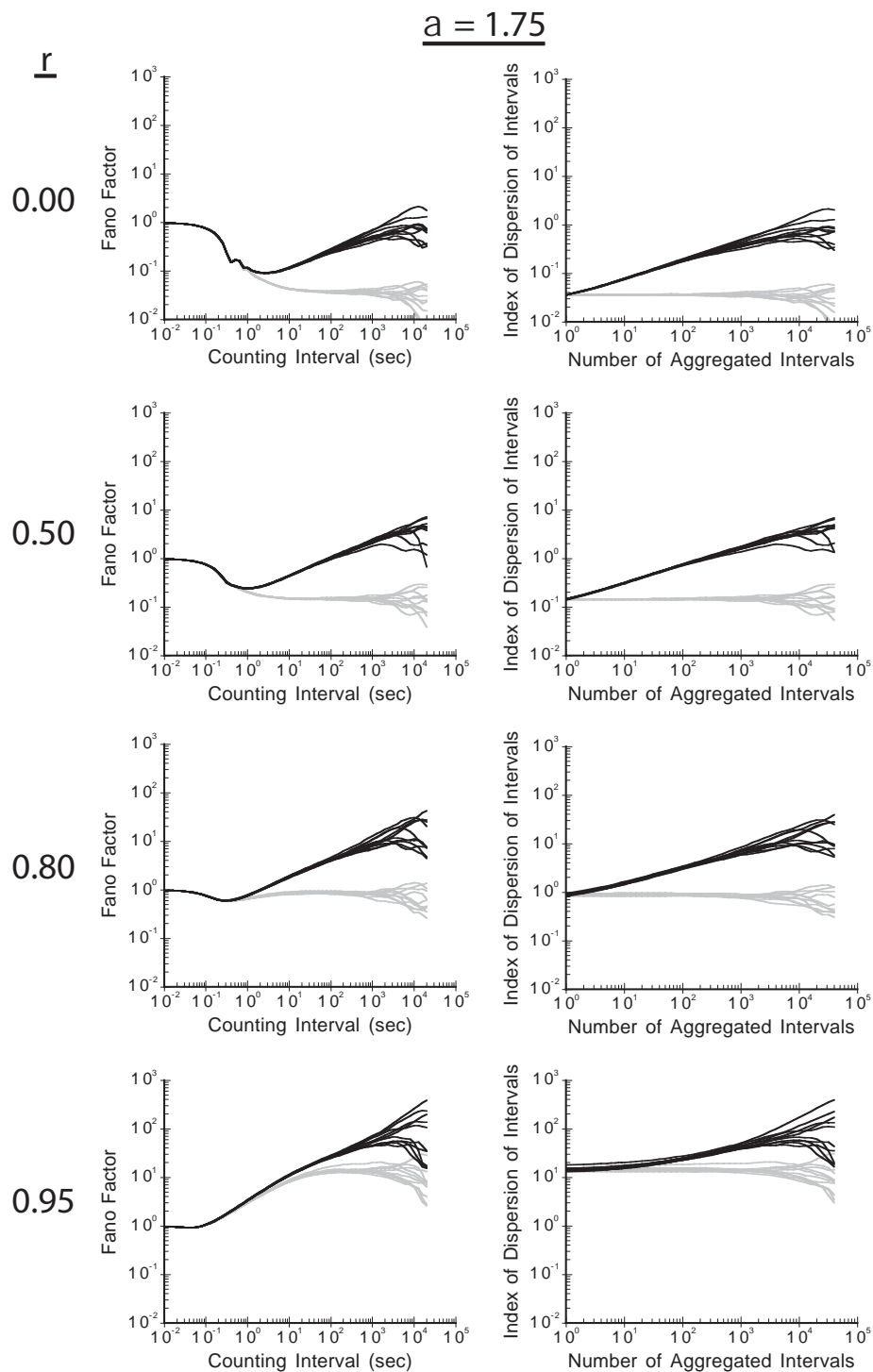
Figure 3.6b: Same as in 3.6a, except for $\alpha = 1.75$.
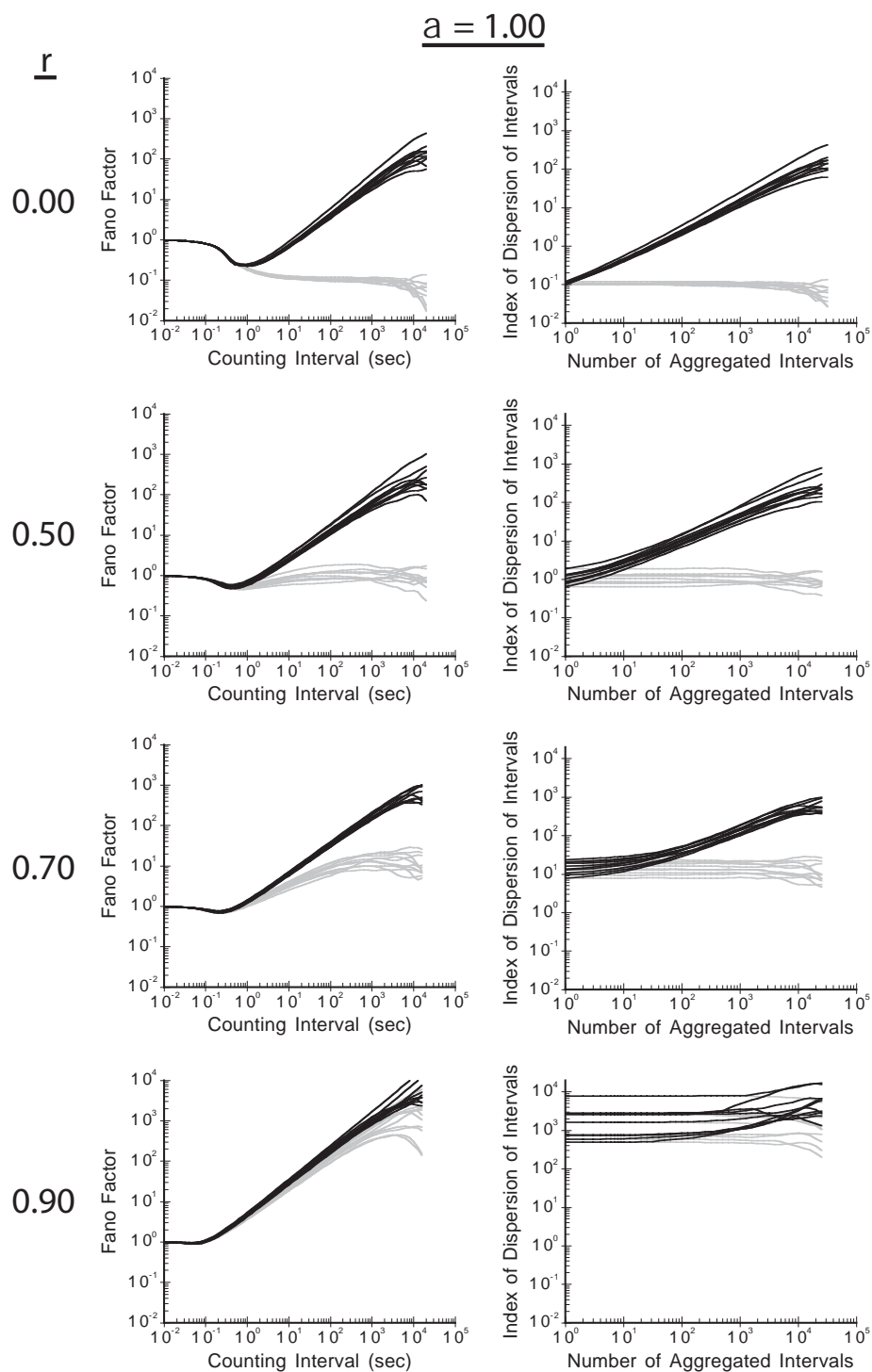
Figure 3.6c: Same as in 3.6a, except for $\alpha = 1.0$. Note that the larger values of r differ slightly from those in 3.6a and 3.6b.

nature of the LRD is different than that produced by inhibition-excitation balance.

To begin with, we observed that the maximum slopes of the IDCs decrease as the value of r increases. In contrast, as the variability of the inputs increases, i.e. as the value of α decreases, the slopes of the IDCs increase. In fact, whereas tightly balanced excitation and inhibition produces IDCs that are horizontal lines, high input variance produces IDCs that are positively sloped lines in double logarithmic coordinates, i.e. power-laws in linear coordinates. Thus, while more balance reduces the dependence between the intervals of the output of the model, lowering the value of α strengthens the interval dependence. Furthermore, the effect of the reduction of α on the variance of interval aggregations is greater at larger aggregation levels. Hence, the range of the interspike interval dependencies produced by the high variability is predominately long in nature, and the IDCs suggest the presence of LRiD. This progression in the original-data IDCs for decreasing values of α coincides with a similar change in the original-data FFCs. As α decreases, the portion of the FFC to the right of the minimum also progressively approximates a positively sloped line (on a double logarithmic plot). This is consistent with Proposition 3.3, which states that LRcD and LRiD always coincide when the interval variance is finite. On the other hand, as expected for a point process with finite-variance intervals, the FFCs for shuffled surrogate data apparently always asymptote to a finite constant, even when r is large and α is small. In addition to increasing as r increases, the asymptotic value of the surrogate-data FFCs increase as α decreases. But, as a result of LRiD, these FFCs still remain below their original-data counterparts. Moreover, the difference between the original-data FFCs and the surrogate-data FFCs at longer counting intervals actually tends to increase as α decreases. This trend may, however, break down as α approaches one, where the mean interspike interval of the input processes becomes infinite.

It is interesting that LRcD in the Pareto RPP inputs effected by infinite interval-variance produces LRcD that is effected by LRiD in the output of the IF model. However, the production of an LRD process by the combination of a number of infinite-variance processes has been seen in other models as well. A common model for the traffic on data networks, such as Ethernet Local Area Networks (LANs), is the sum of a large number of *ON/OFF* sources. The trace in the center of Figure 3.7, if the lower value is zero, would be an example of the value of a single *ON/OFF* source as a function of time. The *ON/OFF* source model has independent *ON*- and *OFF*-times, and, typically, all of the *ON*-times are identically distributed (positive) random variables and all of the *OFF*-times are identically distributed (positive) random variables. The *ON*- and *OFF*-periods are allowed to have different distributions. Thus, this model is also called an "(strictly-)alternating renewal process". Classically, the distributions of the *ON*- and *OFF*-periods had finite variance. However, more recently, it has been shown, both through simulation and analytically, that the sum of a large number of *ON/OFF* processes with long-tailed (i.e. infinite-variance) distributions for the *ON*- and/or *OFF*-periods produces an LRD process (Willinger, Taqqu, Sherman, & Wilson, 1997; Taqqu, Willinger, & Sherman, 1997). Furthermore, in earlier theoretical work, the aggregation of "idealized" *ON/OFF* sources with long-tailed *ON*-/*OFF*-periods was shown to exhibit this effect as well (Mandelbrot, 1969; Taqqu & Levy, 1986; Willinger, Taqqu, Sherman, & Wilson, 1995). In the standard *ON/OFF* model described above each *ON*-period is followed by an *OFF*-period, and

vice-versa. But in the "idealized" form of this model, an *ON*-period (or *OFF*-period) can be followed by either an *OFF*- or an *ON*-period. For related work, see Lowen and Teich (1993c, 1993d) and Heath, Resnick, and Samorodnitsky (1998).

To see the relevance of this work to our present situation, consider the points at which the *ON/OFF* process in Figure 3.7 transitions between states. If the process is a standard (strictly-alternating) *ON/OFF* process, then the black dots (both above and below) represent these points. If it is an "idealized" *ON/OFF* process, then the gray dots might also be transition points. Since the *ON/OFF* process enters the *ON*-state at the points shown above the trace, the top sequence of points might be called the "excitatory" point process. Conversely, the bottom sequence of points, which is composed of the points at which the *ON/OFF* process enters the *OFF*-state, could be called the "inhibitory" point process. In an analogous fashion, if an IF model had a single excitatory input equivalent to the top point process (black dots only) in Figure 3.7 and a single inhibitory input equivalent to the bottom point process (black dots only), then the (sub-threshold) potential of the IF model would be equivalent to the center trace. Moreover, the sum of a large number of *ON/OFF* processes would be similar to the potential in a IF model with many excitatory and inhibitory inputs. Of course, the excitatory and inhibitory RPP inputs in our IF model do not alternate pairwise, as the two point processes do in Figure 3.7. But, when the IF model has a large number of inputs, whose effects are summed, this difference is unlikely to result in significant differences in the long-term statistical structure of the aggregation of *ON/OFF* processes and the potential of the IF model.

In general, then, our simulations of the IF model with Pareto-RPP inputs show that, with proper adjustment of the parameters r and α, this model can produce outputs that simultaneously exhibit both LRD and high interspike interval variability like that found in cortical neurons. More specifically, from the data that was plotted in Figure 3.5, we see that the CV_{ISI} of the output of this model is between 0.5 and 1.5, a typical range of CV_{ISI} estimated from recordings of cortical neurons, for the ranges of r shown in Table 3.2 for each value of α used in our simulations. Combining this information with comparisons between the FFCs obtained from simulations of this model, examples of which are shown in Figure 3.6, and the FFCs obtained from physiological recordings that are shown in Teich et al. (1996), we can evaluate the ability of the Pareto-input model to match the variability and LRD of *in vivo* cortical neurons.

For $2.0 < \alpha \leq 3.0$, when the interval variance of the inputs is finite, a value of r between about 0.7 and 0.9 is required to match physiological CV_{ISI} estimates. For r values near 0.9, these FFCs resemble those of Cell 4 and Cell 7 in Teich et al. (1996), although neither the simulation nor the physiological results suggest the presence of LRcD. These r-values yield CV_{ISI}-values on the upper end of the physiological range, always being larger than 1.0. For lower values of r, the surrogate-data FFCs, and some original-data FFCs as well, tend to asymptote below 1.0. For r values above this range, not only are the CV_{ISI} values too high, but the output of the model is too much like a renewal process, with overlapping original-data and surrogate-data FFCs.

At α-values between 1.5 and 2.0, inclusive, the simulation FFCs with r at the high end of their CV_{ISI}-matching range tend to resemble those of Cell 3 and Cell 19 of Teich et al. (1996). Again, the output at these parameter values will have a CV_{ISI} value above
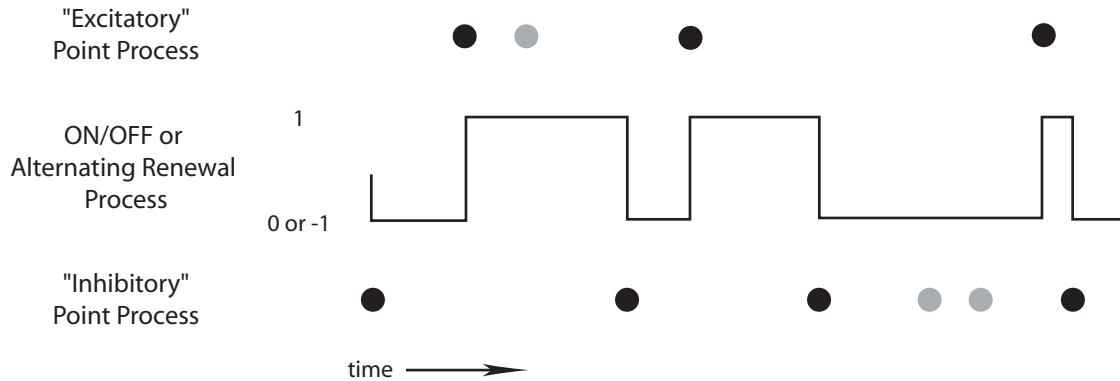
Figure 3.7: An analogy between the effect of the point process inputs on an IF model and a common model in both economics and the study of data network traffic. The center trace is an example of the time plot of the value of what is called either an *ON/OFF* process (if the lower value is "0") or an alternating renewal process (whether the lower value is "0" or "-1"). This continuous process has independent and identically distributed *ON*- and *OFF*-periods. Each point, above or below, designates the time at which the *ON/OFF* process switches states. If only the black points are included, then the center trace represents a standard *ON/OFF* process, or a strictly-alternating renewal process, where an *ON*-period is always followed by an *OFF*-period and vice-versa. If the gray dots are also included, then the center trace represents an "idealized" *ON/OFF* process, where an *ON*-period can be followed by either an *OFF*-period or another *ON*-period. Since the *ON/OFF* process goes into the *ON* state when a point occurs in the upper sequence, this point process is "excitatory". Since it goes into the *OFF* state when a point occurs below, the lower point process is "inhibitory". Thus, this center trace is in some ways analogous to the sub-threshold potential of an integrate-and-fire model with a single excitatory input and a single inhibitory input. See the text for further explanation.

α	Range of r
3.0	$0.75 < r < 0.95$
2.5	$0.70 < r < 0.93$
2.1	$0.68 < r < 0.91$
2.0	$0.66 < r < 0.90$
1.9	$0.64 < r < 0.89$
1.75	$0.61 < r < 0.86$
1.5	$0.57 < r < 0.82$
1.25	$0.47 < r < 0.75$
1.0	$0.20 < r < 0.56$

Table 3.2: The range of the inhibition-excitation ratio r, for each value of α, that leads to values of CV_{ISI} between 0.5 and 1.5 for the integrate-and-fire model with Pareto renewal process inputs. The ranges of r were approximated from plots of CV_{ISI} versus r (see Figure 3.5) obtained from simulations of the model.

1.0, but the FFCs do suggest the presence of LRcD. At lower values of r, LRcD still is present, but the surrogate-data FFCs asymptote at or below 1.0, which may match results from neurons with $CV_{ISI} < 1.0$ if those FFCs were available. At higher values of r, the FFCs still retain a resemblance to Cell 3 and Cell 19 of Teich et al. (1996), but their CV_{ISI}-values are too large. Of course, if r were made close enough to 1.0, the output of the model would again be too renewal-like.

For values of α close to 1.0, the output of the model is still LRcD, but the surrogate-data FFCs asymptote at or below 1.0 for r-values that lead to physiological values of CV_{ISI}, as well as for even lower values of r. Slightly larger values of r lead to FFCs resembling those of Cell 3 and Cell 19 of Teich et al. (1996), while even larger r-values produce renewal-like outputs in the model. The latter effect is likely due to the compounding of the extremely variable inputs (even their mean interval lengths are or are almost infinite) with the effect of tightly balanced amounts of excitation and inhibition. Furthermore, as α approaches one, the inputs are approaching conditions of non-stationarity, which result in a significant increase in the variability our estimates of the Fano factor and the index of dispersion of intervals. This can be seen in the increased "spread" of the curves for $\alpha = 1.0$ in Figure 3.6c.

Although the Pareto-input model is able to produce outputs which share common statistical features with the spike trains in cortical neurons, the type of inputs required are not justified physiologically. In order to produce LRD in the output of the model, the inputs must have infinitely variable intervals, which is not the case for neurons that might serve as inputs to cortical neurons: other cortical neurons or sub-cortical, especially thalamic, neurons. This appears to be the fundamental weakness of the Pareto-input model. In the following section, we will suggest another relatively simple model that remedies this problem.

3.6 Integrate-and-Fire Models with Fractional-Gaussian-Noise-Driven Poisson Process Inputs

In Section 3.5 we considered the IF model with RPP inputs. RPPs and their superpositions are LRD if, and only if, their intervals have infinite variance. In this section, we consider the IF model with inputs that are LRD without having infinitely variable intervals. More specifically, we have chosen to use inputs that are fractional-Gaussian-noise-driven Poisson processes.

3.6.1 The Fractional-Gaussian-Noise-Driven Poisson Process

The stationary, or homogeneous, Poisson process is the simplest stochastic point process. The probability of occurrence of a point in this model is uniform throughout time, parameterized by the rate of occurrence λ, and is completely independent of the past. In addition, the occurrence of multiple points at any one time instant is virtually impossible. Customarily, this is the first model that is used for situations that can be represented as a series of events, due to its simplicity and analytical tractability, and the inputs of the high variability IF models of cortical neurons are no exception. Nevertheless, with more

study of the process to be modeled, one usually finds that more complicated models are eventually required. The inputs of the models considered in Section 3.5 retained some of the temporal independence of the Poisson process, but allowed modification of the interval distribution. The Poisson process can, however, be generalized in such a way that temporal dependence can be specified, but certain distributional properties are retained. This generalization involves replacing the (constant) rate parameter with a time-dependent function $\lambda(t)$, resulting in a nonstationary, or nonhomogeneous, Poisson process. However, to use this as a model, one must know of some deterministic trend in the rate of occurrence of points, in order to define the function $\lambda(t)$.

Often a more useful point process model is produced by replacing the rate parameter with a stochastic process $\Lambda(t)$, yielding what is called a doubly stochastic Poisson process (DSPP)—or, sometimes, called a "Cox process" since it was proposed in a seminal paper by Cox (1955). With the DSPP, we can again have a stationary point process if the stochastic process $\Lambda(t)$ is stationary, but now the point process can be dependent on its past. In essence, the DSPP has a dependency structure equivalent to its stochastic rate process. In particular, for our purposes, if its stochastic rate process is LRD, then so is the DSPP.

Fractional Gaussian noise is an LRD stochastic process that was introduced by Mandelbrot and coworkers (Mandelbrot, 1965; Mandelbrot & Wallis, 1968, 1969a, 1969c, 1969b; Mandelbrot & van Ness, 1968) in order to stochastically model the fluctuations in water levels of the Nile River, which, Hurst (Hurst, 1951) had previously discovered, possessed long memory. Fractional Gaussian noise (fGn) is the increment process of fractional Brownian motion (fBm), which is the only Gaussian process that both is self-similar and has stationary increments. fBm is parameterized by the self-similarity index $0 < H < 1$, where $H = 0.5$ corresponds to the ordinary Brownian motion with independent increments. Thus, fGn can be parameterized by the self-similarity index of the corresponding fBm, with $H = 0.5$ corresponding to the ordinary Gaussian (or "white") noise. When $H < 0.5$, fGn has negative dependence, and when $H > 0.5$, fGn is LRD. Thus, we will only be interested in fGn with H in the range $0.5 \leq H < 1$.

The fractional-Gaussian-noise-driven Poisson process (fGnDP) is created by using a modified form of fGn as the stochastic rate process of a DSPP. Modification of the fGn is necessary for several reasons. First, the rate of a Poisson process cannot be negative, whereas samples of fGn can certainly assume negative values. Second, fGn has mean zero, so that simple truncation of the fGn at zero will significantly change its statistical properties. Third, fGn, like ordinary (white) Gaussian noise, is inherently a discrete process, while the rate process for a DSPP should be defined in continuous time. Let $\{G_H(k), k \in \mathbb{Z}\}$ denote standard fGn, i.e. with a mean of zero and a variance of one. Then, in order to remedy the aforementioned incongruencies, we let the rate process of a DSPP be[3]

$$\Lambda(t) = \max\left\{0, \lambda + \sigma G_H\left(\left\lfloor \frac{t}{\tau} \right\rfloor\right)\right\}, \tag{3.19}$$

---

[3]This rate function is not actually a stationary stochastic process since time zero is always at the beginning of a τ-length sampling interval. However, this "small" non-stationarity facilitated mathematical analysis and did not affect our results in test simulations due to the length of the simulations and analysis methods.

where λ and σ are positive constants and $\lfloor x \rfloor$ is the largest integer less than x. Assuming that $\sigma \ll \lambda$, so that the probability that $\lambda + \sigma G_H < 0$ is negligible, the mean rate of the fGnDP with the rate process in (3.19) is $E\{\Lambda(t)\} \approx \lambda$. In addition, if $N(t)$ is a DSPP with the rate process in (3.19), then the variance of $N(t)$ will be approximately λt plus a term proportional to σ^2.

Teich et al. (1990) suggested using "fractal-noise-driven" Poisson process models for neural spike trains of the auditory nerve in order to account for the long-range dependence, or "fractal behavior", that had been previously discovered (Teich, 1989). In subsequent papers, the theory and application of this model to the auditory nerve and other peripheral sensory-system neurons was further developed (Teich, 1992; Teich & Lowen, 1994; Lowen & Teich, 1993a, 1995, 1996a, 1997; Kumar & Johnson, 1993; Thurner et al., 1997). The fractional-Gaussian-noise-driven Poisson process[4] is, arguably, the simplest of these models. Furthermore, since the theory of fractional Gaussian noise has been well developed in many different contexts, and since the other "fractal" noises that were suggested as the driving noise in these models converge to fractional Gaussian noise under appropriate conditions, the fGnDP is a natural starting point for using "fractal-noise-driven" Poisson process models.

The fGnDP was used as a model of the spike trains in the auditory nerve because it is LRcD (for $H > 0.5$), but, unlike LRD renewal processes, it is also LRiD and the variability of its inputs is finite and similar to a Poisson process. Thus, its FFCs and IDCs resemble those in Figure 3.2b. These properties are shared by the auditory nerve, as well as most sub-cortical neurons that have been studied with respect to such properties. Thus, a further justification of the IF model with fGnDP inputs is the fact that those thalamic neurons that have been studied with respect to LRD and that also project into the cortex possess these properties as well.

3.6.2 The Superposition of Fractional-Gaussian-Noise-Driven Poisson Processes

The only difference between the present model and the model of Section 3.5 is the form of the input processes. Thus, much of the reasoning in that section regarding general properties of the model apply here as well. In particular, we still expect the fGnDP-input model to exhibit negative correlation on small time scales, due to the integration mechanism, and this negative correlation will presumably manifest itself in a negatively sloped FFC for small counting windows. Also, excitation-inhibition balance will, most likely, still produce memory over longer ranges by producing high interval variability. Finally, additional memory properties of the output, if present, should be evident in the superposition of the inputs. Therefore, we will next consider the properties of the superposition of fGnDPs.

The superposition of two independent DSPPs is also a DSPP, as proved in the following theorem.

---

[4] In this literature, the *fractional*-Gaussian-noise-driven Poisson process is called the *fractal*-Gaussian-noise-driven Poisson process. We have used the former in accordance with mainstream mathematical and analytical literature on *fractional* Gaussian noise.

Theorem 3.5. *Let $N_1(t)$ and $N_2(t)$ be independent doubly stochastic Poisson processes with (independent) rate processes $\Lambda_1(t)$ and $\Lambda_2(t)$, respectively. Then the superposition $N(t) = N_1(t) + N_2(t)$ is a doubly stochastic Poisson process with rate process $\Lambda(t) = \Lambda_1(t) + \Lambda_2(t)$.*

Proof. The probability generating functional of a doubly stochastic Poisson process $N(t)$ with stochastic rate process $\Lambda(t)$ is (e.g. Cox & Isham, 1980, p. 71; Daley & Vere-Jones, 1988, Proposition 8.5.I)

$$G_N[\xi] = E_\Lambda \left\{ \exp \left[\int_{-\infty}^{\infty} (\xi - 1)\Lambda(t)\, dt \right] \right\},$$

where E_Λ denotes expectation with respect to Λ. Therefore, if $N_1(t)$ and $N_2(t)$ are independent doubly stochastic Poisson processes with independent stochastic rate processes $\Lambda_1(t)$ and $\Lambda_2(t)$, respectively, then the probability generating functional of their superposition $N(t) = N_1(t) + N_2(t)$ is (e.g. Cox & Isham, 1980, p. 41; Daley & Vere-Jones, 1988, Proposition 7.4.VII)

$$
\begin{aligned}
G_N[\xi] &= G_{N_1}[\xi] G_{N_2}[\xi] \\
&= E_\Lambda \left\{ \exp \left[\int_{-\infty}^{\infty} (\xi - 1)\Lambda_1(t)\, dt \right] \right\} \cdot E_\Lambda \left\{ \exp \left[\int_{-\infty}^{\infty} (\xi - 1)\Lambda_2(t)\, dt \right] \right\} \\
&= E_\Lambda \left\{ \exp \left[\int_{-\infty}^{\infty} (\xi - 1)\big(\Lambda_1(t) + \Lambda_2(t)\big)\, dt \right] \right\}.
\end{aligned}
$$

Thus, $N(t)$ is a doubly stochastic process with rate process $\Lambda(t) = \Lambda_1(t) + \Lambda_2(t)$. □

This result is easily extended to any finite sum of independent DSPPs by repetitive application of the previous theorem, yielding the following corollary.

Corollary 3.6. *Let $N_1(t), N_2(t), \ldots, N_m(t)$ be independent doubly stochastic Poisson processes with (independent) rate processes $\Lambda_1(t), \Lambda_2(t), \ldots, \Lambda_m(t)$, for finite integer m. Then the superposition $N(t) = \sum_{i=1}^{m} N_i(t)$ is a doubly stochastic Poisson process with rate process $\Lambda(t) = \sum_{i=1}^{m} \Lambda_i(t)$.*

Another theorem, concerning the sum of independent fractional Gaussian noises, is therefore useful with regard to the fGnDP-input IF model.

Theorem 3.7. *Let $G_1(k)$ and $G_2(k)$ be two independent standard (i.e. zero mean, unity variance) fractional Gaussian noises, each with identical Hurst index H. Then, their weighted sum $G(k) = \sigma_1 G_1(k) + \sigma_2 G_2(k)$, $\sigma_1, \sigma_2 > 0$, is also a fractional Gaussian noise with Hurst index H, mean zero, and variance $\sigma_1^2 + \sigma_2^2$.*

Proof. Let $Z = \{Z_j, j = \ldots, -1, 0, 1, \ldots\}$ be any stationary sequence. The sequence of transforms $\{T_N, N = 1, 2, 3, \ldots\}$ defined, for each N, as

$$T_N : Z \to T_N Z = \{(T_N Z)_i, i = \ldots, -1, 0, 1, \ldots\},$$

where

$$(T_N Z)_i = \frac{1}{N^H} \sum_{j=iN+1}^{(i+1)N} Z_j, \quad i = \ldots, -1, 0, 1, \ldots,$$

is called the *renormalization group with index H* (Samorodnitsky & Taqqu, 1994, p.338f). T_N transforms the original sequence into a sequence composed of renormalized sums of N adjacent components of the original sequence. According to Corollary 7.2.13 of Samorodnitsky and Taqqu (1994), fractional Gaussian noise is the *only* Gaussian fixed point of the renormalization group, where Z is by definition a fixed point of the renormalization group if $T_N Z \overset{d}{=} Z$ for all $N \geq 1$.

Now let $G_1 = \{G_1(k), k = \ldots, -1, 0, 1, \ldots\}$ and $G_2 = \{G_2(k), k = \ldots, -1, 0, 1, \ldots\}$ be two standard fractional Gaussian noises, each with Hurst index H, and let

$$G = \sigma_1 G_1 + \sigma_2 G_2 = \{\sigma_1 G_1(k) + \sigma_2 G_2(k), k = \ldots, -1, 0, 1, \ldots\},$$

for constants σ_1 and σ_2. Also let $\{T_N, N = 1, 2, 3, \ldots\}$ be the renormalization group with Hurst index H equal to the Hurst indices of G_1 and G_2.

Each T_N is clearly a linear transformation, so, for each $N \geq 1$,

$$T_N G = T_N(\sigma_1 G_1 + \sigma_2 G_2) = \sigma_1(T_N G_1) + \sigma_2(T_N G_2).$$

Since G_1 and G_2 are independent and are both fixed points of the renormalization group,

$$T_N G \overset{d}{=} \sigma_1 G_1 + \sigma_2 G_2 = G.$$

Thus, G is also a fixed point of the renormalization group. Furthermore, for each $k = \ldots, -1, 0, 1, \ldots$, $\sigma_1 G_1(k)$ has a Gaussian distribution with mean zero and variance σ_1^2, and $\sigma_2 G_2(k)$ has a Gaussian distribution with mean zero and variance σ_2^2. Thus, $G(k)$, for each k, also has a Gaussian distribution with mean zero, and its variance is $\sigma_1^2 + \sigma_2^2$. Finally, since G is Gaussian and a fixed point of the renormalization group, it must be fractional Gaussian noise. $\square$

This theorem can also be extended to any finite sum by its repetitive application.

Corollary 3.8. *Let $G_i(k), i = 1, 2, \ldots, m$, be m independent standard fractional Gaussian noises, each with identical Hurst index H. Then, their weighted sum $G(k) = \sum_{i=1}^{m} \sigma_i G_i(k)$, where $\sigma_i > 0$ for $i = 1, 2, \ldots, m$, is also a fractional Gaussian noise with Hurst index H, mean zero, and variance $\sum_{i=1}^{m} \sigma_i^2$.*

Nevertheless, according to (3.19), the rate process of an fGnDP is not a linear function of an fGn. In order to produce a valid rate process, the linear function of fGn

$$\lambda + \sigma G_H \left(\left\lfloor \frac{t}{\tau} \right\rfloor \right)$$

must be truncated below at zero. If this truncation were unnecessary, then, using Corollaries 3.6 and 3.8, the sum of m independent fGnDPs with rate processes

$$\Lambda_i(t) = \lambda + \sigma G_{H,i} \left(\left\lfloor \frac{t}{\tau} \right\rfloor \right), i = 1, 2, \ldots, m,$$

would be an fGnDP with rate process

$$\Lambda(t) = \sum_{i=1}^{m} \Lambda_i(t) = m\lambda + \sqrt{m}\, \sigma G_H \left(\left\lfloor \frac{t}{\tau} \right\rfloor \right). \tag{3.20}$$

Hence, the mean rate of this fGnDP would be $m\lambda$ and the variance $m\sigma^2$.

Statistically, however, it is desirable that the rate process of fGnDP be as much like fGn as possible. The truncation process was only used since the rate of a Poisson process cannot be negative. Thus, ideally, we could consider the superposition of m fGnDPs to be an fGnDP with the rate process (3.20) truncated at zero. In other words, the rate process of the superposition would be

$$\Lambda(t) = \max\left\{0, m\lambda + \sqrt{m}\,\sigma G_H\left(\left\lfloor\frac{t}{\tau}\right\rfloor\right)\right\}. \tag{3.21}$$

In order to compare the difference between this "ideal" situation and the "real" superposition of fGnDPs with rate processes of the form (3.19), consider the sum of two independent fGnDPs. Suppose that the rate processes of these fGnDPs are

$$\Lambda_i(t) = \max\left\{0, \lambda + \sigma G_{H,i}\left(\left\lfloor\frac{t}{\tau}\right\rfloor\right)\right\}, \quad i = 1, 2.$$

Then the rate process of their superposition is

$$\begin{aligned}
\Lambda(t) &= \Lambda_1(t) + \Lambda_2(t) \\
&= \max\left\{0, \lambda + \sigma G_{H,1}\left(\left\lfloor\frac{t}{\tau}\right\rfloor\right)\right\} + \max\left\{0, \lambda + \sigma G_{H,2}\left(\left\lfloor\frac{t}{\tau}\right\rfloor\right)\right\} \\
&= \max\left\{0, \lambda + \sigma G_{H,1}\left(\left\lfloor\frac{t}{\tau}\right\rfloor\right), \lambda + \sigma G_{H,2}\left(\left\lfloor\frac{t}{\tau}\right\rfloor\right), 2\lambda + \sigma(G_{H,1} + G_{H,2})\left(\left\lfloor\frac{t}{\tau}\right\rfloor\right)\right\}, \tag{3.22}
\end{aligned}$$

while the "ideal" situation yields a rate process of

$$\Lambda(t) = 2\lambda + \sqrt{2}\,\sigma G_H\left(\left\lfloor\frac{t}{\tau}\right\rfloor\right). \tag{3.23}$$

Thus, since $G_{H,1} + G_{H,2}$ is distributed like $\sqrt{2}G_H$, the "ideal" case and the "real" case are different when either $\sigma G_{H,1} < -\lambda$ or $\sigma G_{H,2} < -\lambda$, but not when both are true. As we assumed previously, these occurrences should be rare. In addition, when they do occur, the differences that they produce should usually be small, since big differences would necessitate that $\lambda + \sigma G_{H,i} \ll 0$ for the process i that has an untruncated negative rate. Therefore, the superposition of m independent fGnDPs with rate processes of the form (3.19) is well approximated by a single fGnDP with the rate process in (3.21).

3.6.3 Simulation Procedures for the IF Model with fGnDP Inputs

For the IF model with fGnDP inputs, we ran simulations similar to those for the renewal-input model. Again, the simulations were 100,000 seconds in duration, and ten independent simulations were run for each set of parameter values. Each simulation used 100 excitatory inputs and $100r$ inhibitory inputs. The IF mechanism has a reset potential of zero and a threshold of unity, and inputs caused instantaneous increases or decreases of $1/40$ in the integration potential. The rate of each input was calculated via (3.16), yielding a nominal output rate of 2.5 spikes per second.

Each input of this model is specified by its stochastic rate process

$$\Lambda_i(t) = \max\left\{0, \lambda + \sigma G_{H,i}\left(\left\lfloor\frac{t}{\tau}\right\rfloor\right)\right\},$$

under the assumption that they are all statistically identical. The parameter λ, the mean rate of the fGnDP if the effect of truncation is neglected, was specified by the rate of the inputs that produced a nominal output rate of 2.5 spikes per second. In previous studies (unpublished), we found that the value of $\sigma = 30$ worked well for modeling neural spike trains when the rate was $\lambda = 100$. Since the variance of the counts should be additive with respect to the rate, this suggested that $\sigma = 3\sqrt{\lambda}$. Finally, the sample time τ, i.e. the length of the intervals over which the rate of the Poisson processes remained constant, was set to 0.1 seconds. This value was chosen by balancing the cost of simulation time with the condition that this value not have a significant effect on the results.

Therefore, the fGnDP-input model had only two parameters that were left for us to vary, the inhibition-excitation ratio r and the Hurst index H of the inputs. For r, we used the same set of values that was used for the Pareto-input model: 0.0, 0.5, 0.7, 0.8, 0.9, 0.95. We desired to use a set of values for H that spanned the range $0.5 \leq H < 1.0$, which include all values for which fGn is uncorrelated or LRD, but not degenerate. Thus, we chose the values 0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, and 0.95.

For the simulations of the fGnDP-input model, we used the result from Section 3.6.2 that the superposition of independent fGnDPs may be approximated by a single fGnDP. Thus, for each simulation run, we only needed to produce two fGnDPs, one for the excitatory inputs and one for the inhibitory inputs. Both of these fGnDPs had Hurst indices of H, and the rate processes were

$$\Lambda_E(t) = \max\left\{0, 100\lambda + 3\sqrt{100\lambda}\,G_{H,E}\left(\lfloor 10t\rfloor\right)\right\}$$

for the excitatory process and

$$\Lambda_I(t) = \max\left\{0, 100r\lambda + 3\sqrt{100r\lambda}\,G_{H,I}\left(\lfloor 10t\rfloor\right)\right\}$$

for the inhibitory process.

Simulating fGn is not a trivial undertaking, and many different methods have been suggested, some approximate and some exact, beginning with the early suggestions of Mandelbrot and Wallis (1969a) and Mandelbrot (1971). Some of the more common methods include Cholesky decomposition, direct approximation of integral representations, the Durbin-Levinson algorithm, algorithms involving fast Fourier transforms (FFTs), and the random midpoint displacement method. In addition, recent research has also been focused on the production of LRD processes, such as fGn, using wavelet transforms (e.g. Abry, Flandrin, Taqqu, & Veitch, 2000). For a concise overview of these, and other, methods with brief descriptions and discussion of the pros and cons of some of these methods, and references, see Bardet, Lang, Oppenheim, Philippe, and Taqqu (2002).

The random midpoint displacement method (see, e.g., Lau, Erramilli, Wang, & Willinger, 1995), although fast and popular, is quite inaccurate, and approximation of integral representations of fBms can also be problematic. Although it is an obvious and exact method, the Cholesky decomposition method (see, e.g., Beran, 1994) is exceedingly

slow, with a computational complexity of $\mathcal{O}(N^3)$ for a sample of length N. In the past, we have typically used the Durbin-Levinson algorithm (Taqqu, Teverovsky, & Willinger, 1995; Brockwell & Davis, 1991), which is also an exact method, but only requires $\mathcal{O}(N^2)$ operations. Nevertheless, since each of the 1200 simulated fGns (60 sets of parameters $\times$ 10 simulation runs per set $\times$ 2 fGns per run) requires 10^6 samples (10^5 seconds $\times$ 10 samples per second), we still found the computation time for this algorithm prohibitive. FFT methods, however, only have an algorithmic complexity of $\mathcal{O}(N \log_2 N)$, which is an immense improvement over $\mathcal{O}(N^2)$ for very large N. Although most popular FFT methods (e.g. Paxson, 1997; Ledesma & Derong, 2000) are approximate (and, often, quite inexact), an exact FFT method has been developed (Davies & Harte, 1987; Beran, 1994; Bardet et al., 2002). While the approximate FFT methods are attempts to match the frequency spectrum of fGn, the obvious choice for a Fourier transform method, it turns out that applying the FFT in an algorithm meant to match the covariance structure of fGn produces an exact method. This exact method is based upon embedding the covariance matrix in a circulant matrix, but the use of the FFT circumvents any matrix computations. The main drawback FFT methods, however, is that they require large amounts of memory for long samples. We were, however, able to meet these requirements with 1.5 gigabytes of RAM, and thus used the exact FFT method to produce fGns in the simulations of the fGnDP-input model.

After simulating the fGnDP-input IF model with all combinations of the values of the two parameters r and α, we calculated estimates of the CV_{ISI}, the FFC, and the IDC for each simulation run in the same manner as was done for the renewal-input model. This included estimating the FFC and IDC of the surrogate data, produced by shuffling the interspike intervals, for each simulation run.

3.6.4   Simulation Results for the IF Model with fGnDP Inputs

Figure 3.8 shows the estimated values of CV_{ISI} from all of our simulations of the fGnDP-input model plotted versus the inhibition-excitation ratio r. Column (a) contains graphs of the mean CV_{ISI}'s calculated across the ten simulations at each combination of parameter values for r and H. Means for the same value of H are connected by lines. Like the Gaussian-input model and the Pareto-input model, the graph of CV_{ISI} as a function of r for any particular value of H has an increasing, convex shape. More specifically, these curves are well-approximated by the square root of a rational function with a pole at $r = 1$. Furthermore, the values of CV_{ISI} all lie above the curve for the Poisson-input model (dashed line), and they increase with H. Since, like the Pareto-input model, the intervals in the superpositions of the inputs are positively correlated for this model, the fact that CV_{ISI} is always greater (at least for $0.5 \leq H < 1$) than in the Poisson-input model is to be expected. The positive correlation between CV_{ISI} and H is also not surprising, given the results for the Pareto-input model. In the Pareto-input model, decreasing the value of α strengthened the dependence between intervals in the output, and, we suspect, the input-superpositions, of the model. This was also associated with increases in the value of CV_{ISI}. In the present model, increasing the value of H will certainly strengthen the dependency between intervals in the input-superpositions. Thus, we should expect that such increases in H will lead to stronger dependence in the output

as well, and that this be associated with increases in the value of CV_{ISI}. In the IDCs below, we will see that increases in H do indeed strengthen the dependence between the output intervals. These results thus indicate that for higher values of H the fGnDP-input model requires less excitation-inhibition balance to achieve any particular value of CV_{ISI}. As for the Pareto-input model, this means that, for high values of H, the CV_{ISI} of the output of the fGnDP-input model is substantially larger than physiological estimates at even moderate degrees of balance.

Column (b) of Figure 3.8 shows, for a subset of the values of H, the individual estimates of CV_{ISI} for each simulation. The lines again connect the mean value of these estimates. Thus, like Figures 3.3 and 3.5, there are ten symbols, one for each simulation run, for each of the different combinations of the two parameters r and H. Here again the symbols are tightly grouped for low values of r and weaker LRD, i.e. low values of H. Also, in similarity to the other two sets of simulations, the variance of estimates of CV_{ISI} increases as either r or H is increased. Again, this is a result of the fact that the model approaches a non-stationary state as r or H approach the value of one.

Figure 3.9 shows the FFCs and IDCs for simulations of the IF model with fGnDP inputs for several representative combinations of the values for the parameters H and r. The curves for Hurst indices of $H = 0.5$, $H = 0.7$, and $H = 0.9$ are shown in 3.9a, 3.9b, and 3.9c, respectively. For each of these values of the Hurst index, FFC/IDC sets are shown for the inhibition-excitation balance ratios $r = 0.0$, $r = 0.5$, $r = 0.8$, and $r = 0.95$. The entire set of FFCs and IDCs from all of our simulations are shown in Appendix B.3.

For these curves, as should always be the case, the FFCs approach a value of one at very small counting interval lengths, and the mean surrogate-data IDCs are horizontal. Due to the lack of very short intervals in the output of this model, caused by the integration mechanism, the initial trend of the FFCs is downward. These initial downward trends are comparable to those seen in the FFCs of the Gaussian-input and Pareto-input models. In contrast, the original-data IDCs for the fGnDP-input model, when $H > 0.5$, have an initial, and indeed enduring, upward trend. In fact, these IDCs all closely resemble a power-law function, being very linear on a double logarithmic plot. Like the Pareto-input model, but in contradistinction to the Gaussian-input model, a superposition of the inputs to the fGnDP model has positively correlated interspike intervals. Not only does this positive correlation produce the positive slopes of the IDCs, but it also produces positive slopes in the original-data FFCs for medium and large counting intervals. Thus, except when H equals 0.5 or when H is close to 0.5 and r is small, the original-data FFCs have an asymptotic value larger than one.

Theoretically, the output of the fGnDP-input model with $H = 0.5$ is a renewal process. fGn is simply the common "white" Gaussian noise when $H = 0.5$. Thus, since this fGn has no memory and Poisson processes have no memory, the fGnDPs used as inputs to the IF model will have no memory when $H = 0.5$. Thus, in this sense, it is very similar to the Poisson-input IF model, with the integration mechanism, which is reset at the occurrence of each output spike, being the only component possessing memory. The results of our simulations are in agreement with this prediction. As is evident in Figure 3.9a, the original-data FFCs and IDCs are nearly identical to the surrogate-data ones. This is the behavior that we expect from a renewal process. Due to the presence of the fGn, each input to this model should, however, be more variable than a Poisson process, which
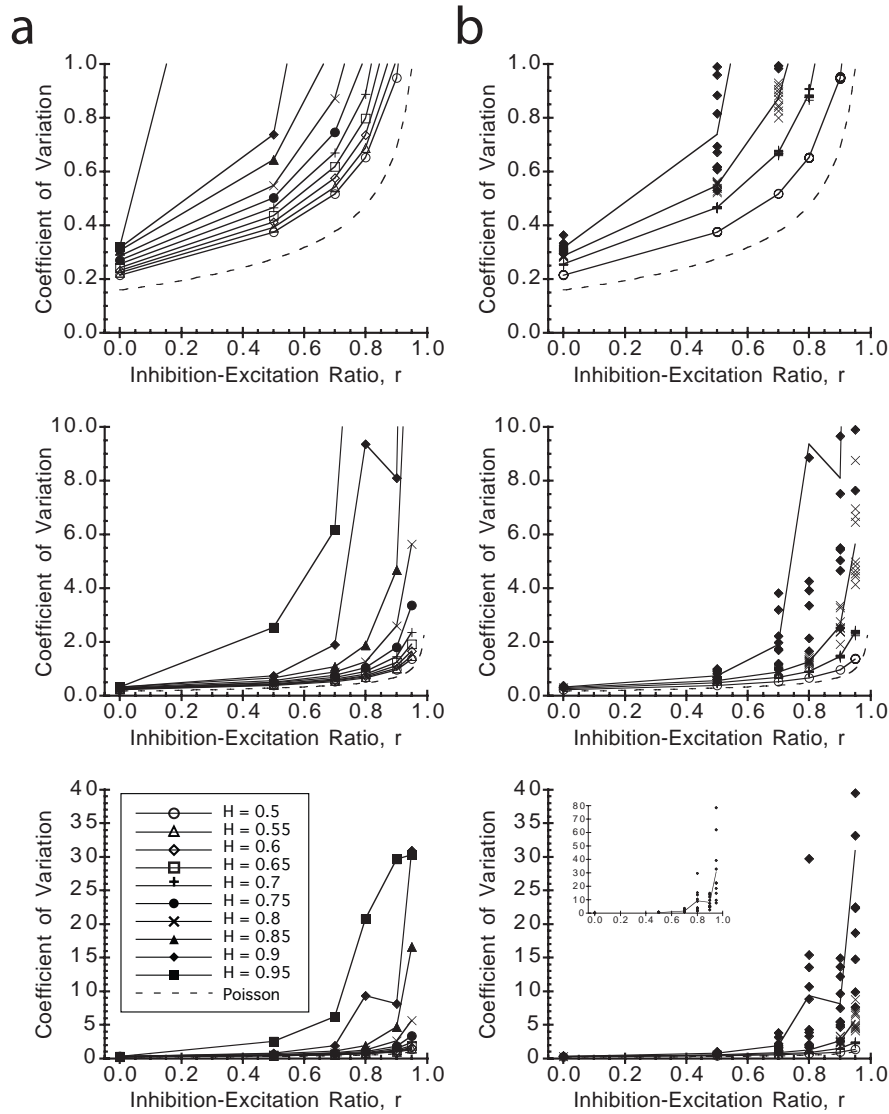
Figure 3.8: The coefficient of variation of interpoint intervals as a function of the inhibition-excitation ratio r for the output of an integrate-and-fire model with inputs that are fractional-Gaussian-noise-driven Poisson processes. For each combination of the parameters, r and the Hurst index H of the fractional Gaussian noise, ten separate simulation runs were conducted. In column (a), the mean values of the estimates are plotted for each combination of parameters. Each solid line connects the means for a single value of H, and the dashed line is the theoretical curve for the model with Poisson process inputs. All three sets of axes contain the same data, but each displays them on a different vertical scale. In column (b), the estimates calculated from each individual simulation run are plotted for a subset of the values of H. Solid and dashed lines are the same as in column (a), and all three large sets of axes contain the same data. The inset on the bottom set of axes contains only the data for $H = 0.9$, revealing two additional data points that are out of range of the larger set of axes.
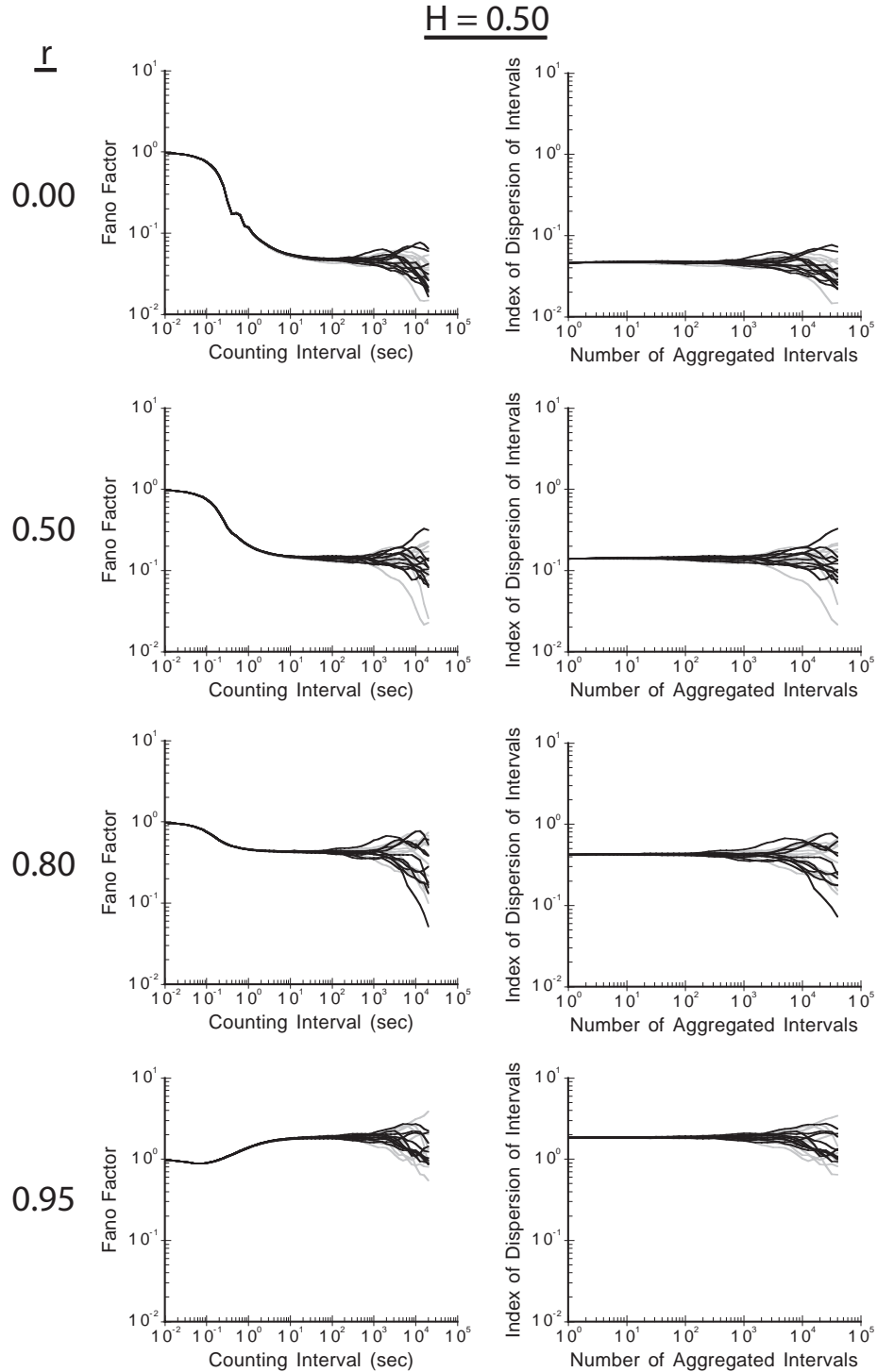
Figure 3.9a: FFCs and IDCs estimated from simulations of the IF model with fGnDP inputs with Hurst index $H = 0.5$. Each set of axes contains ten curves calculated from original data (black) and ten curves calculated from the corresponding shuffled surrogate data (gray). For each value of the inhibition-excitation ration r, each individual FFC in the left set of axes was calculated from the same data as one of the IDCs in the right set of axes.

H = 0.70

r

0.00

0.50

0.80

0.95
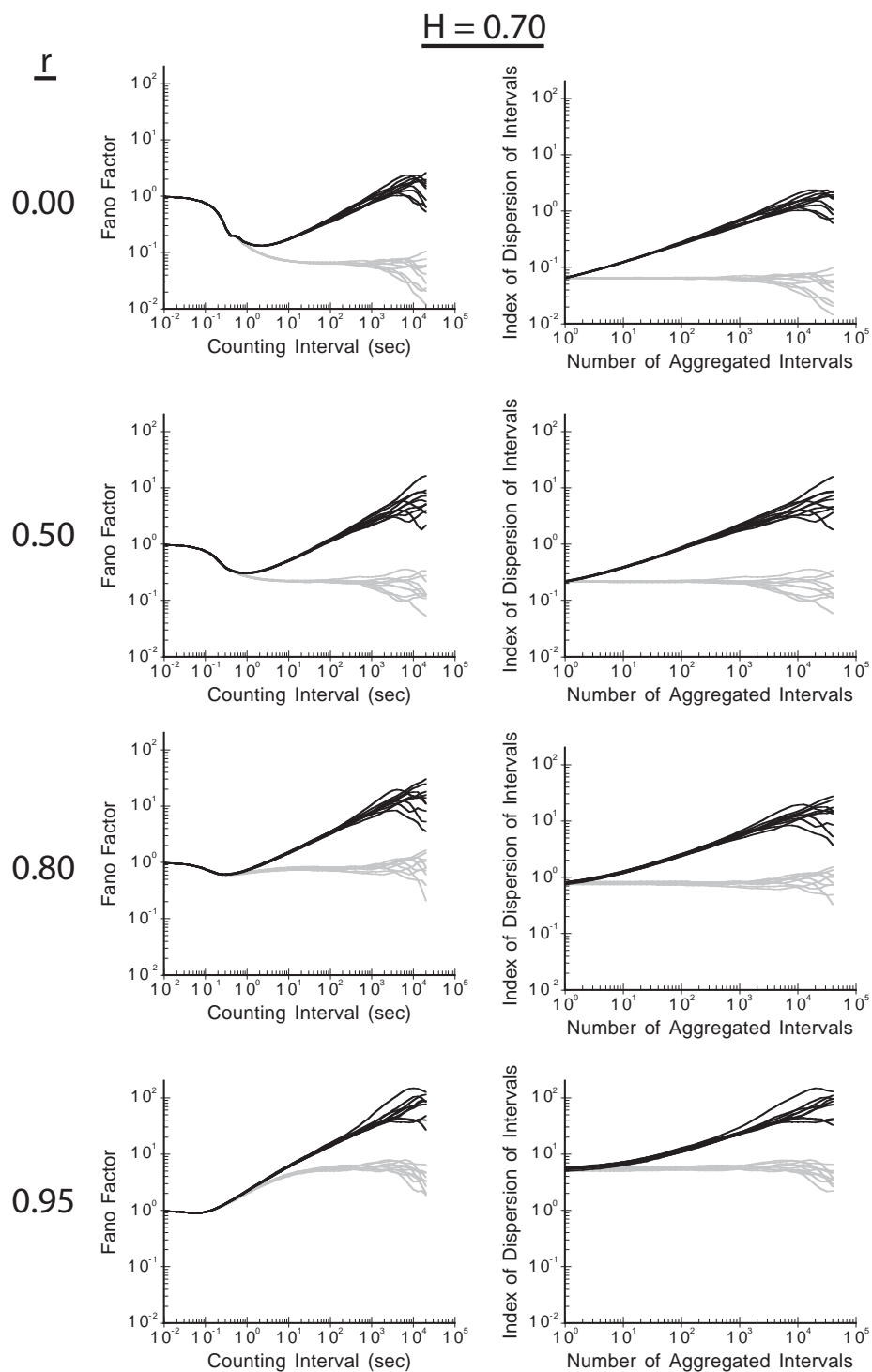
Figure 3.9b: Same as in 3.9a, except for $H = 0.7$.
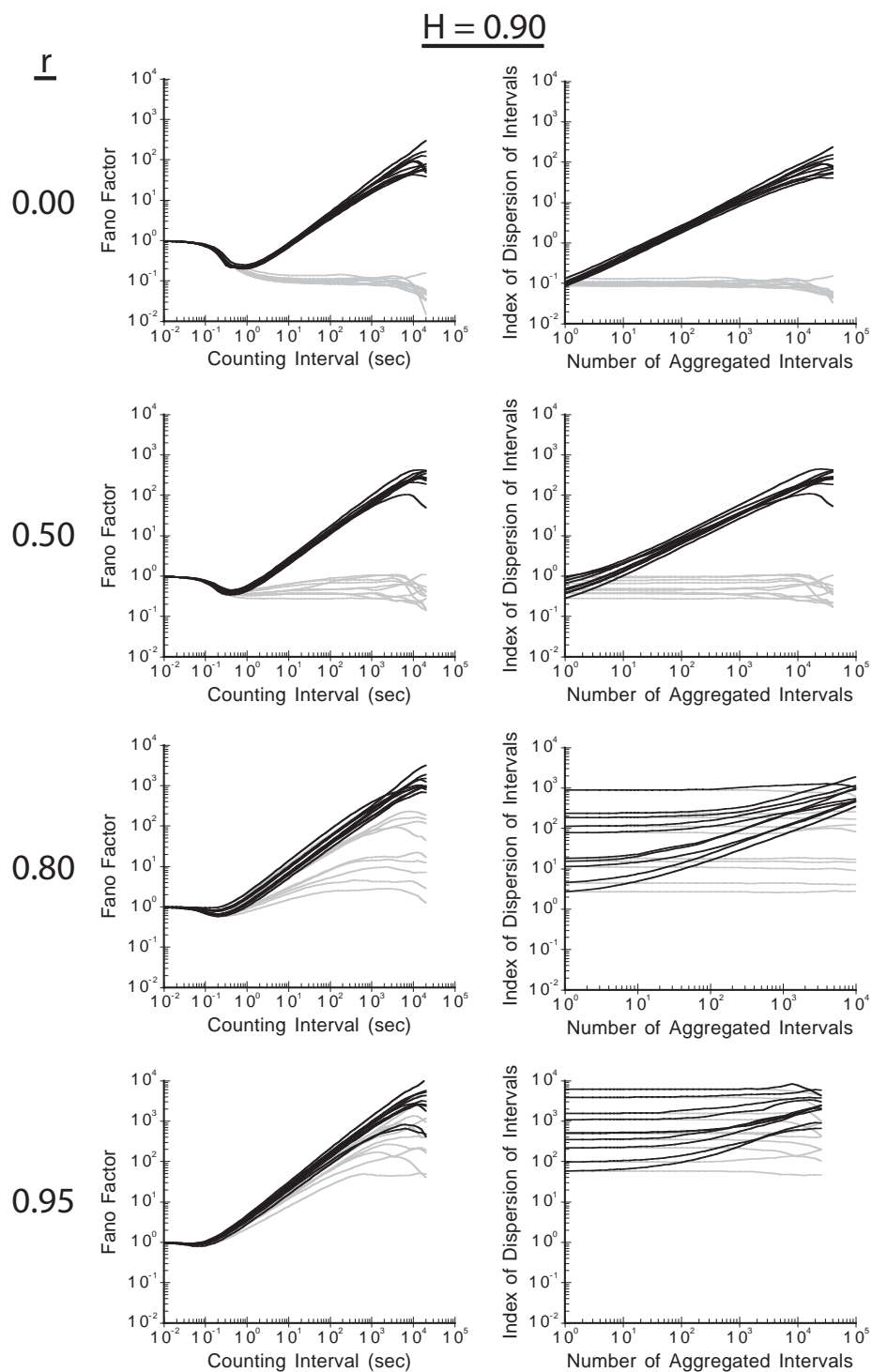
Figure 3.9c: Same as in 3.9a, except for $H = 0.9$.

should manifest itself in higher variability at the output of the model. Indeed, as we saw in Figure 3.8, the variance of the output of the fGnDP-input model for $H \geq 0.5$ is always greater than that for the Poisson-input model having the same r-value.

When $H > 0.5$, the inputs of the fGnDP model, as well as the superpositions of the inputs, are LRD. This was also true of the Pareto-input model when $\alpha \leq 2.0$. The LRD in the Pareto inputs and superpositions, however, came in the form of LRcD with infinite interval-variance and no LRiD, whereas in the fGnDP inputs and superpositions it is in the form of LRcD with LRiD and finite interval-variance. But this distinction does not seem to be important for the IF model, at least with regard to the statistical procedures that we have used. This is demonstrated in the striking similarity between the results from the Pareto model with $\alpha = 1.75$ in Figure 3.6b and those from the fGnDP model with $H = 0.7$ in Figure 3.9b. Thus, disregarding the effect of the interaction between excitation and inhibition, in the fGnDP-input model the LRiD at the input propagates through the model, while in the Pareto-input model the infinite interval-variance of the inputs is converted into LRiD by the model. In either case, the result at the output seems to be just about the same.

As we saw in Section 3.5.3, the potential in the Pareto-input IF model is in some ways similar to the aggregation of *ON/OFF* source models. At a minimum, both of these systems convert the infinite variance of intervals between "points" into LRD in a semi-continuous time series. Since the aggregation of infinite-variance *ON/OFF* sources asymptotically approaches fBm (Taqqu & Levy, 1986; Willinger et al., 1995, 1997; Taqqu et al., 1997) , the "integral" of fGn, the correspondence between the Pareto-input model and the fGnDP-input model is not unusual. In fact, Lowen and others have developed a DSPP model that is driven by fractal binomial noise, which is the equivalent of the aggregation of many infinite-variance *ON/OFF* sources, to model both neural activity (Lowen & Teich, 1993a, 1995; Thurner et al., 1997) and network traffic (Ryu & Lowen, 1996, 1997; Ryu, 1997; Ryu & Lowen, 1998, 2000). Predictably, this fractal-binomial-noise-driven Poisson process asymptotically approaches an fGnDP. In a neuron, the correlate of the *ON/OFF* sources may be ion channels (Lowen & Teich, 1993b, 1993d, 1993c; Lowen, Liebovitch, & White, 1999; DeFelice & Isaac, 1993), since it is suspected that the durations of their opened and closed states are long-tailed (see, e.g., Liebovitch & Sullivan, 1987; Liebovitch & Toth, 1990a, 1990b; Liebovitch & Koniarek, 1992; Liebovitch, 1996; Liebovitch & Todorov, 1996; Liebovitch, Scheurle, Rusek, & Zochowski, 2001).

The effect on the output of the fgnDP-input model of balancing the amounts of excitation and inhibition is comparable to its effect on the other models that we have considered. As r increases toward one, the effect of the correlation in the inputs is gradually overwhelmed by the high variability of the excitation-inhibition interaction. The output will continue to exhibit LRcD, but it will progressively become more a result of high interval variability rather than of LRiD. However, the surrogate-data FFCs all, except perhaps when H and r are both very close to one where the approach to non-stationarity creates unusually large variance in our estimates, asymptote to a finite constant, consistent with the intervals in the output having finite variance. The output interval-variance does increase with r, causing the asymptotic constant of the surrogate-data FFCs to move upwards, but this FFC always remains below its original-data counterpart because of the presence of LRiD. Also, in a manner analogous to that in the Pareto-input model, the

difference between the original-data FFCs and the surrogate-data FFCs at long counting intervals increases as H increases. But, this trend again seems to break down, in this case as H approaches one.

Pareto RPPs are non-LRD for any α greater than two, but fGnDPs are only non-LRD when $H = 0.5$. In addition, the Pareto RPPs will produce positive correlation between the intervals of the output for any α, including $\alpha > 2$, while intervals in the output of the fGnDP model with $H = 0.5$ are independent. Thus, these two models behave quite differently in their non-LRD state. The Pareto model is more flexible, allowing a range of different FFC shapes and differences between the original-data curves and the surrogate-data curves. On the other hand, only the asymptotic value of the FFCs can be adjusted, by changing the value of r, for the fGnDP model, and original-data and surrogate-data FFCs are always identical. Therefore, the fGnDP with $H = 0.5$ is not even an improvement over the Poisson-input model, even for non-LRD data. The fGnDP model is, however, more flexible than the Pareto model in producing weaker short-term dependence when LRD is present in the output. Due to the positive correlation that is present in the non-LRD Pareto model, the original-data FFCs and IDCs are already significantly different than their surrogate-data counterparts on shorter time scales when $\alpha = 2.0$, the first value at which it is LRD. In contrast, the original-data curves for the fGnDP gradually differentiate themselves, at all time scales, from the surrogate-data curves as LRD is introduced into the output. Thus, although these two models can produce outputs with similar statistical features, the statistical characteristics of their outputs are adjustable in different ways.

In short, then, our simulations show that, with proper adjustment of the parameters r and H, the fGnDP-input model can produce outputs that are like cortical spike trains in that they possess both LRD and intervals with finite variance. The typical range of values for CV_{ISI} estimates in cortical spike trains is 0.5 to 1.5. Table 3.3 shows the approximate range of the inhibition-excitation ratio r, for each value of H in our simulations, that produces values for the CV_{ISI} in this range. Using this data and comparisons of the FFCs from our simulations, some of which are shown in Figure 3.9, with those shown in Teich et al. (1996), the ability of the fGnDP-input IF model to match statistical properties of cortical spike trains can be more directly evaluated.

For low values of H, the value of CV_{ISI} for the fGnDP model is within the physio-logical range for r between about 0.6 and 0.9. This is very similar to the Pareto model at α-values below, but close to, 2.0. For r values near 0.9, the FFCs for $H \approx 0.6$ resemble those of Cell 4 and Cell 7 in Teich et al. (1996). For the Pareto model, however, such FFCs were created when α was greater than 2.0. The critical difference, however, is that the $H \approx 0.6$ fGnDP model is LRD, while the $\alpha > 2.0$ Pareto model is not. However, since the slope of the FFCs is so shallow for the $H \approx 0.6$ fGnDP model, the LRD in the output is difficult to distinguish using an FFC calculated from any reasonable length sample of the process.

For H in the neighborhood of 0.65 to 0.85, the upper limit of r-values needed to produce physiological CV_{ISI}'s is $0.8 - 0.9$. For this combination of parameters, the FFCs of the fGnDP model resemble those of Cell 3 and Cell 19 of Teich et al. (1996). These FFCs suggest the presence of LRcD, but the surrogate-data FFCs asymptote to a finite value above one. If r is reduced, then this asymptotic value drops below one, which does

H	Range of r
0.5	$0.69 < r < 0.96$
0.55	$0.66 < r < 0.95$
0.6	$0.63 < r < 0.94$
0.65	$0.59 < r < 0.92$
0.7	$0.55 < r < 0.90$
0.75	$0.50 < r < 0.87$
0.8	$0.45 < r < 0.81$
0.85	$0.32 < r < 0.77$
0.9	$0.25 < r < 0.64$
0.95	$0.06 < r < 0.28$

Table 3.3: The range of the inhibition-excitation ratio r, for each value of H, that leads to values of CV_{ISI} between 0.5 and 1.5 for the integrate-and-fire model with fGnDP inputs. The ranges of r were approximated from plots of CV_{ISI} versus r (see Figure 3.8) obtained from simulations of the model.

not match the results in Teich et al. (1996), although the LRcD is still present. This case may, however, match the FFCs of neurons with $CV_{ISI} < 1.0$ if they were available. Similar to those of the Pareto model, as r is increased, the FFCs still resemble the physiologically measured FFCs, but the values of CV_{ISI} become too large.

As H increases to about 0.9 and above, the FFCs and CV_{ISI}'s can no longer be matched to those measured physiologically. When these models have physiological values of CV_{ISI}, they are still LRcD, but the surrogate-data FFCs asymptote at too low of a value. Like the Pareto model with $\alpha \approx 1.0$, moderately large values of r tend to result in FFCs that suggest that the output of the model is renewal-like. Here, again, this is most likely due to the models approach to non-stationarity, which occurs at $H = 1$, with the effect of tightly balanced excitation and inhibition. Also, much greater variability in the estimates of the Fano factor and the index of dispersion of intervals are evident in this parameter range.

Therefore, like the Pareto-input model, the fGnDP-input model can produce outputs that share common statistical features with the spike trains in cortical neurons. However, whereas the inputs to the Pareto model do not seem to be justified physiologically, the inputs to the fGnDP are known to be statistically similar to the spike trains of neurons that project into cortex.

3.7 Conclusion

In Chapter 2, we were able to show that a very large class of models meant to explain the high variability of cortical spike trains was incapable of also producing LRD, another statistical feature of cortical spike trains. Here we have considered another group of highly variable cortical models, as well as suggesting a new type of model, and have studied them through the use of simulations using many different combinations of parameter values.

The first group of models consists of an integrate-and-fire mechanism with inputs that are renewal point processes. This model is capable of producing correlation between the intervals of the superpositions of inputs that then causes the intervals in the output of the model to be similarly correlated. If the intervals of the inputs are distributed according to a positive-Gaussian distribution, this correlation is negative. We expect this result to be true of any interval distribution that possesses a shorter tail than the exponential distribution, the interval distribution of a Poisson process. Although the Gaussian-input model can produce any value of CV_{ISI} if the inhibition-excitation ratio r is properly adjusted, the negative correlation created by these inputs hinder the production of LRD. Thus, in some sense, the Gaussian-input model is inferior to the standard Poisson-input model.

If the intervals of the RPP inputs are distributed according to a Pareto distribution, then positive correlation is produced in the IF model. We expect this to also be true of any interval distribution with a tail that is longer than the exponential distribution. Furthermore, if the tail of the distribution is long enough that the variance is infinite, then both the inputs and the output of the model are LRD. However, the input RPPs will be LRcD, but not LRiD, while the output will be both LRcD and LRiD. This "conversion", so to speak, is produced by the integration process in the model, and has precedent in other models that aggregate processes with infinite variance. We found that by proper adjustment of α, a parameter of the Pareto distribution that affects the length of its tail, and the inhibition-excitation ratio r, we could produce spike trains with the Pareto-input IF model that had interval variance and LRD similar to spike trains recorded from cortical neurons. The main drawback to this model was that its inputs do not seem physiologically justified, since they have intervals that have infinite variance but are not correlated.

We therefore suggested a model with inputs that are LRD, but are statistically similar to actual neurons. The fractional-Gaussian-noise-driven Poisson process has elsewhere been used as a model for sensory neurons in the periphery, such as primary auditory neurons, to model the LRD present in their spike trains. The same statistical characteristics present in these neurons are known to exist in many sub-cortical sensory pathways, including certain neurons that project into the cortex. Thus, this process seems to be a reasonable choice for the inputs to the IF model of cortical neurons. We found that, as for the Pareto RPP inputs, the fGnDP inputs did indeed produce LRD in the output of the IF model. However, in the fGnDP both the inputs and the outputs had LRcD and LRiD, and neither had intervals with infinite variance. By proper adjustment of the Hurst index H, a parameter that modulates the strength of the LRD present in an fGnDP, and the inhibition-excitation ratio r, we were able to produce output spike trains with the fGnDP-input IF model that approximated the LRD and variability of cortical spike trains. Although the output of this model is not necessarily more similar to that of cortical neurons than the Pareto-input IF model, the physiological justification of its fGnDP inputs is reason enough to favor it over the Pareto model.

The one class of models of high variability in cortical neurons that is not covered by the work in this chapter and Chapter 2 is the group of models that produces variability through network dynamics. These models consist of large networks of interconnected integrate-and-fire neurons. Usher and his colleagues (Usher et al., 1994, 1995) demonstrated that the spike trains in individual units of their version of this model have power

spectrums that approximate a power-law at low frequencies, a characteristic that is essentially equivalent to having LRD. (See Section 2.3 for further details regarding the relationship between the power spectrum and long-range dependence.) However, although they can find LRcD, their analyses are not capable of distinguishing between LRcD produced by high interval-variance and LRcD produced by LRiD. Thus, in future work, we would like to apply the analysis methods used in this chapter to simulations of their model.

Chapter 4

Estimating the Moments of Long-Range Dependent Spike Trains

ABSTRACT Sample statistics for spike counts are commonly used to characterize neurons or to assess their ability to carry information about the stimulus. However, use of these statistics is rarely associated with an evaluation of the expected quality of the estimates obtained from them. Relatively recently, long-range dependence has been observed in many neurons. In other types of stochastic processes, long-range dependence is known to have adverse effects upon such sample statistics. Therefore, in this chapter, we attempt to gauge the reduction in quality, caused by long-range dependence, of three common statistical estimators: the sample mean, the sample variance, and the sample standard deviation. In particular, we examine their expected value and variability for the fractional-Gaussian-noise-driven Poisson process. Other experimental studies have demonstrated that this process is a good model of long-range dependence in the spike trains of many neurons. We conclude that the variability of the sample mean is likely to be much larger in the presence of long-range dependence than would be expected otherwise. Furthermore, in the presence of long-range dependence, the sample variance and the sample standard deviation underestimate the actual variance and standard deviation of spike counts. We discuss the implications of these findings for common experimental methods and results.

4.1   Introduction

The nervous system is able to quickly encode and process a remarkable amount of information about external stimuli. At present, no semblance of a general theory that explains how this is accomplished exists. In particular, although significant progress has been made in the study of some specialized neural systems, the meaning of the activity in sensory neurons, as understood by other neurons, is largely a mystery. From another perspective, this means that the properties of this neural activity that convey information to other neurons have not been fully identified.

The output activity of most neurons consists of a temporal sequence of discrete events that are more or less identical. Each of these events is called an action potential, or, colloquially, a spike, and is essentially pulsatile in nature. The theory or assumption that neural information is encoded in the rate of occurrence of action potentials, or the "firing rate", is by far the most pervasive explanation of neural coding and processing. One of the underlying implications of the theory of rate coding is that the trial to trial variability that is observed in neural responses to presumably identical experimental conditions is considered by the system to be noise. Thus, this "noise" places limits on the ability of the nervous system to detect and categorize external stimuli. Some of the noise may, of course, be removed by averaging, but even that which can be averaged out is taxing the system by requiring the use of some of its resources. Therefore, the form of this noise and its statistical properties have significant implications for the structure of information content within neural activity and the capabilities and limits of neural processing.

In estimating and representing the stochastic nature of neural activity within the context of rate-coding theory, researchers are wont to estimate the mean and variance of the firing rate, or, equivalently, of the count in a fixed-length interval of time. However, although so common as to usually not incur scrutiny, these seemingly innocuous measurements can be quite misleading if their underlying assumptions are violated. For each estimator that is used to measure a moment of the underlying process, its reliability requires that the following be true: (i) the value of the estimator is on average equal to the value of the moment that it is estimating; (ii) as the number of observations increases, the value of the estimator converges (in a sense that will be made explicit later) to the value of the actual moment; (iii) and this convergence occurs fast enough that the value of the estimator calculated from a limited set of observations is a reasonably good approximation to the value of the actual moment. Of course, these somewhat informal prescriptions have formal mathematical complements that, in turn, place restrictions on the underlying process that is producing the observations used in the calculation of an estimate. For the standard estimators used for the mean and variance, the validity of the estimators is predicated on the following: (i) the mean and variance of the underlying process exist and are finite; (ii) the observations are taken from a stationary process and are thus identically distributed; (iii) and the observations are statistically independent. Furthermore, determination of the rate of convergence of these estimators is commonly established on the basis of the supposition that the observations have a Gaussian (normal) distribution.

Rarely, in practice, are the requirements of these presuppositions met. Often, however, deviations from these ideals are small, in the sense that they have limited effects on the results. But this is not always the case. In this paper we will investigate the validity of these assumptions in relation to estimating the mean and variance of neural spike counts and firing rates. In particular, we will consider them in relation to the long-range autocorrelations that seem to often be present in neural spike trains. In this context, the property of possessing long-range autocorrelations has been referred to as "fractal behavior" or as possessing "fractal patterns", but it is more generally (and formally) known as having long memory or long-range dependence. Since examples where the presence of long-range dependence frustrates moment estimation and related statistical procedures in standard stochastic processes, such as stochastic time series, are plentiful (see, for example, Beran, 1994), the questions that we shall raise have significant bearing

on the validity of moment estimation for neural firing rates and the adequacy of the theory of rate coding.

4.2 Long-Range Dependence: Theory and its Relevance to Neural Activity

4.2.1 Short-Range and Long-Range Dependence

Classical statistical methodologies have primarily been developed and validated for short-range dependent processes. A stationary stochastic process $\{X_i : i \in \mathbb{Z}\}$ is said to have short-range dependence if it possesses the following essentially equivalent properties (Cox, 1984):

(i) $\displaystyle\lim_{n\to\infty} \sum_{j=0}^{n} Cov\{X_0, X_j\} = C_1;$

(ii) $S_X(0) = C_2$, where $S_X(\omega)$ is the spectral density (or power spectrum) of $\{X_i\}$;

(iii) $\dfrac{Var\left\{\sum_{j=1}^{n} X_j\right\}}{n} \sim C_3$ as $n \to \infty.$

where the C_k's are some finite constants. Processes that have no dependence also fall into this class since, for example, $C_1 = 0$ for such a process. The critical notion denoted by the term "short-range dependence" is that the correlations at least decay quickly as the separation between the random variables increases, not that these correlations necessarily exist.

According to the properties above, a short-range dependent process has (i) quickly decaying autocorrelations and (ii) a spectral density that is finite-valued at zero frequency. Also, according to (iii), asymptotically, the variance of its sample mean decays as one over the number of sample values or faster.

The negations of the three properties for short-range dependent processes above yield a description of long-range dependent processes. Thus, for a stationary stochastic process $\{X_i : i \in \mathbb{Z}\}$ that is long-range dependent (Cox, 1984),

(i') $\displaystyle\lim_{n\to\infty} \sum_{j=0}^{n} Cov\{X_0, X_j\} = \infty;$

(ii') $\displaystyle\lim_{\omega\to 0} S_X(\omega) = \infty;$

(iii') $\displaystyle\lim_{n\to\infty} \dfrac{Var\left\{\sum_{j=1}^{n} X_j\right\}}{n} = \infty.$

Hence, (i') its autocorrelations decay slowly with separation of the random variables, and (ii') its spectral density has a pole at zero frequency. Furthermore, (iii') the variance of its sample mean decays more slowly than one over the number of sample values.

The concept of long-range dependence is important because, over the last half-century, its presence has been detected in many natural and physical processes. Hence, a growing number of models for long-range dependence have been developed over that time as well. And, since it can have significant and adverse consequences on statistics developed for short-range dependent processes (see, e.g., Beran, 1994; Cox, 1984), the presence of long-range dependence is important to identify in data.

Any of the three previous properties can be applied as a definition of long-range dependence. Property (i') seems to be the obvious choice, since it refers directly to the typical measure of dependence: covariance. Yet, in practice direct estimates of the covariance tend to be noisy. This is especially problematic when the covariances are small, which, perhaps contrary to intuition, can be the case for long-range dependent processes. Long-range dependence is an asymptotic property that is the result of the cumulative effect of many correlations, not the result of a single or small set of large correlations.

The variance of property (iii') can usually be measured more reliably than covariances. The reason for this, as well as the connection between property (iii') and property (i'), is easily seen in the basic identity

$$Var\left\{\sum_{j=1}^{n} X_j\right\} = \sum_{j=1}^{n} Var\left\{X_j\right\} + 2\sum_{i<j} Cov\left\{X_i, X_j\right\}. \tag{4.1}$$

Thus, this variance is associated with the *sum* of covariances, rendering its estimates more stable than those of individual covariances. Finally, property (iii') is to be preferred over property (ii') due to the greater ease with which it is estimated and understood.

Thus, we will use the following definition for long-range dependence in general stationary stochastic processes.

Definition 4.1. A stationary stochastic process $\{X_i : i \in \mathbb{Z}\}$, for which X_i has finite variance for all i, exhibits long-range dependence (LRD) when

$$\limsup_{n\to\infty} \frac{Var\left\{\sum_{i=1}^{n} X_i\right\}}{n} = \infty.$$

For this definition to be valid, the stochastic process must have finite variance so that the second-order properties upon which long-range dependence is based are well-defined. Although infinite variance often occurs in conjunction with long-range dependence, as will become apparent in following sections, this situation is not relevant in the context of the point processes that we will be considering.

4.2.2 Long-range Dependence in Point Processes

In order to define long-range dependence in point processes, we must relate Definition 4.1 to properties that are particular to point processes. We will only be considering one-dimensional stochastic point processes on the real line, which will represent time. These point processes can be characterized by either their interpoint distances or the numbers of points in any arbitrary set of intervals on the real line. Thus, there are actually two types

of long-range dependence that can occur in point processes, one associated with interpoint intervals and one associated with counts.

Since the interpoint intervals of a point process form a stochastic process, long-range dependence in these intervals is easily defined in a way analogous to Definition 4.1.

Definition 4.2. *(Daley & Vesilo, 1997)* [1] A stationary point process $N(\cdot)$ on the real line exhibits long-range interval dependence (LRiD) when the stationary sequence of interpoint intervals $\{Y_i\}$, with finite variances, is LRD in the sense that

$$\limsup_{n\to\infty} \frac{Var\left\{\sum_{i=1}^{n} Y_i\right\}}{n} = \infty.$$

In order to apply Definition 4.1 to the counts of a stochastic point process, we might consider the stochastic process formed by the counts in a sequence of T-length intervals for some fixed $0 < T < \infty$, e.g. the sequence with elements

$$N_i(T) = N\big((i-1)T, iT\big],$$

for $i = 1, 2, \ldots$. Thus, the definition might look something like

$$\limsup_{n\to\infty} \frac{Var\left\{\sum_{i=1}^{n} N_i(T)\right\}}{n} = \infty,$$

which begs the question of how to choose the value of T. However,

$$\sum_{i=1}^{n} N_i(T) = \sum_{i=1}^{n} N\big((i-1)T, iT\big] = N(0, nT].$$

Thus, using this relation, substituting $t = nT$, and recognizing that the multiplication or division of an expression by a finite, positive constant does not change whether or not its limit is infinite, a definition without arbitrary variables such as T is obtained.

Definition 4.3. *(Daley & Vesilo, 1997)* A second-order stationary point process $N(\cdot)$ on the real line exhibits long-range count dependence (LRcD) when

$$\limsup_{t\to\infty} \frac{Var\left\{N(0, t]\right\}}{t} = \infty.$$

The discussion in Section 3.3 suggests that LRcD is the "general" LRD property of point processes, and that the statistical origin of LRcD in any particular point process may reside in either LRiD or the infinite variance of interpoint intervals or in both of these.

---

[1]The original definition of Daley and Vesilo (1997) does not contain the phrase "with finite variances", though the finiteness of the variances of the intervals is necessary and may have been implied.

4.2.3 Long-Range Dependence in Neural Spike Trains

As summarized in Section 2.3, LRD has been observed in the spike trains of many different neural systems. In these studies the spike trains are often said to exhibit other properties, such as fractalness, self-similarity, or 1/f fluctuations, but these properties imply the presence of LRD. LRD has been investigated primarily in the sub-cortical and cortical sensory systems of mammals, although it has also been found in the visual system of certain insects (Turcott et al., 1995). But, in practically all instances where analysis of the spike trains was sensitive to LRD-like properties, it provided evidence of the existence of LRD. The one exception is the peripheral vestibular system (Teich, 1989), which has very different statistical properties than most sensory systems. This suggests that LRD may be a common property of neural spike trains.

The sub-cortical locations where LRD has been identified include the primary auditory nerve (Teich, 1989; Teich & Lowen, 1994; Lowen & Teich, 1996b), the lateral superior olive (Teich et al., 1990; Turcott et al., 1994), the retina and lateral geniculate nucleus (Teich, 1996; Teich et al., 1997; Lowen et al., 2001), and the ventrobasal neurons of the thalamus (Kodama et al., 1989). In all, but the last, the analysis of the spike trains revealed that the LRcD in the spike trains was due to correlations in the sequence of interspike intervals and that the variance of these intervals was lower than that expected of a Poisson process. The study of Kodama et al. (1989) neither supports nor refutes this conclusion.

In the cortex, LRD has been found in primary visual cortex (Teich et al., 1996), somatosensory cortex (Wise, 1981), the mesencephalic reticular formation (Yamamoto et al., 1986; Gruneis et al., 1989, 1990, 1993), and the hippocampus (Mushiake et al., 1988; Kodama et al., 1989). The study in the primary visual cortex revealed that the LRcD present in its neural spike trains was also due to LRiD, but that the variance of the interspike intervals was greater than that of a Poisson process, though still finite. The remaining cortical studies did not produce any results relevant to the question of the statistical origin of the LRcD.

4.3 Statistical Theory

4.3.1 The Distinction between Population Moments and Sample Moments

Often, in applied settings, statistics become conflated with the population moments that they estimate. To draw the distinction, consider the "random machine" shown in Figure 4.1. A stochastic system is composed of a particle floating freely in a pool. All possible states of the system form a set, called the sample space Ω. The precise state of the system at any time, $\omega \in \Omega$, is sent by "magical" sensors in the pool to the function machine X, which calculates a particular value from the state of the system. The "function machine" is analogous to a random variable, which is a function on a sample space. For example, the output would be $X_1 = X(\omega_1)$ when the stochastic system is in the state ω_1, $X_2 = X(\omega_2)$ when the system is in the state ω_2, and so on. The function machine then outputs these values, and although it could possibly output its values continuously,
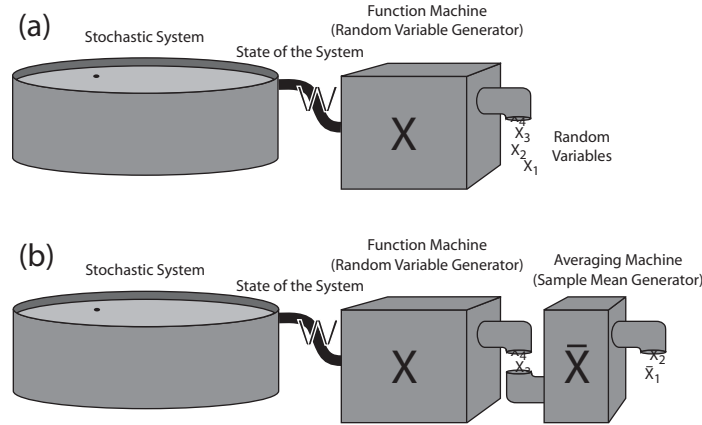
Figure 4.1: A hypothetical "random machine". A floating particle moves freely in the pool at the left. This constitutes a stochastic system. "Magical" sensors in the pool send all information (ω) about the state of the system to the function machine (X). The function machine determines a specified value, e.g. the distance of the particle from the edge of the pool, from the state of the system and outputs this value every second, say. An observer only has access to the output values of the function machine and wishes to study the floating particle system. In (a), the observer looks at each of the individual output values of the function machine. In (b), the observer has decided to use an averaging machine ($\bar{X}$), another type of "function machine", that collects the output values of X at its input and outputs the average values $\bar{X}_k = \sum_{i=1}^{k} X_i$. See text for further explanation.

for ease of exposition, let us suppose that it outputs one every second, say. Also, suppose that these values at the output of the function machine are the only data that an observer has available for studying the system. This situation is shown in Figure 4.1a.

The population moments, in this case, are directly determined by those properties, intrinsic to the stochastic system, that influence the behavior of the parameters of the system that affect the value of the random variable output. They convey information about the expected behavior of the sequence of output values but are not directly available to our hypothetical observer. If this observer is interested in the value of a particular population moment, then, at best, he can infer its value from a set of output values, the random variables. To do this, he uses a statistic, a function of a set of random variables, that is a good estimator of the relevant population moment. For instance, let's say that the observer interested in the mean. Then he might, as shown in Figure 4.1b, place an "averaging machine" so that it collects the output values from the "function machine", averages them, and emits these average values. Hopefully, then, these averages approximate the population mean well.

The sample moments are a particular set of statistics, and these statistics, or statistics based upon them, are usually good estimators of the population moments. For example, suppose that we have a sample of n values, $X_1, X_2, \ldots, X_n$, from the output of a stochastic process, and, furthermore, suppose that it is reasonable to think that the values in this sample are statistically independent. Then, the sample mean, which is the same as the

first-order sample moment

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_n,$$

is a good estimator of the population mean. When estimating the population variance, however, the best estimator is not the second-order (central) moment

$$m_2 = \frac{1}{n} \sum_{i=1}^{n} (X_n - \bar{X})^2,$$

since its expected value is not equal to the population variance. But the best estimator, s^2, is a function of the second-order (central) moment:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_n - \bar{X})^2 = \left(\frac{n}{n-1} \right) m_2.$$

This is what is ordinarily called the sample variance.

In general, how well a particular statistic estimates the population moment is determined by the unobservable system. Thus, in choosing a statistic to estimate a particular population moment, one is implicitly, if not explicitly, relying on a model of the unobservable system. Especially for lower order moments, e.g. the mean and variance, the class of models for which certain estimators are, in some sense, optimal may be very large. However, the behavior of such estimators may vary considerably across this class. The most familiar behavioral characteristics of common moment estimators have usually been derived under the assumption that the observations within a sample are independent. Examples are that the variance of the sample mean is equal to the population variance divided by the size of the sample and that the expected value of sample variance is equal to the population variance. But if the independence assumption is violated, these need not be true.

4.3.2 The Quality of Statistical Estimators

The quality of a statistic used to estimate a parameter of a stochastic process can be assessed by considering its bias, consistency, and variability. In the previous section, we hinted at the concept of the bias of an estimator, which is based upon the intuition that a statistic should on average be equal to the parameter that it is estimating. If $\hat{\theta}$ is an estimator of the parameter θ, then the bias of $\hat{\theta}$ is $E\{\hat{\theta}\} - \theta$. Thus, it is desirable that the bias be zero, or that the statistic be unbiased. If the bias is positive, then the statistic tends to overestimate the parameter, and if it is negative, then the statistic tends to underestimate the parameter. Furthermore, even if a statistic is biased, it may be be asymptotically unbiased, which means that its bias will go to zero as the sample size increases. Although such a statistic may not be ideal, it may be the best when other competing factors, such as its variability, are considered.

Not only is it desirable that a statistic be equal to the value of the estimated parameter, but the probability that it is close to the parameter value should increase as the sample size increases. This notion is made mathematically rigorous with the concept of

consistency. A statistic is said to be consistent if it converges in probability to the value of the estimated parameter. Loosely, consistency means that as the sample size increases the distribution of the statistic better approximates a Dirac delta function with its mass located at the value of the parameter. In other words, the probability accumulates at the parameter value. In addition to knowing that a statistic is consistent, it is helpful to know how fast it converges to the parameter value. In order to have a precise gauge of this, we would need to know the full distribution of the statistic, which is usually not available or obtainable. Thus, to assess the speed of convergence, we can consider the variance of the estimator. Clearly, if a statistic is consistent, then its variance will asymptotically approach zero.

Thus, to assess the quality of a statistic, or estimator, we should look at its bias and its variance. That these are the two critical properties that should be considered is supported by the fact that the mean squared error of an estimator is equal to the sum of its variance and the square of its bias. Assuming again that the statistic $\hat{\theta}$ is an estimator of the parameter θ, the mean squared error (MSE) of this estimator is

$$
\begin{aligned}
MSE &\equiv E\left\{(\hat{\theta} - \theta)^2\right\} \\
&= E\left\{\hat{\theta}^2 - 2\hat{\theta}\theta + \theta^2\right\} \\
&= E\left\{\hat{\theta}^2 - 2\left(E\{\hat{\theta}\}\right)^2 + \left(E\{\hat{\theta}\}\right)^2 + \left(E\{\hat{\theta}\}\right)^2 - 2\hat{\theta}\theta + \theta^2\right\} \\
&= E\left\{\hat{\theta}^2 - 2\hat{\theta}E\{\hat{\theta}\} + \left(E\{\hat{\theta}\}\right)^2\right\} + \left(E\{\hat{\theta}\}\right)^2 - 2E\{\hat{\theta}\}\theta + \theta^2 \\
&= E\left\{(\hat{\theta} - E\{\hat{\theta}\})^2\right\} + \left(E\{\hat{\theta}\} - \theta\right)^2 \\
&= \quad Variance \quad + \quad Bias^2
\end{aligned}
$$

4.4   Models of Long-Range Dependent Spike Trains

4.4.1   The Fractional-Gaussian-Noise-Driven Poisson Process

In order to develop a theory of the performance of estimators for firing rates in neurons, we must have a model of the statistical characteristics of spike trains. In particular, for our purposes, the model must include the long-range dependence observed in spike trains. We will use the fractional-Gaussian-noise-driven Poisson process (fGnDP), which was introduced in Section 3.6.1, as our model. The fGnDP was one of the models developed for the spike trains in auditory nerve and other sub-cortical sensory neurons once long-range dependence was observed empirically Teich (1989), Teich et al. (1990), Teich (1992), Teich and Lowen (1994), Lowen and Teich (1993a, 1995, 1996a, 1997), Kumar and Johnson (1993), Thurner et al. (1997). Thus, it should be relevant to most sub-cortical sensory neurons. Furthermore, since it is long-range interval dependent but has finite-variance interpoint intervals, as we saw was the case for cortical neurons, the general results obtained from this model should be applicable to these neurons as well.

Unless stated otherwise, our model will be a doubly stochastic Poisson process with

the stochastic rate process

$$\Lambda(t) = \max\left\{0, \lambda + \sigma G_H\left(\left\lfloor\left|\frac{t_0 + t}{\tau}\right|\right\rfloor\right)\right\}, \tag{4.2}$$

where $\{G_H(k), k \in \mathbb{Z}\}$ is standard fractional Gaussian noise, λ and σ are positive constants, t_0 is a random variable uniformly distributed on the interval $[0, \tau)$ that is fixed for each sample function, and $\lfloor x \rfloor$ is the largest integer less than x. The random variable t_0 is necessary in order for the process to be stationary, but has no effect otherwise.

In the following section, we will give some results for this model. In the derivation of those results, we have assumed that $\sigma \ll \lambda$, so that the probability that $\lambda + \sigma G_H < 0$ is negligible. Thus, any results derived will necessarily, by virtue of this assumption, be approximate. However, in order to obtain other results, it was necessary to make a further approximation. The autocovariance function of the standard fGn process $G_H(k)$ is (e.g. Beran, 1994, p. 52; Samorodnitsky & Taqqu, 1994, p. 333)

$$\gamma_G(k) = E\{G_H(0)G_H(k)\} = \frac{1}{2}\left\{(k+1)^{2H} - 2k^{2H} + |k-1|^{2H}\right\}, \quad k = 0, 1, 2, \ldots. \tag{4.3}$$

Thus, with the only approximating assumption being that the truncation effect of the "max" function is negligible, the autocovariance of the rate process in (4.2) is

$$\begin{aligned}
\gamma_\Lambda(t) &= E\left\{\left(\Lambda(t_0) - \lambda\right)\left(\Lambda(t_0 + t) - \lambda\right)\right\} \\
&= E\left\{\left(\sigma G_H\left(\left\lfloor\left|\frac{t_0}{\tau}\right|\right\rfloor\right)\right)\left(\sigma G_H\left(\left\lfloor\left|\frac{t_0 + t}{\tau}\right|\right\rfloor\right)\right)\right\} \\
&= \sigma^2 E\left\{G_H\left(\left\lfloor\left|\frac{t_0}{\tau}\right|\right\rfloor\right)G_H\left(\left\lfloor\left|\frac{t_0 + t}{\tau}\right|\right\rfloor\right)\right\} \\
&= \sigma^2 E_{t_0}\left\{\gamma_G\left(\left\lfloor\left|\frac{t_0 + t}{\tau}\right|\right\rfloor - \left\lfloor\left|\frac{t_0}{\tau}\right|\right\rfloor\right)\right\}, \quad \text{for } t \geq 0,
\end{aligned} \tag{4.4}$$

where E_{t_0} is the expectation with respect to the random variable t_0. However, the conversion from the discrete samples of fGn to the stochastic rate process in continuous time creates difficulties for some analytical derivations. Thus, in these instances, we will use the following continuous approximation to (4.3):

$$\begin{aligned}
\tilde{\gamma}_G(s) &= \frac{1}{2}\left\{\left(\frac{s}{\tau} + 1\right)^{2H} - 2\left(\frac{s}{\tau}\right)^{2H} + \left|\frac{s}{\tau} - 1\right|^{2H}\right\} \\
&= \frac{1}{2\tau^{2H}}\left\{(s + \tau)^{2H} - 2s^{2H} + |s - \tau|^{2H}\right\}, \quad \text{for } s \geq 0.
\end{aligned} \tag{4.5}$$

Thus, $\tilde{\gamma}_G(k\tau) = \gamma_G(k)$ for $k = 0, 1, 2, \ldots$. The corresponding approximation to the autocovariance of the rate process is

$$\tilde{\gamma}_\Lambda(t) = \frac{\sigma^2}{2\tau^{2H}}\left\{(s + \tau)^{2H} - 2s^{2H} + |s - \tau|^{2H}\right\}, \quad \text{for } t \geq 0. \tag{4.6}$$

4.4.2 Moments and Moments of Statistics for the Fractional-Gaussian-Noise-Driven Poisson Process

In this section, we will, for the most part, only give results, since most of the derivations or proofs are cumbersome but not conceptually helpful in subsequent results nor in the following sections. Those proofs that are not contained in this section can, however, be found in Appendix C.

The mean and variance must exist (i.e. be finite) if they are to be estimated, and we must know their values in order to compare them to values of their estimators. In the following theorems, we derive approximate formulae for the mean and variance of the count in an interval of length T for an fGnDP.

Theorem 4.1. *Let $N(\cdot)$ be an fGnDP with rate process*

$$\Lambda(t) = \max\left\{0, \lambda + \sigma G_H\left(\left\lfloor\frac{t}{\tau}\right\rfloor\right)\right\}, \tag{4.7}$$

where $G_H(k)$ is standard fractional Gaussian noise with Hurst index H. Assuming that the right term is negative with a probability of nearly zero, the mean count in an interval of length T is

$$E\{N(T)\} \approx \lambda T,$$

and the variance of the count in the same interval is given by the following:

(i) If $0 < T < \tau$, then

$$Var\{N(T)\} \approx \lambda T + \sigma^2 T^2\left(1 + \frac{2^{2H-1} - 2}{3} \cdot \frac{T}{\tau}\right); \tag{4.8a}$$

(ii) If $T \geq \tau$ and $n = \left\lfloor\frac{T}{\tau}\right\rfloor$, then

$$Var\{N(T)\} \approx \lambda T + \frac{\sigma^2 \tau^2}{6}\left\{\left(\frac{T}{\tau} - n\right)^3 (n+2)^{2H}\right.$$
$$- \left[3\left(\frac{T}{\tau} - n\right)^2\left(\frac{T}{\tau} - n - 1\right) - \left(3\left(\frac{T}{\tau} - n\right) + 1\right)\right](n+1)^{2H}$$
$$\left. + \left[3\left(\frac{T}{\tau} - n\right)^2\left(\frac{T}{\tau} - n - 2\right) + 4\right]n^{2H} - \left(\frac{T}{\tau} - n - 1\right)^3 (n-1)^{2H} - 2\right\}. \tag{4.8b}$$

Corollary 4.2. *Let $N(\cdot)$ be the fGnDP in Theorem 4.1. If T is an integer multiple of τ, then the variance of the counting measure is*

$$Var\{N(T)\} \approx \lambda T + \frac{\sigma^2 \tau^2}{6}\left\{\left(\frac{T}{\tau} + 1\right)^{2H} + 4\left(\frac{T}{\tau}\right)^{2H} + \left(\frac{T}{\tau} - 1\right)^{2H} - 2\right\}.$$

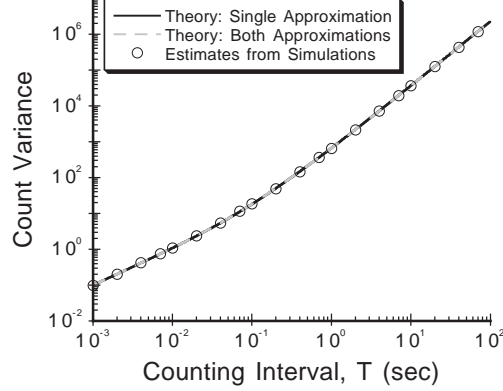*Proof.* Substitute $\frac{T}{\tau}$ for n in (4.8b). $\qquad\square$

Figure 4.2: Comparison between different approximations to the variance of the counts for an fGnDP. The black, solid line is the curve for (4.8), which was obtained assuming only that negative values of the non-truncated rate process were negligible. The gray, dotted line is the curve for (4.9), which, in addition, included the autocovariance approximation of (4.6). The open circles are estimates from 10,000 independent simulations. The parameter values used in all cases were $H = 0.9$, $\lambda = 100$ spikes/s, $\sigma = 30$ spikes/s, and $\tau = 0.1$ s.

Theorem 4.3. *Let $N(\cdot)$ be the fGnDP in Theorem 4.1. Using the approximation (4.6) to the autocovariance function of its rate process, the count variance of N can be approximated by*

$$Var\left\{N(T)\right\} \approx \lambda T + \frac{\sigma^2 \tau^2}{2(H+1)(2H+1)}$$
$$\times \left\{\left(\frac{T}{\tau} + 1\right)^{2(H+1)} - 2\left(\frac{T}{\tau}\right)^{2(H+1)} + \left|\frac{T}{\tau} - 1\right|^{2(H+1)} - 2\right\}. \quad (4.9)$$

Figure 4.2 contains a comparison between the approximations of (4.8), (4.9), and estimates, using the sample variance, from independent simulations of an fGnDP. The two analytical approximations are practically identical, and the results from simulations are consistent with these theoretical results. This comparison provides evidence that the approximation (4.6) produces relatively little additional error into our approximation of the variance, which supports the use of (4.6) when (4.4) is too unwieldy.

The following describes the paradigm that we will consider as the typical method of estimating counts and rates in neural spike trains, and defines the necessary parameters and statistics. A diagram of this paradigm is shown in Figure 4.3. The variable names specified in this definition will be used in a consistent manner throughout the remainder of this paper.

Definition 4.4. Let $N(\cdot)$ be a stationary, orderly point process in time. Fix the values $S > 0$, the length of the repetition interval, and $0 < T \leq S$, the length of the counting interval. Define the sequence of counts in T-length intervals starting every S units as

$$N_i(T) = N\big((i-1)S + T\big) - N\big((i-1)S\big), \ i = 1, 2, 3, \ldots.$$
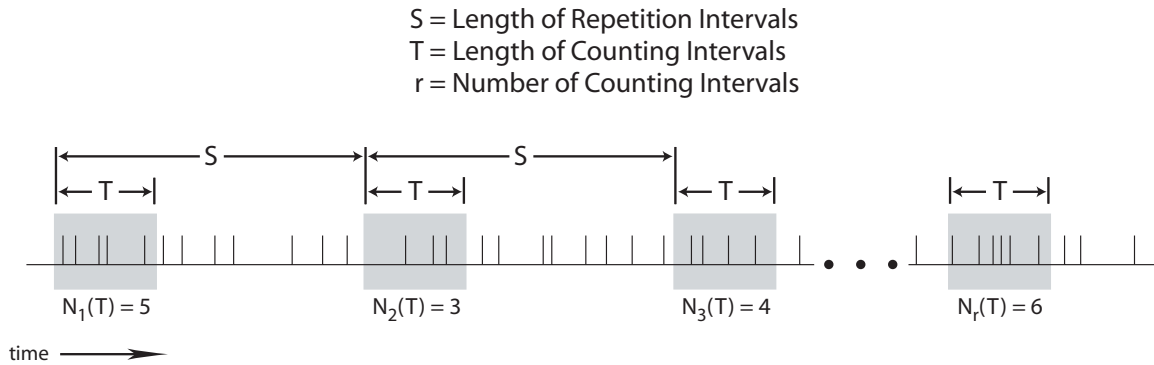
Figure 4.3: Diagram of the standard method for estimating statistics of spike counts or rates from a spike train, as described in Definition 4.4. Time is represented by horizontal distance, and each vertical line represents the occurrence of a single spike. Gray areas designate the intervals on which the spikes are counted, and the number of spikes occurring in a single gray box is one sample count.

Let $m_r(T)$ be the sample mean of $r \in \mathbb{N}$ consecutive counts in this sequence, i.e.

$$m_r(T) = \frac{1}{r} \sum_{i=1}^{r} N_i(T),$$

and let $s_r^2(T)$ be their sample variance, i.e.

$$s_r^2(T) = \frac{1}{r-1} \sum_{i=1}^{r} \Big(N_i(T) - m_r(T) \Big)^2.$$

In the following two theorems we find the expected value and variance of the sample mean, the statistic used to estimate the population mean of the counts.

Theorem 4.4. *Let $N(\cdot)$ be an fGnDP as in Theorem 4.1, and define the sample mean of the counts as in Definition 4.4. Then the expected value of the sample mean, assuming only that truncation of the rate process is negligible, is*

$$E\{m_r(T)\} = E\{N(T)\} \approx \lambda T.$$

Thus, the sample mean is an unbiased estimator of the mean count.

Theorem 4.5. *Let $N(\cdot)$ be an fGnDP as in Theorem 4.1, and define the sample mean of the counts as in Definition 4.4. Assume that $S \geq \tau$. Then, using the approximation (4.6) to the autocovariance function of the rate process of N, the variance of the sample mean*

count is approximately

$$Var\{m_r(T)\} \approx \frac{\lambda T}{r} + \frac{\sigma^2}{2r\tau^{2H}(2H+1)(H+1)}$$

$$\times \left\{ \left[(T+\tau)^{2(H+1)} - 2\left(T^{2(H+1)} + \tau^{2(H+1)} \right) + |T-\tau|^{2(H+1)} \right] \right.$$

$$+ \frac{1}{r}\sum_{k=1}^{r-1}(r-k)\left[(kS+T+\tau)^{2(H+1)} - 2(kS+T)^{2(H+1)} + |kS+T-\tau|^{2(H+1)} \right.$$

$$- 2\left[(kS+\tau)^{2(H+1)} - 2(kS)^{2(H+1)} + |kS-\tau|^{2(H+1)} \right]$$

$$+ (kS-T+\tau)^{2(H+1)} - 2(kS-T)^{2(H+1)} + |kS-T-\tau|^{2(H+1)} \left.\left. \right] \right\}. \quad (4.10)$$

Finally, the following theorem gives the expected value of the sample variance, the statistic used to estimate the population variance of the counts. We do not, however, have an analytical formula for the variance of this estimator.

Theorem 4.6. *Let $N(\cdot)$ be an fGnDP as in Theorem 4.1, and define the sample mean and sample variance of the counts as in Definition 4.4. Assume that $S \geq \tau$. Then, using the approximation (4.6) to the autocovariance function of the rate process of N, the expected value of the sample variance is approximately*

$$E\left\{ s_r^2(T) \right\} \approx Var\{N(T)\} - \left(\frac{1}{r(r-1)} \right) \cdot \left(\frac{\sigma^2}{2\tau^{2H}(2H+1)(H+1)} \right)$$

$$\times \sum_{k=1}^{r-1}(r-k)\left\{ (kS+T+\tau)^{2(H+1)} - 2(kS+T)^{2(H+1)} + |kS+T-\tau|^{2(H+1)} \right.$$

$$- 2\left[(kS+\tau)^{2(H+1)} - 2(kS)^{2(H+1)} + |kS-\tau|^{2(H+1)} \right]$$

$$+ (kS-T+\tau)^{2(H+1)} - 2(kS-T)^{2(H+1)} + |kS-T-\tau|^{2(H+1)} \left. \right\} \quad (4.11)$$

Corollary 4.7. *Under the conditions of Theorem 4.6, the sample variance $s_r^2(T)$ of the counts is a negatively biased estimator of the count variance if the fGnDP is long-range dependent, i.e. if $H > 0.5$. In other words, the sample variance, on average, underestimates the true variance of the counts of the point process.*

The standard deviation of spike counts is often more useful than the variance since it is in the same units as the mean. The sample standard deviation, used to estimate the true standard deviation of the process, is just the square root of the sample variance, $s_r(T)$. However, as the following lemma shows, even when the sample variance is unbiased, the sample standard deviation is negatively biased.

Lemma 4.8. *Let s^2 be an unbiased estimator of the variance σ^2 of a stochastic process or random variable that is nonconstant with positive probability. Then s is a negatively biased estimator of the standard deviation σ.*

Proof. The variance of a random variable is always nonnegative and is equal to zero only if the random variable is constant with probability one. Thus, for the statistic s of the given stochastic process or random variable,

$$Var\{s\} > 0.$$

Now, also,

$$Var\{s\} = E\left\{(s - E\{s\})^2\right\} = E\{s^2\} - \left(E\{s\}\right)^2 = \sigma^2 - \left(E\{s\}\right)^2,$$

since s^2 is an unbiased estimator of σ^2. Thus,

$$\sigma^2 - \left(E\{s\}\right)^2 > 0,$$

or

$$\left(E\{s\}\right)^2 < \sigma^2.$$

Taking the square root of both sides, we get

$$E\{s\} < \sigma,$$

which means that s underestimates the standard deviation. $\square$

Furthermore, the following corollary shows that the square root of the expected value of the sample variance can be used to place an upper bound on the expected value of the sample standard deviation, if the former is negatively biased. Thus, in some sense, the situation in Corollary 4.7 becomes worse when an estimate of the standard deviation is needed.

Corollary 4.9. *Let s^2 be a negatively biased estimator of the variance σ^2 of a stochastic process or random variable that is nonconstant with positive probability. Then, s is a negatively biased estimator of the standard deviation σ, and (the magnitude of) its bias is larger than $\sigma - \sqrt{E\{s^2\}}$.*

Proof. In this case, we have

$$Var\{s\} = E\left\{(s - E\{s\})^2\right\} = E\{s^2\} - \left(E\{s\}\right)^2 > 0.$$

So,

$$E\{s^2\} - \left(E\{s\}\right)^2 > 0$$
$$\left(E\{s\}\right)^2 < E\{s^2\} < \sigma^2$$
$$E\{s\} < \sqrt{E\{s^2\}} < \sigma,$$

which means that $\sigma - E\{s\} > \sigma - \sqrt{E\{s^2\}} > 0.$ $\square$

4.4.3 Refractory-Modified Fractional-Gaussian-Noise-Driven Poisson Process

The most obvious difference between the output of an fGnDP and real neural spike trains is the absence of very short interspike intervals in the latter. In neurophysiology, this absence is usually attributed to refractoriness in the neural spiking mechanism. The refractory period of a neuron can be divided into two parts. Immediately following a spike, there is a relatively short period (e.g. approximately 0.7 milliseconds in primary auditory neurons) when another spike cannot occur. This is referred to as the *absolute* refractory period. Following the absolute refractory period is a much longer period called the *relative* refractory period. During this period the probability of a spike occurring is less than if the previous spike had not occurred, but is not zero.

No general theory of refractoriness exists for neurons, but Young and Barta (1986) have suggested a simple model to approximate the effect of refractoriness on spike trains. Their model produces relative-refractoriness-like behavior in a Poisson process by introducing independent and exponentially distributed absolute refractory periods after each event, or "spike", that is produced by the Poisson process. Thus, since the "absolute refractory period" can be arbitrarily small, the model does not have an absolute refractory period in the stationary sense. However, since the absolute refractory period of neurons is so short, their model produces an effect that is still very similar to the entire refractory period of neurons. In their paper, Young and Barta (1986) show that an exponentially distributed absolute refractory period (or "deadtime") with a mean of 4 ms produces results that are a good approximation to the refractoriness seen in primary auditory neurons.

Since the effect of the refractory period of a neuron on its statistical attributes may be significant, we will also test our results on an fGnDP with the refractory model of Young and Barta (1986). Simulations of this refractory-modified fGnDP (RM-fGnDP) are produced in the following manner. First, a spike train is produced by the fGnDP model. Then the length of an absolute refractory period is drawn from an exponentially distributed random variable with a mean of 4 ms. Any spike that occurs within this length of time after the first spike in the spike train is discarded. Once this absolute refractory period ends, the first spike to occur remains in the spike train. A second refractory period length is then chosen, and this (absolute) refractory period is applied immediately after this spike. This process is repeated until the end of the spike train is reached. The end result is a spike train consisting of all of the spikes that were not discarded.

4.5 Effect of Long-Range Dependence on Sample Statistics

We have shown analytically that the sample mean is an unbiased estimator of the population mean of the counts for an fGnDP, so long-range dependence does not change the bias of the sample mean. Results from simulations of an fGnDP (not shown) support this theoretical result. The variability of this statistic, however, is affected by long-range dependence. For r independent and identically distributed observations, the standard

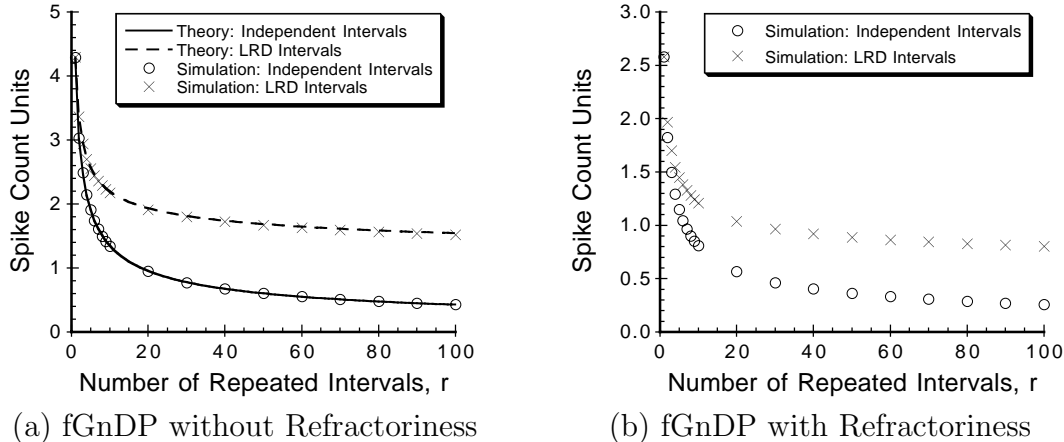(a) fGnDP without Refractoriness   (b) fGnDP with Refractoriness

Figure 4.4: The variability of the sample mean count as a function of the number of counting intervals for an fGnDP without (a) and with (b) refractoriness. The solid line in (a) is the theoretical standard deviation of the sample mean calculated from r independent counting intervals. The equation for this curve is given in (4.12). The circles, in both (a) and (b), are estimates of this standard deviation calculated from 10,000 independent sets of spike trains, with each set consisting of r independent simulations of an fGnDP. The dashed line in (a) is the theoretical standard deviation of the sample mean count calculated according to Definition 4.4. The $\times$'s, in both (a) and (b), are estimates of this standard deviation calculated from 10,000 independent simulations of an fGnDP. The parameter values for the fGnDP, theoretical and simulated, were $H = 0.9$, $\lambda = 100$ spikes/s, $\sigma = 30$ spikes/s, and $\tau = 0.1$ s. The counting interval length, T, was 0.1 s, and the repetition interval length, S, was 1 s. The refractory period was exponential distributed with a mean of 4 ms. The average spike rate for the RM-fGnDP simulations was approximately 71 spikes/s.

deviation of the sample mean is

$$\sqrt{Var\{\bar{X}\}} = \frac{\sigma}{\sqrt{r}}, \tag{4.12}$$

where σ is the standard deviation of the common distribution. In Figure 4.4(a) this result (solid line) is compared to the standard deviation of the sample mean count calculated according to Definition 4.4 for an fGnDP (dashed line) with Hurst index[2] $H = 0.9$. The latter curve is the square root of the theoretical result for the variance of the sample mean count given in Theorem 4.5. The symbols are the estimates of the standard deviation of the sample mean, where each value of the sample mean is calculated from r independent simulations of an fGnDP (circles) or from a single, longer simulation of an fGnDP ($\times$'s). Thus, our simulations agree very well with the theoretical results.

---

[2]Estimates of the Hurst index, which is equal to $(D + 1)/2$, where D is the "fractal dimension", for neural spike trains predominately fall in the interval from about 0.75 to 0.95 with a tendency toward the upper portion of this interval (Teich, 1989; Teich et al., 1990; Teich, 1992; Kelly, Johnson, Delgutte, & Cariani, 1996; Lowen & Teich, 1996b; Teich et al., 1997; Lowen et al., 2001). Therefore, since such estimates are also usually negatively biased (e.g. Lowen & Teich, 1993a), we have chosen a value, 0.9, in the upper portion of this range.

As expected, the positive dependence in the fGnDP reduces the rate of decay of the variability of the sample mean with the number of observations, or counting intervals. Thus, if the observations are independent, then using 100 observations instead of just one reduces the standard deviation by a factor of ten. For observations from an fGnDP, however, this only reduces the standard deviation by a factor of about 2.75. In this case, the standard deviation of this estimator is over three times larger when calculated from 100 counting intervals in the same spike train than when calculated from 100 independent counting intervals. Furthermore, since the slope of the curve for the dependent counting intervals is noticeably shallower near $r = 100$ than that for the independent counting intervals, this ratio will increase dramatically as even more observations are included.

Although analytical results are not available for the refractory-modified fGnDP, simulation results for this model are shown in Figure 4.4(b). The meaning of the symbols in this figure are identical to the meaning of those in Figure 4.4(a), only now the estimates are derived from simulations of an RM-fGnDP. Due to the exclusion of some spikes, the average rate of the RM-fGnDP was about 71 spikes per second, instead of 100 spikes per second like the original fGnDP. Comparing these two figures, it appears that the standard deviations of the sample means for an RM-fGnDP are about 60% of those for an fGnDP. Thus, much, but perhaps not all, of the reduction in variability of this estimator may be due to the reduction in spike rate. The relationship between the variability for dependent intervals and the variability for independent intervals, however, is very similar to the results for an fGnDP. Again, the standard deviation of the sample mean for 100 dependent intervals is over three times larger than that for 100 independent intervals. Furthermore, the slope of the dependent-interval curve is also much shallower than the slope of the independent-interval curve at $r = 100$. Thus, refractoriness does not appear to alter the effect of the dependence in fGnDPs on the variability of the sample mean.

In contrast to the sample mean, we found analytically that the sample variance and the sample standard deviation of the counts of an fGnDP, calculated from a single spike train, are biased estimators of the population variance and the population standard deviation, respectively. Both of these estimators are negatively biased, meaning that they tend to underestimate the population values, and the sample standard deviation is, in some sense, more biased than the sample variance. Figure 4.5(a) corroborates these results.

The solid line in Figure 4.5(a) is the theoretical value of the population standard deviation as determined by the square root of (4.8). This line denotes the value that one would like to obtain from the sample standard deviation, or the square root of the sample variance, and that, obviously, does not depend on the parameters of the estimator, such as r. In other words, if the estimators were unbiased, this would be the expected value of the sample standard deviation and the square root of the expected value of the sample variance for all values of r. Since the sample variance is unbiased if the observations are independent, the values of the square root of the sample variance (squares) calculated from independent intervals of simulations do indeed lie close to this line at all values of r. However, since the sample standard deviation is negatively biased when the sample variance is unbiased (see Lemma 4.8), the values of the sample standard deviation (circles) always lie below this line. The difference between the sample standard deviation and the actual standard deviation is quite large at low values of r, but decreases quickly as r

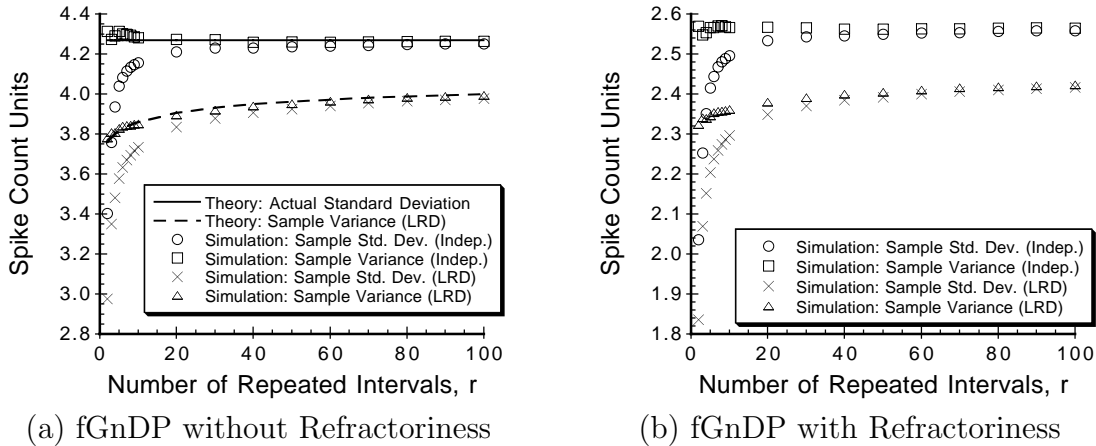(a) fGnDP without Refractoriness (b) fGnDP with Refractoriness

Figure 4.5: The variability and estimates of the variability of the counts for an fGnDP without (a) and with (b) refractoriness as a function of the number of counting intervals. The solid line in (a) is a constant curve at the theoretical value of the standard deviation of the counts, which was calculated by taking the square root of (4.8). The dashed line in (a) is the square root of the theoretical expected value of the sample variance in Definition 4.4, which was calculated by taking the square root of (4.11). The circles, in both (a) and (b), are the average values of the sample standard deviation calculated from 10,000 independent sets of simulated spike trains, each set consisting of r independent spike trains. The squares are the square roots of the average values of the sample variance calculated from the same simulations. The $\times$'s are the average values of the sample standard deviation calculated according to Definition 4.4 from 10,000 independent simulated spike trains, and the triangles are the square roots of the average values of the sample variance from the same simulations. All parameter values are the same as Figure 4.4.

increases.

The dashed line is the square root of the theoretical expected value of the sample variance of an fGnDP, calculated according to Definition 4.4, which is given in (4.11). Estimates of this value (triangles), calculated from simulations of an fGnDP, agree with the theoretical curve very well. Both of these results confirm that the sample variance calculated from a single spike train output of the fGnDP model is indeed negatively biased. While the actual value of the standard deviation, i.e. the square root of the variance, for these parameters is about 4.27, the square root of the expected value of the sample variance ranges from slightly below 3.8 for two intervals up to 4.0 for 100 intervals. These are underestimates by over 11% and by 6.3%, respectively. Although this graph suggests that the sample variance is asymptotically unbiased, the bias decays quite slowly due to the long-range correlations in the process. The $\times$'s are estimates, obtained from simulations, of the expected value of the sample standard deviation for dependent counting intervals from the fGnDP. They reveal, as expected from Corollary 4.9, that the sample standard deviation is even more biased than the sample variance when their biases are compared in the same units. The additional bias of the sample standard deviation is quite large for small numbers of observations, but quickly decays as more observations are

used. In fact, the amount and rate of decay of this additional bias is comparable to the bias of the sample standard deviation for independent observation intervals.

Figure 4.5(b) displays the corresponding results from simulations of an RM-fGnDP. The symbols in this figure represent the same quantities as in Figure 4.4(a), the only difference being the spike-train-producing process itself. As for variability of the sample mean, the relationships in Figure 4.5(b) are very similar to those in Figure 4.4(a) for the fGnDP without refractoriness. Furthermore, just like we found for the variability of the sample mean, the values for the RM-fGnDP are about 60% of the values for the fGnDP. Again, much of this reduction may be simply due to the reduction in spike rate. Thus, the general effect of the dependence in the output of an fGnDP on the expected values of the sample standard deviation and the sample variance does not seem to be altered by refractoriness.

Finally, in Figure 4.6, we consider the variability of the sample standard deviation and the sample variance. We have no theoretical results in this case, so our conclusions will have to be based on estimates from simulations of an fGnDP. According to these results, in contrast to the standard deviation of the sample mean, the standard deviation of the sample standard deviation is practically the same for $r \leq 100$ whether the counting intervals are independent (circles) or long-range dependent ($\times$'s) for both the fGnDP (Figure 4.6(a)) and the RM-fGnDP (Figure 4.6(b)). However, the standard deviation of the sample standard deviation does decay with increasing r in a manner similar to the decay of the standard deviation of the sample mean. Furthermore, although not conclusive, the present results suggest that the standard deviation of the sample standard deviation might actually decay more slowly for LRD intervals than for independent intervals. The evidence for this can be seen in Figure 4.6(a), where the values for the LRD intervals are lower than those for the independent intervals for $r < 20$, nearly equal to those for the independent intervals in the middle of the graph, and slightly higher than those for the independent intervals when $r > 80$. If this is the case, then the variability of the sample standard deviation will be noticeably larger for LRD intervals than for independent intervals for large numbers of observations. Although this is likely to be the case, the differences in our results up to $r = 100$ are too small to support a definitive judgement.

As for the sample standard deviation, Figures 4.6(a) and 4.6(b) suggest that the variability of the sample variance seems to be nearly the same for $r \leq 100$ whether the counting intervals are independent (squares) or long-range dependent ($+$'s) and whether the model is an fGnDP or an RM-fGnDP. In these figures, this is observed in the values of the square root of the standard deviation of the sample variance. Also, the decay of the variability of the sample variance is quite similar to the decay of the variability of the sample standard deviation. Furthermore, the rate of this decay at large values of r seems to be slower for LRD intervals than for independent intervals as well. The most striking aspect of these graphs, however, is that the square root of the standard deviation of the sample variance is significantly larger than the standard deviation of the sample standard deviation. But it is unclear whether any meaningful conclusion can be derived from this observation. One might be tempted to conclude that the sample variance is more variable than the sample standard deviation, but in this context it is difficult to make the phrase "more variable" have any precise meaning beyond the observable relationship.

(a) fGnDP without Refractoriness
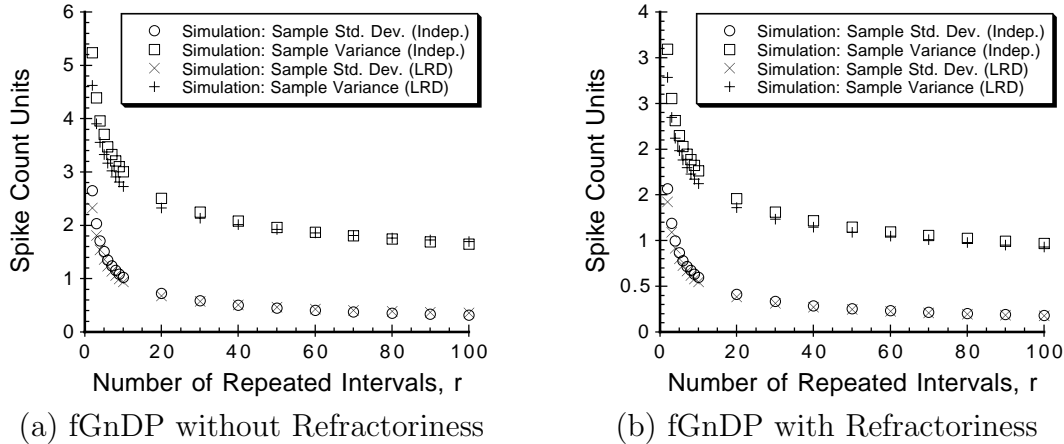
(b) fGnDP with Refractoriness

Figure 4.6: The variability of the sample variance and sample standard deviation of counts as a function of the number of counting intervals for an fGnDP without (a) and with refractoriness (b). The circles, in both (a) and (b), are the values of the standard deviation of the sample standard deviation calculated from 10,000 independent sets of simulated spike trains, each set consisting of r independent spike trains. The squares are the square roots of the values of the standard deviation of the sample variance calculated from the same simulations. The $\times$'s are the values of the standard deviation of the sample standard deviation calculated according to Definition 4.4 from 10,000 independent simulated spike trains, and the triangles are the square roots of the values of the standard deviation of the sample variance from the same simulations. All parameter values are the same as Figure 4.4.

4.6 Implications for Statistical Analysis of Neurophysiological Spike Trains

4.6.1 Estimation of the Spike Rate from a Single Counting Interval

The spontaneous firing rate, i.e. the spike rate in the absence of stimulation, is a basic attribute of neurons that is often estimated from the spike count in a long counting interval. The spontaneous rate of neurons can be used as an indicator by which to group neurons that have other characteristics in common as well. For example, the spontaneous rates of primary auditory neurons are correlated with their thresholds to sound (Liberman, 1978; Geisler, Deng, & Greenberg, 1985), the shapes of their response functions for tonal stimuli (Sachs & Abbas, 1974; Palmer & Evans, 1980; Sachs, Winslow, & Sokolowski, 1989), certain of their morphological characteristics (Liberman, 1982; Leake, Snyder, & Hradek, 1993; Kawase & Liberman, 1992), and the patterns of their projections into the cochlear nucleus (Ryugo & Rouiller, 1988; Ryugo & May, 1993; Leake & Snyder, 1989; Leake, Snyder, & Merzenich, 1992; Liberman, 1991, 1993). Furthermore, although rare, long counting intervals are sometimes used to measure the response of neurons to stimuli that are stationary (in some sense). For instance, in Geisler et al. (1985), the average

spike rate of auditory neurons was measured using a single, long counting interval at each intensity of a tonal stimulus. By making these measurements at many different stimulus intensities, they produced rate-intensity functions for each neuron, which describe the relationship between average spike rate and the intensity of the tonal stimulus that produced it. The much more common method for measuring rate-intensity functions will be described in Section 4.6.2.

What effect does the presence of LRD in spike trains have upon the types of measurements mentioned above? Since these measurements are based on the sample mean, our previous results indicate that they should not be biased. However, the variability of these measurements should be larger than one would expect if the spike trains had little or no temporal correlation. In order to gauge this increase in variability, we can compare the variability of the sample mean for an LRD fGnDP model with that from non-LRD point process models.

In Figure 4.7, the standard deviations of spike rate estimates are compared for fGnDPs with Hurst index $H = 0.9$, fGnDPs with $H = 0.5$, and Poisson processes for nominal spike rates of 5 spikes per second and 100 spikes per second. The theoretical standard deviations for the fGnDPs were calculated by dividing the square root of (4.8), the variance of the count in an interval of length T, by the length T. The variance of the count in an interval for a Poisson process is equal to the mean count, so for the Poisson processes the standard deviation of the spike rate estimate from a counting interval of length T is

$$\frac{\sqrt{Var\{N(T)\}}}{T} = \frac{\sqrt{E\{N(T)\}}}{T} = \frac{\sqrt{\lambda T}}{T} = \sqrt{\frac{\lambda}{T}}, \tag{4.13}$$

where λ is spike rate of the process.

These plots show that LRD has a substantial effect on the variability of estimates of the spike rate, regardless of the length of the counting interval. For both these nominal rates, the standard deviation of the rate estimate for a one second counting interval when $H = 0.9$ (LRD) is twice that when $H = 0.5$ (no LRD), and this ratio increases to eleven for a counting interval of 100 seconds. If the LRD case is compared to a Poisson process, this ratio increases to about 2.6 at one second and about 15 at 100 seconds. The estimates for the non-LRD fGnDP are more variable than for the Poisson process due to the additional variability produced by the fGn.

An intuition for the practical effect of this excess variability can be gained from Table 4.1, where we have given examples of 95% confidence intervals for spike rate estimates based on counting intervals of different lengths for the fGnDP. These confidence intervals are based upon the assumption of a Gaussian distribution for the spike counts, which is a common assumption and has been justified in other studies (Young & Barta, 1986; Delgutte, 1987; Viemeister, 1988, e.g.). This is probably a good approximation for a nominal rate of 100 spikes per second and for a nominal rate of 5 spikes per second if the counting interval is long; otherwise, the lower limit of zero spikes will cause significant deviation from Gaussianity. We calculated each of these confidence intervals by finding the values at which the Gaussian distribution with mean λT, where λ is the nominal spike rate, and variance given by (4.8) is equal to 0.025 and 0.975 and dividing these values by T. The bounds on a 95% confidence interval are approximately two standard deviations

(a) fGnDP with Spike Rate of 5 sp/s

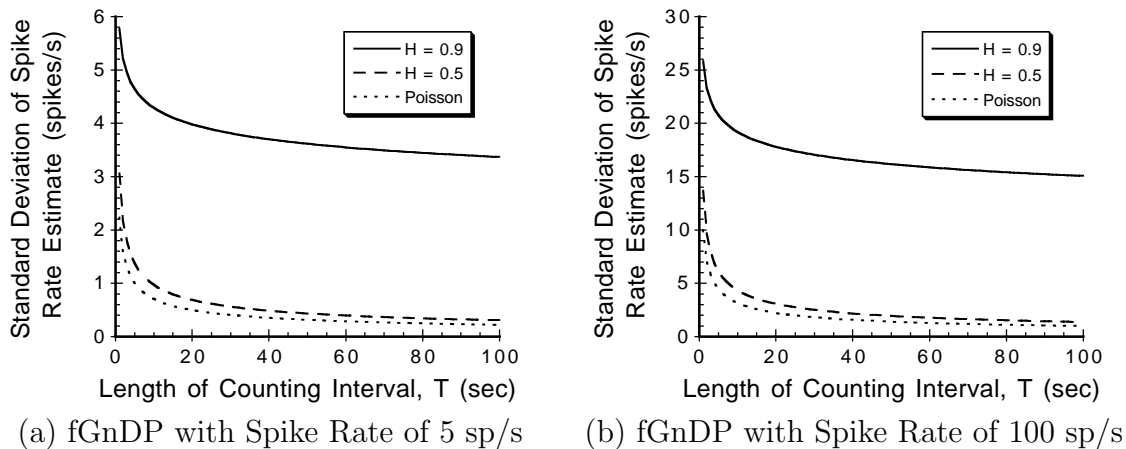(b) fGnDP with Spike Rate of 100 sp/s

Figure 4.7: The standard deviation of rate estimates calculated from single, long counting intervals for fGnDPs and Poisson processes with nominal rates $\lambda = 5$ spikes/s (a) and $\lambda = 100$ spikes/s (b), as a function of the counting interval length. Solid lines are data for fGnDPs with Hurst index $H = 0.9$; medium-dashed lines are data for fGnDPs with $H = 0.5$; and small-dashed lines are data for Poisson processes. The curves for the fGnDPs were calculated by dividing the square root of (4.8) by the counting interval length T. The parameter values for the fGnDPs were $\tau = 0.1$ s, either $H = 0.5$ or $H = 0.9$, and $\sigma = \sqrt{45} \approx 6.71$ spikes/s (a) or $\sigma = 30$ spikes/s (b). The curves for the Poisson processes were calculated from (4.13).

above and below the mean. When the spike rate is 5 spikes per second, the lower bounds of the confidence intervals for shorter counting intervals are negative; in these cases, the lower bounds are not given in the table since the spike rate must be nonnegative.

Both the size and the relatively slow reduction in this size with increasing counting interval length are troublesome. Even with a counting interval of 20 seconds, the length of these confidence intervals are about 13 spikes/s when the average spike rate is 5 spikes/s and 70 spikes/s when the average spike rate is 100 spikes/s. Furthermore, increasing the counting interval to 3600 seconds (one hour!) decreases the size of the confidence interval by less than half, to about 9 spikes/s and 40 spikes/s, respectively.

This variability in spike rate intervals may, in large part, explain the broad range of spontaneous rates observed within groups of primary auditory neurons that seem to be distinct based upon other characteristics. Based upon anatomical differences and an apparent multimodal distribution of spontaneous rates (SRs), Liberman (1982) suggested that auditory neurons can be divided into three classes: low SR neurons (SR < 0.5 spikes/s), medium SR neurons (0.5 < SR < 18 spikes/s), and high SR neurons (SR > 18 spikes/s). Our present results suggest that these classes may be much more distinct with respect to the underlying properties contributing to SR than our measurements of SR lead us to believe. In fact, it is possible that if we could measure the "true SR", we might only obtain one or a few values within each group.

Counting Interval	95% Confidence Interval at 5 spikes/sec	95% Confidence Interval at 100 spikes/sec	References
5 sec.	SR < 14.1	59.3 < SR < 140.7	Winter et al. (1990)
			Ohlemiller et al. (1991)
			Geisler et al. (1985)
10 sec.	SR < 13.4	62.4 < SR < 137.6	
15 sec.	SR < 13.0	64.0 < SR < 136.0	Liberman (1982, 1991, 1993)
20 sec.	SR < 12.8	65.1 < SR < 135.0	Relkin ... (1987, 1988, 1991)
			Jackson and Relkin (1998)
30 sec.	SR < 12.5	66.6 < SR < 133.4	Ryugo ... (1988, 1993)
50 sec.	SR < 12.0	68.3 < SR < 131.7	Wang and Sachs (1993)*
60 sec.	SR < 12.0	68.9 < SR < 131.1	Kiang et al. (1965)*
100 sec.	SR < 11.6	70.5 < SR < 129.5	
900 sec.	SR < 10.3	76.3 < SR < 123.7	
1800 sec.	0.1 < SR < 9.9	77.9 < SR < 122.0	
3600 sec.	0.4 < SR < 9.6	79.4 < SR < 120.6	

Table 4.1: Confidence intervals, based on the fGnDP model, of spontaneous spike rate (SR) estimates from different length counting intervals. Some representative studies of the auditory nerve that used certain length counting intervals to estimate (spontaneous) rate are listed in the rightmost column. The counting interval listed for a study marked by an asterisk is the minimum length that was used, but no further information on the distribution of the counting intervals is given. The parameters used were $H = 0.9$, $\tau = 0.1s$, $\sigma = \sqrt{45}$ for a spike rate of 5 spikes/s, and $\sigma = 30$ for a spike rate of 100 spikes/s. See the text for further explanation.

4.6.2 Estimation of Spike Rate from Multiple Repeated Counting Intervals

Rate-intensity functions and isointensity contours (or response areas), for tonal stimuli, are two other common measurements from auditory neurons. Both of these types of curves require many individual measurements of the mean spike rate of a neuron in response to a pure-tone stimulus of a particular frequency and a particular intensity. The difference between them lies in how the parameters of the stimulus are varied to obtain a single curve. For a single rate-intensity function, the intensity of the stimulus is varied while the frequency remains fixed, whereas for a single isointensity contour, the frequency of the stimulus is varied while the intensity remains fixed.

In either case, the typical experimental paradigm for estimating the mean spike rate for a single frequency-intensity combination is as described in Definition 4.4 and depicted in Figure 4.3. When estimating these spike rates, the tonal stimuli are on during the counting intervals (gray areas in Figure 4.3) and off at other times.[3] This

---

[3]Actually, the start of the counting interval is often delayed by a small amount with respect to the beginning of the stimulus in order to ignore the transitory response to the onset of the stimulus. The

Stimulus Duration	Repetition Interval	Number of Repetitions	References
500 ms	1.0 s	4–8	Geisler et al. (1974)
400 ms	1.5 s	∼4	Sachs and Abbas (1974)
200 ms	1.0 s	∼7[†]	Winslow and Sachs (1988)
100 ms	400 ms	∼10	Winter et al. (1990)
100 ms	150-981 ms	3,5,10	Ohlemiller et al. (1991)
100 ms	400 ms	5–20	Cooper and Yates (1994)
100 ms	1.0 s	10	Jackson and Relkin (1998)
50 ms	100 ms	10	Liberman (1978)
50 ms	200 ms	10–20	Winter and Palmer (1991)

Table 4.2: Parameters for some representative studies of rate-intensity functions or isointensity contours (for tonal stimuli) of primary auditory neurons. ([†]Actually only one counting interval was used at each intensity, but the rate-intensity functions were smoothed with a triangular weighting function that incorporated the spike counts for seven adjacent intensities.)

intermittent pattern is used in order to minimize the effects of adaptation of the neuron to the stimulus. The estimate of the mean spike rate is then given by $m_r(T)/T$.

In Table 4.2 the parameters used to measure either rate-intensity functions or isointensity contours in several representative studies are given. Stimulus duration (or counting interval), repetition interval, and number of repetitions in this table correspond, respectively, to the parameters T, S, and r in Definition 4.4 and Figure 4.3.

Table 4.3 contains 95% confidence intervals for the spike rate estimated by repeated intervals for fGnDPs with a nominal spike rates of 5 spikes per second and 100 spikes per second. It contains confidence intervals for ranges of counting interval lengths, repetition interval lengths, and numbers of repetitions that are representative of the ranges found in the studies listed in Table 4.2. In addition, we included data for 100 repetitions, which is significantly larger than the number of repetitions used in any of the studies mentioned in Table 4.2 but is still within the realm of practical possibility. We calculated each of the confidence intervals in this table by finding the values at which the Gaussian distribution with mean λT, where λ is the nominal spike rate, and variance (for the sample mean) given by (4.5) is equal to 0.025 and 0.975 and dividing these values by T. When the spike rate is 5 spikes per second, the lower bounds of the confidence intervals for all of the examples in this table are negative and are, therefore, not given.

The lengths of these confidence intervals are in the general neighborhood of those for the sample mean when only one long counting interval is used. Furthermore, their reduction with increasing counting interval, increasing repetition interval, or increasing number of repetitions is equally slow. Generally, for the parameters used in the studies shown in Table 4.2, 95% of spike rate estimates will cover an interval from zero up to 14

---

counting intervals are then shortened accordingly. This delay, however, which is only a few milliseconds at most, is very short relative to the length of the entire counting interval. Since it should, therefore, have very little effect on our results, we will not incorporate it into our calculations.

Counting Interval	Repetition Interval	Number of Repetitions	95% Confidence Interval at 5 spikes/sec	95% Confidence Interval at 100 spikes/sec
100 ms	150 ms	5	R < 17.4	44.5 < R < 155.5
100 ms	150 ms	10	R < 16.0	51.0 < R < 149.0
100 ms	150 ms	20	R < 14.9	55.9 < R < 144.0
100 ms	150 ms	100	R < 13.1	63.8 < R < 136.2
100 ms	400 ms	5	R < 16.8	47.4 < R < 152.6
100 ms	400 ms	10	R < 15.2	54.3 < R < 145.7
100 ms	400 ms	20	R < 14.1	59.3 < R < 140.7
100 ms	400 ms	100	R < 12.4	67.0 < R < 133.0
100 ms	1.0 s	5	R < 16.3	49.6 < R < 150.4
100 ms	1.0 s	10	R < 14.6	56.9 < R < 143.1
100 ms	1.0 s	20	R < 13.5	62.1 < R < 137.9
100 ms	1.0 s	100	R < 11.8	69.7 < R < 130.3
100 ms	1.5 s	5	R < 16.1	50.5 < R < 149.5
100 ms	1.5 s	10	R < 14.4	58.0 < R < 142.0
100 ms	1.5 s	20	R < 13.2	63.2 < R < 136.8
100 ms	1.5 s	100	R < 11.5	70.8 < R < 129.2
500 ms	1.0 s	5	R < 14.4	57.8 < R < 142.2
500 ms	1.0 s	10	R < 13.6	61.6 < R < 138.4
500 ms	1.0 s	20	R < 12.9	64.7 < R < 135.3
500 ms	1.0 s	100	R < 11.6	70.4 < R < 129.6
500 ms	1.5 s	5	R < 14.2	58.9 < R < 141.1
500 ms	1.5 s	10	R < 13.3	62.8 < R < 137.2
500 ms	1.5 s	20	R < 12.6	65.9 < R < 134.0
500 ms	1.5 s	100	R < 11.4	71.5 < R < 128.5

Table 4.3: Confidence intervals, based on the fGnDP model, of spike rate (R) estimates for repeated counting intervals. The parameters listed in the table are the counting interval T, the repetition interval S, and the number of repetitions r. The other parameters used were $H = 0.9$, $\tau = 0.1s$, $\sigma = \sqrt{45}$ for a spike rate of 5 spikes/s, and $\sigma = 30$ for a spike rate of 100 spikes/s. See the text for further explanation.

Stimulus Duration	Repetition Interval	Number of Repetitions	References
250 ms	250 ms	20	Geisler et al. (1985)
200 ms	1.0 s	40–100[†]	Young and Barta (1986)
51.2 ms	51.2 ms	1000	Teich and Khanna (1985)
204.8 ms	204.8 ms	250	Teich and Khanna (1985)
50 ms	120 ms[‡]	25	Delgutte (1987)
50 ms	200 ms	10–20	Winter and Palmer (1991)

Table 4.4: Parameters for some representative studies that attempted to estimate the variability of spike counts (or spike rates) of primary auditory neurons. ([†]Actually 80–100 *stimulus* repetitions were used, but data in "regions of noticeable background rate trends or fluctuations" were eliminated from calculations./ [‡]The repetition interval was not specified in this study. However, in two other approximately contemporaneous studies by the author (Delgutte, 1990a, 1990b), individual stimuli were repeated every 120 ms.)

to 16 spikes per second when the actual spike rate is 5 spikes per second. If there were no temporal dependence, these estimates would cover an interval about half this size or less. When the actual spike rate of the fGnDP model is 100 spikes per second, the 95% confidence interval extends from 50 to 60 spikes per second at the low end to 140 to 150 spikes per second at the high end for the parameters used in typical experimental studies. In this case, as well, if the LRD were not present, the confidence intervals would be half this size or less.

4.6.3 Estimation of the Variability of Spike Counts for Repetitions of a Stimulus

In order to assess how well the nervous system could, if it were (in some sense) optimal, distinguish stimuli based upon only the average spike rate of afferent neurons, it is not sufficient to have only a description of the relationship between average spike rate and the relevant stimulus parameters. The variability of the spike rate, or, more precisely, the variability of the spike count for a pertinent counting interval length, must also be known. The typical experimental paradigm for estimating this variability is identical to the repeated-stimulus paradigm described in Section 4.6.2, only the sample variance (see Definition 4.4) is used instead of the sample mean. The parameters used in several studies of primary auditory neurons to estimate the variance, or standard deviation, of spike counts are given in Table 4.4.

The results of Figure 4.6 suggest that LRD has little or no effect on the variability of these estimates for the numbers of repetition intervals used in these studies. However, we did see that LRD produces a marked negative bias in the sample variance and the sample standard deviation. Figure 4.8 indicates the expected extent of this bias, based on the fGnDP model, for the studies listed in Table 4.4. These graphs show the relationship between the variability of the spike count, actual or estimated, and the mean spike count for the fGnDP model. The graph for each study contains a curve for the actual standard

(a) Parameters from Geisler et al. (1985)

(b) Parameters from Young and Barta (1986)

(c) Parameters from Teich and Khanna (1985) (T = 51.2 ms)

(d) Parameters from Teich and Khanna (1985) (T = 204.8 ms)

(e) Parameters from Delgutte (1987)

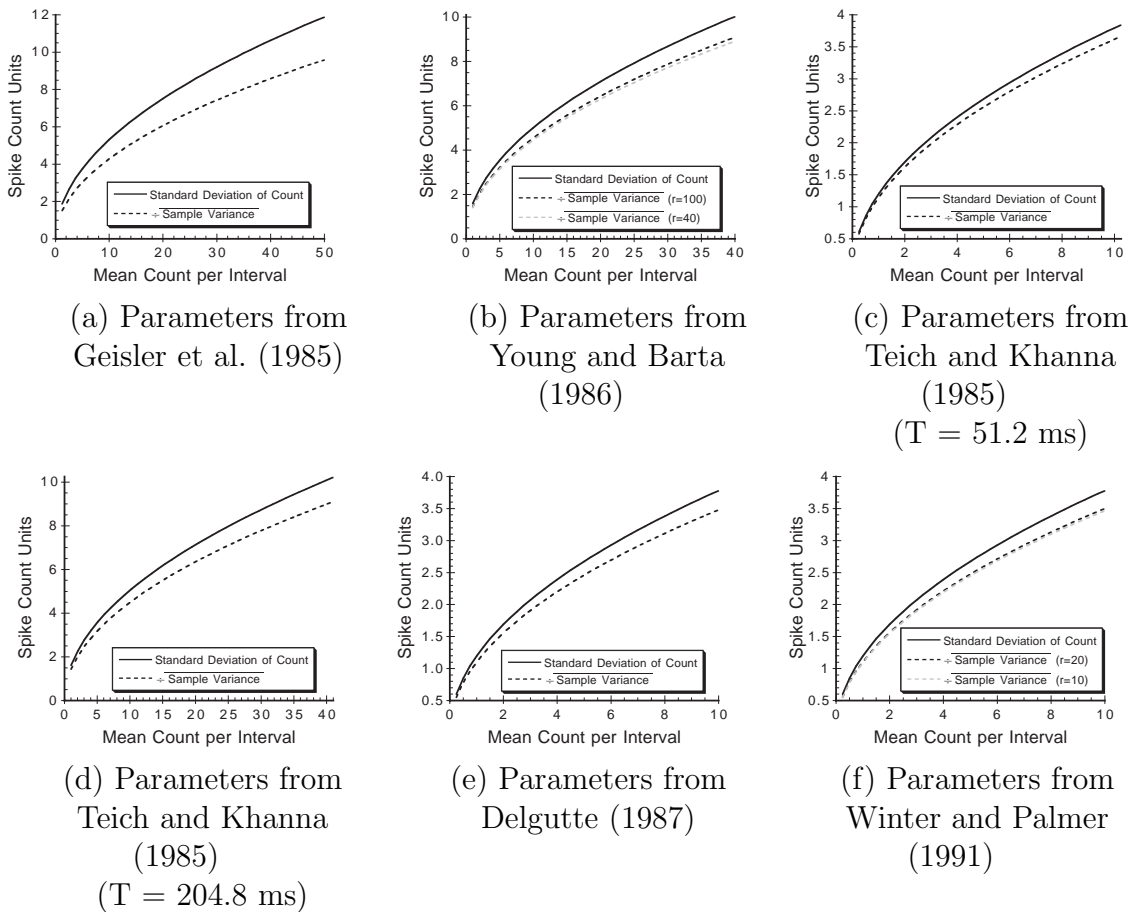(f) Parameters from Winter and Palmer (1991)

Figure 4.8: The bias of estimates of the spike count variance for the fGnDP model using parameters from the studies listed in Table 4.4. Each curve for the actual standard deviation of the count was calculated by taking the square root of (4.8) and is displayed as a solid line. Each curve for the square root of the expected value of the sample variance of the count was calculated by taking the square root of (4.11) and is displayed as a dashed line. The parameter values for the fGnDP were $H = 0.9$, $\tau = 0.1$ s, and $\sigma = \sqrt{9\lambda}$ spikes/s, where λ is the average spike rate. The counting interval length T, the repetition interval length S, and the number of repetitions r differ for each graph and are given in Table 4.4.

deviation of the spike count (solid line) and one or two curves for the expected values of the sample variance (dashed lines), transformed into units of "spikes" by taking the square root. The actual standard deviation was calculated as the square root of (4.8) for the length of counting interval, T, used in the given study. The curve associated with the expected value of the sample variance was calculated by taking the square root of (4.11).

In all of these graphs, the sample variance is noticeably biased. Although not insignificant, relative to the absolute values, the bias does not seem to completely undermine the utility of this estimator. Since it is constant with respect to the mean count, the ratio of the square root of the expected sample variance to the actual standard deviation can be used to make meaningful quantitative comparisons between the bias under the different conditions represented in Figure 4.8. The ratios for these graphs are 0.81 for Figure 4.8(a),

0.91 and 0.89 for Figure 4.8(b), 0.95 for Figure 4.8(c), 0.89 for Figure 4.8(d), 0.92 for Figure 4.8(e), and 0.93 and 0.92 for Figure 4.8(f). Thus, we would expect the most bias (by percentage) to be present in the measurements of Geisler et al. (1985). But, since they did not show any of these measurements, we can neither confirm nor disprove that their values for the spike count variance are significantly lower than would be expected based on the estimates from the other studies. The expected bias, in units of counts, for the remainder of the studies is between 5% and 11%.

We should note, however, that although we are measuring bias in "spikes" these numbers are descriptive of the bias in the sample *variance.* But, for the most part, the sample standard deviation is actually being used in the aforementioned studies in the calculations from which the results or conclusions are drawn. As we saw in Section 4.5, the sample standard deviation is more biased than the sample variance. Hence, our predictions of the bias in these studies will be smaller than the actual bias.

Two of the previously mentioned papers (Young & Barta, 1986; Winter & Palmer, 1991) contain scatter plots of their measurements of the sample standard deviation versus the mean spike count per counting interval (their Figures 6a and 5, respectively). In addition, a third paper (Delgutte, 1987) contains a plot of the average sample variance versus the mean spike count (Figure 6B). As we would expect, since our theoretical results do not incorporate any refractoriness, the magnitudes in Figure 4.8 are noticeably larger than those in these references. However, using the relative amount of bias in these graphs, we can speculate on the actual standard deviation or variance of the spike counts that they were attempting to estimate. Our analysis of the fGnDP model suggests that there was at least a 9% to 11% bias in the sample standard deviation in Young and Barta (1986). At their largest mean spike counts, approximately 25 to 30 spikes per interval, the average value of their sample standard deviation measurements was about 3.8, well below the values of 5 to 5.5 expected from a Poisson process. This suggests that the actual (average) standard deviation was at least $3.8/0.9 = 4.2$. This value is still below that expected of a Poisson process, but if their cloud of data were shifted accordingly, as if their measurements were being "unbiased", the upper data points would lie on or perhaps above the curve for the Poisson process. Since they used a much shorter counting interval, Winter and Palmer (1991) obtained standard deviation estimates that were well below those expected of a Poisson process. The spread of their estimates, however, was larger, in a relative sense, than that of Young and Barta (1986), and, thus, the largest of their estimates were equal to or above theoretical values for the Poisson process. Our analysis suggests that their bias was less than that of Young and Barta (1986), but it was still probably at least 7% or 8%. Nevertheless, such a bias for measurements in the range of one to three is quite small compared to the spread of their measurements. The bias in the measurements of Delgutte (1987), according to our estimates, was probably approximately the same as that of Winter and Palmer (1991), who used the same counting interval length, a longer repetition interval, but fewer repetitions. A rough comparison between the data from these two papers, which must be converted into the same units by squaring one or taking the square root of the other, reveals that the averages of their statistical measurements are probably very similar. Since Delgutte (1987) did not show raw data, but only averages, we can not make a comparison of the spreads of their data.

4.7   Discussion

Measuring the sample mean and sample variance of natural processes is so commonplace that they are often confounded with the actual "population" mean and variance. However, the ability of these statistics to produce good approximations to the quantities that they are estimating is substantially affected by certain properties of the process itself. For temporally dynamic processes, one of these properties is the dependence of their future states on their states in the present and the past. In the last half-century or so, a growing number of natural processes have been found to possess dependence on the past that decays particularly slowly (see Beran, 1994, for a good collection of examples). This property is usually called "long-range dependence" or "long-term memory" and can have significant detrimental effects on the behavior of sample statistics.

As in many other areas of experimental science, the sample mean and sample variance statistics are generally utilized to evaluate neurophysiological action-potential sequences (spike trains) without discussion of the expected quality of the calculated estimates. However, in the past ten to fifteen years, it has become apparent that the spike trains of many, if not most, neurons are long-range dependent (LRD) (see Section 4.2.3 for references). Although long-range dependence in neural spike trains has been investigated in a number of different studies, its repercussions on estimates of the mean and variance of spike counts has not been investigated. This, therefore, is what we have attempted to do in this study. In particular, we examined the effect of long-range dependence on the bias and variability of both the sample mean and the sample variance.

This necessitated the use of a model of neural spike trains. We chose to use the fractional-Gaussian-noise-driven Poisson process, which is long-range dependent and is a good model of neural spike trains, at least in sub-cortical systems (see Section 4.4.1 for references). In addition to the mean and variance of the counts in intervals of arbitrary length, we obtained approximate analytical expressions for the expected value and variability of the sample mean and for the expected value of the sample variance. Expressions for the variability of the sample variance and for the behavior of the sample standard deviation, however, are not available. These were investigated via simulations of the model.

The sample mean statistic is unbiased even in the presence of LRD. Furthermore, from our simulations, it also appears that LRD does not have much of a detrimental effect on the variability of the sample variance, at least for the range of parameters typical of neurophysiological recordings. Trends in this data suggest, however, that in other parameter ranges, for instance when a very large number of repeated counting intervals are employed, the variability of the sample variance statistic may be significantly amplified by the presence of LRD.

The variability of the sample mean statistic and the bias of the sample variance and sample standard deviation statistics are, nevertheless, negatively influenced by LRD. Both theory and simulations support the conclusion that the decay of the variability of the sample mean, as more data is included, is much slower in the presence of LRD than in its absence. This effect is such that, even based upon relatively moderate amounts of data, the standard deviation of the sample mean can be two to three times larger when LRD is present than in the absence of dependence. With the inclusion of more data, the ratio

between the standard deviation of the sample mean with and without LRD becomes even larger.

LRD also causes the sample variance and the sample standard deviation to underestimate the actual variance and standard deviation of spike counts. This bias was noticeable, but not exceedingly large in a relative sense. For typical parameters that are used in practice, the bias for the sample variance is usually between 5% and 11% when calculated in units of "spikes". From theoretical considerations, we know that the bias of estimates of the standard deviation will be even larger than this. Furthermore, simulations imply that the bias of the sample standard deviation will be about twice as large as this for small quantities of measurements, but that its additional bias will decrease rapidly as more data is incorporated into the statistical calculation.

The fact that the sample mean is so variable in the presence of LRD may suggest that the broad distribution of spontaneous rate measurements across auditory neurons is not due to a broad distribution in their properties, but is instead simply indicative of an inherent uncertainty in our measurements. In this case, it is possible that, in harmony with characterizations based on other measurable properties, there are only a few (possibly three) different "types" of auditory neurons, each with a distinct long-term spontaneous rate. The variability of this statistic also warrants caution in interpreting rate-intensity functions and isointensity contours on an absolute scale, since the position of the entire curve will be quite variable. In particular, this has implications for certain methods of measuring detection thresholds that are based upon such measurements.

The biases of the sample variance and the sample standard deviation do not seem to be as troublesome. However, one should be mindful of their presence when coming to conclusions based on these measurements. In addition, in combination with other forms of biasing, they may produce significant errors in calculations. One such form of biasing is the removal of "non-stationary" segments of spike trains (e.g. Young & Barta, 1986; Lansky & Radil, 1987; Kelly et al., 1996). LRD is known to produce apparent trends in stochastic processes that are actually stationary (see Beran, 1994, for a good discussion of this effect). Thus, it is not only possible, but probable, that a stationary LRD spike train will contain segments that appear to contain trends. Removal of these supposedly non-stationary segments might then further reduce estimates of the variability of counts from this spike train.

Chapter 5

Summary and Discussion

It has been approximately fourteen years since the discovery of long-range dependence in neural systems, albeit under different names, has been widely published (Gruneis et al., 1989; Teich, 1989).[1] During this time a relatively small number of studies have investigated this property of neural spike trains. However, for the most part, these investigations have been focused on revealing its existence, attempting to measure its strength, developing models, and fitting models to the basic trends of the statistics used to reveal LRD in neural recordings. A few studies have also suggested possible sources for this long-range dependence. But, the import of long-range dependence to both the way in which the nervous system operates and the way in which one investigates it is largely unknown. This gap was noted, with regard to physiological models of psychophysical results in the auditory system, in the review chapter of Delgutte (1996): "renewal process models might suffice for modeling most psychophysical tasks, at least until the significance of fractal effects [(i.e. long-range dependence)] is better elucidated." Unfortunately, since that time, little has been done to narrow this gap.

While entire books have been written on the subject of long-range dependence in stochastic time series (e.g. Beran, 1994; Doukhan, Oppenheim, & Taqqu, 2003), work on long-range dependence, or similar properties, of point processes seems relatively disjointed, to the extent that even the nomenclature associated with properties equivalent to long-range dependence varies significantly among different groups of researchers. Daley and his coworkers have more recently begun to form a mathematically sound theory of long-range dependence in point processes, but, being modest in size and published in an abstract mathematical context, this work is likely to be either unknown or inaccessible to a majority of neuroscientists.

In this thesis, we aimed to accomplish two general goals. First, we have attempted to form a more cohesive theory and understanding of long-range dependence in point processes, that is both well-connected with relevant mathematical literature and the larger literature on long-range dependence in stochastic time series. Second, we have applied this theory to some significant statistics-related questions in neuroscience. This work, we hope, is suggestive of areas within neuroscience where long-range dependence should be taken

---

[1]An earlier, but much more obscure, study (Wise, 1981), noted long-range-dependent-like properties in neurons as well.

into account and how this may be done.

5.1 Cortical Processing

The significance of Softky and Koch's (1992, 1993) challenge to the well-established and ubiquitous integrate-and-fire model is apparent in the long string of studies conducted in response to it. The strong reaction to their study is also indicative of a wide-spread and firmly established belief that the nervous system transmits and processes information based on a rate code. This is strikingly apparent in Shadlen and Newsome (1994). In this paper, the authors fervently argue for the theory that information is transmitted by neurons via their spike rate and that neural processing, in general, consists in integrating a large number of inputs, each with small influence on the integrating neuron. In countering the assertions of Softky and Koch to the contrary, Shadlen and Newsome proposed an answer to the dilemma that Softky and Koch raised that has since become a common assumption in cortical neurophysiology: that the amount of inhibitory input to a cell is tightly balanced with the amount of excitatory input.

Yet the assertions of both Softky and Koch (1992, 1993) and Shadlen and Newsome (1994) are based on the simplified integrate-and-fire model with Poisson process inputs and on the assumption that the coefficient of variation is an adequate measure of the variability of spike trains. Use of the simple model is understandable, since it facilitates analysis. However, conclusions based upon it would be more convincing if the analysis provided additional evidence that this model incorporates all of the critical dynamics of the modeled system that are relevant to the question being addressed. The coefficient of variation of interspike intervals alone does not do this.

In Chapter 2, we suggested this type of additional analysis. We proposed that the high-variability cortical neurons, originally only evaluated based on their interspike-interval coefficient of variation, be tested as well for the presence of long-range dependence, a property that, as we have shown in this thesis, is related statistically to the variability of a spike train. We then showed that a large portion of these models share a single shortcoming that undermines their ability to meet both the interspike-interval coefficient of variation and long-range dependence requirements. All of these models produce outputs that are renewal point processes, which must have infinite interspike-interval variability in order to possess long-range dependence. But this is inconsistent with the variability measured from actual cortical neurons.

In Chapter 3, we considered two other types of models that do not produce renewal-process outputs. Both of these models differed from the models considered in Chapter 2 in that their inputs were no longer Poisson processes. The first type of model had renewal-point-process inputs. We found that, with properly distributed interspike intervals at the input, this model could meet both the interspike interval variability and long-range dependence requirements. However, this success required that the inputs have infinite interspike-interval variability, which contradicts known properties of the inputs to cortical neurons. This, therefore, makes the renewal-process-input model fairly unconvincing as a model of cortical processing.

Hence, we suggested another model. In this model, the inputs were fractional-

Gaussian-noise-driven Poisson processes, which have many statistical properties in common with the inputs to cortex and, albeit to a lesser degree, cortical neurons themselves. This model was also able to produce interspike-interval variability and long-range dependence of a nature comparable to that found in cortical neurons. However, since the statistical character of its inputs better match those of real neurons, this model seems preferable to the renewal-process-input model.

These two models are possible solutions to the paradox raised by Softky and Koch (1992, 1993) and make the tight balance between excitation and inhibition suggested by Shadlen and Newsome (1994) unnecessary. Moreover, unlike the excitation-inhibition balance model, they possess long-range dependence like that found in cortical neurons. This suggests that the explanation of high cortical variability that these two models supply is in some ways superior to the balance model, and, therefore, that the assumption of a tight balance between excitation and inhibition is less credible than commonly thought.

The models of Chapter 3 were not suitable for complete mathematical analysis, and, hence, were analyzed on the basis of simulations. Thus, it was necessary to develop methods to discern different types of long-range dependence in point-process-like data. These methods, and the theoretical foundation developed around them, should be useful in future studies of long-range dependence in neural spike trains. Moreover, as we have shown the significance of long-range dependence to the study of cortical spike trains, particularly with regard to their variability, it is hoped that more such studies would be carried out.

5.2   Statistics of Spike Counts

While moments of spike counts are usually measured and used without hesitation, the analysis of Chapter 4 suggests that, in light of the long-range dependence present in many, or perhaps, most neurons, these statistics should be employed with a bit more caution. In particular, we considered the quality of the estimates obtained from the sample mean, sample variance, and sample standard deviation. We found that the sample mean is quite variable when long-range dependence is present and that this variability decays very slowly as more data is incorporated into the calculation. Furthermore, we showed that the sample variance and sample standard deviation underestimate, on average, the actual variance and standard deviation.

Through the use of the fractional-Gaussian-noise-driven Poisson process model, we were able to estimate the variability present in some representative spike rate measurements from the auditory nerve. These estimates suggest that such measurements should be interpreted with much caution. In particular, with regard to measurements of spontaneous rate, we provided evidence that the wide distribution of spontaneous rates could possibly be an artifact of the variability of the sample mean. If this is the case, then the "real" spontaneous rates of auditory neurons may be even more consonant with the distinct classification of neurons that is possible from other properties, such as response functions, morphology, and anatomical projections, than has been previously thought.

The negative bias of the sample variance and the sample standard deviation is, in

some ways, more disconcerting than the variability of the sample mean. At least for the sample mean, although the probability that it will differ from the true value is high, it will on average be correct. For the sample variance and and sample standard deviation, we expect, in the presence of long-range dependence, that the estimates will on average be incorrect. Fortunately, although not insignificant, the bias of these statistics is likely to be relatively small, in general, being on the order of ten percent (in units of spikes). Unfortunately, we do not have an analytical bound on the bias of the standard deviation, which is, whether explicitly stated or not, often the statistic that is used in studies involving the variability of spike counts, since it is in the same unit of measure as the mean. Thus, when studying the variability of spike counts, we suggest that one consider the effect that adding about ten or twenty percent to the standard deviation might have on conclusions drawn from the sample standard deviation (i.e. the square root of the sample variance).

5.3   The Theory of Rate Coding

One of the underlying motivations to the studies in this thesis was to provide evidence against the assertion that information is transmitted and processed through the spike rates of neurons. Though the nervous system may use spike rate for some more global background information, this method of transmitting information seems far too slow and inefficient to transmit and process the immense amount of information that it must handle in a rapid manner. As an example, Thorpe and his coworkers (Thorpe, Fize, & Marlot, 1996; Thorpe & Gautrais, 1997; Gautrais & Thorpe, 1998; Thorpe, Delorme, & Van Rullen, 2001; Van Rullen & Thorpe, 2001) have provided evidence of this in the visual system. They concluded, based on event-related potentials in humans performing a categorization task, that the visual system is able to process complex visual stimuli in under 150 milliseconds. They argue that, given the number of synapses through which the information must pass, that the nervous system would require either "excessively long observation periods incompatible with the speed of sensory processing or excessively large numbers of redundant neurones, incompatible with the anatomical constraints imposed by sensory pathways" (Gautrais & Thorpe, 1998) to make use of a rate code. The constraints can be met in a physiologically realistic way, however, by certain "temporal codes" (e.g. Thorpe et al., 2001; Van Rullen & Thorpe, 2001).

On the whole, our results have not convincingly refuted the theory of rate coding. The integration mechanism in the integrate-and-fire neuron model is representative of the types of mechanisms that would be necessary to process rate-encoded information. Although in Chapter 2 we found that many of these models were incompatible with the combination of high variability and long-range dependence found in cortical neurons, in Chapter 3 we presented examples of such models that are successful in producing these two effects. The results of Chapter 4 suggest that averaging over long intervals is very inefficient in the presence of long-range dependence. However, to the extent that the nervous system can combine the output of many neurons on a short time scale, it may be capable of transmitting information by means of average spike rates.

Appendix A

The Derivative of the Hazard Rate for the Positive Gaussian Distribution

We wish to prove that the hazard function for the positive Gaussian distribution is monotone increasing. Recall that the positive Gaussian probability density function is

$$f(t) = \frac{2}{\pi \, \mu} \, e^{-\frac{t^2}{\pi \, \mu^2}},$$

where $\mu > 0$ is the expected value. Thus, the (cumulative) distribution function is

$$F(t) = \int_0^t f(s) \, ds = \mathrm{erf}\left(\frac{t}{\mu \, \sqrt{\pi}}\right).$$

We will also require, below, the derivative of the probability density function

$$f'(t) = -\frac{4 \, t}{\pi^2 \, \mu^3} \, e^{-\frac{t^2}{\pi \, \mu^2}}.$$

The hazard function of any distribution $F(t)$ is given by

$$z(t) = \frac{f(t)}{1 - F(t)}.$$

Thus, its derivative is

$$\frac{d}{dt} \, z(t) = \frac{f^2(t) + f'(t) \left(1 - F(t)\right)}{\left[1 - F(t)\right]^2}.$$

Thus, in order to determine the sign of the derivative of the hazard function, we can restrict ourselves to the finding the sign of the numerator in this expression, as long as $F(t) < 1$ for all t. For the positive Gaussian distribution, we have

$$f^2(t) + f'(t)\left[1 - F(t)\right] = \frac{4}{\pi^2 \, \mu^2} \, e^{-\frac{t^2}{\pi \, \mu^2}} \left[e^{-\frac{t^2}{\pi \, \mu^2}} - \frac{t}{\mu} \, \mathrm{erfc}\left(\frac{t}{\mu \, \sqrt{\pi}}\right)\right].$$

Thus, the sign of the derivative of the hazard function is the same as the sign of the expression

$$e^{-\frac{t^2}{\pi \mu^2}} - \frac{t}{\mu} \operatorname{erfc}\left(\frac{t}{\mu \sqrt{\pi}}\right).$$

In order to prove that the hazard function is monotone increasing or decreasing, we must show that this expression is either positive for all $t > 0$ or negative for all $t > 0$. By substituting $y = t/\mu$, we find that this problem is equivalent to showing that the expression

$$e^{-\frac{y^2}{\pi}} - y \operatorname{erfc}\left(\frac{y}{\sqrt{\pi}}\right). \tag{A.1}$$

is always positive or always negative for $y > 0$.

We begin with the inequality

$$\begin{aligned}
-\frac{d}{dx}\frac{1}{x}e^{-\frac{x^2}{\pi}} &= -\left[-\frac{1}{x^2}e^{-\frac{x^2}{\pi}} + \frac{1}{x}e^{-\frac{x^2}{\pi}}\left(-\frac{2}{\pi}x\right)\right] \\
&= \frac{1}{x^2}e^{-\frac{x^2}{\pi}} + \frac{2}{\pi}e^{-\frac{x^2}{\pi}} \\
&> \frac{2}{\pi}e^{-\frac{x^2}{\pi}}.
\end{aligned}$$

Integrating both sides of this inequality from y to ∞, we get

$$\int_y^\infty \frac{2}{\pi}e^{-\frac{x^2}{\pi}}\,dx = \frac{2}{\sqrt{\pi}}\int_{\frac{y}{\sqrt{\pi}}}^\infty e^{-t^2}\,dt = \operatorname{erfc}\left(\frac{y}{\sqrt{\pi}}\right)$$

and

$$\int_y^\infty -\frac{d}{dx}\frac{1}{x}e^{-\frac{x^2}{\pi}}\,dx = \frac{1}{y}e^{-\frac{y^2}{\pi}}.$$

Thus, we have that

$$\frac{1}{y}e^{-\frac{y^2}{\pi}} > \operatorname{erfc}\left(\frac{y}{\sqrt{\pi}}\right),$$

or, equivalently, that

$$e^{-\frac{y^2}{\pi}} > y \operatorname{erfc}\left(\frac{y}{\sqrt{\pi}}\right),$$

for all $y > 0$. Thus, expression (A.1) is always positive (when $y > 0$), and, hence, the hazard rate of the positive Gaussian distribution is monotone increasing.

Appendix B

FFCs and IDCs from Integrate-and-Fire Model
Simulations

B.1 Integrate-and-Fire Model with Positive-Gaussian-Distributed Inputs
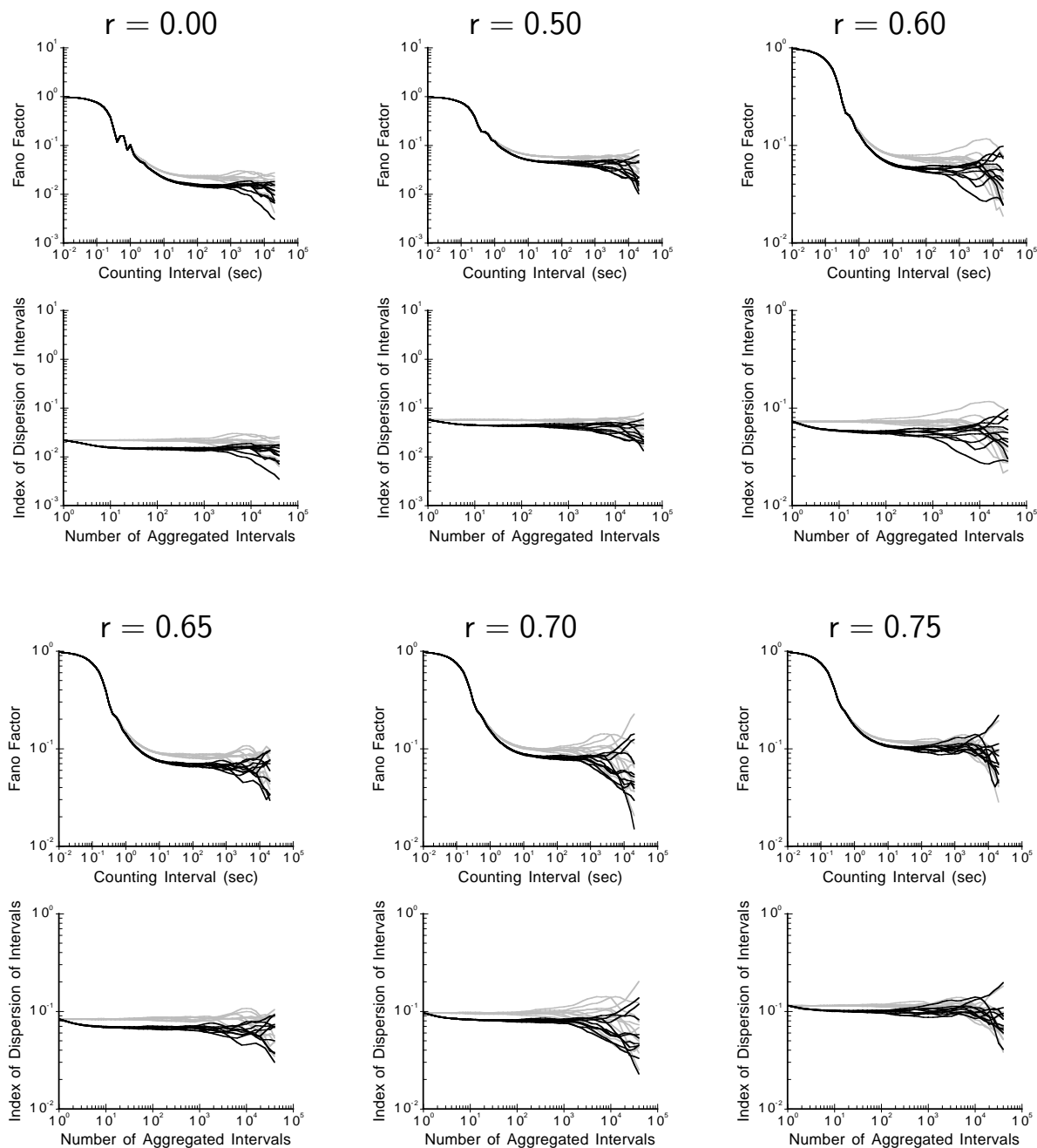
Figure B.1: FFCs and IDCs estimated from simulations of the IF model with positive-Gaussian inputs. Each set of axes contains ten curves calculated from original data (black) and ten curves calculated from the corresponding shuffled surrogate data (gray). For each value of the inhibition-excitation ratio r, each inidividual FFC in the left set of axes was calculated from the same data as one of the IDCs in the right set of axes.

Figure B.1: FFCs and IDCs estimated from simulations of the IF model with positive-Gaussian inputs (cont.)
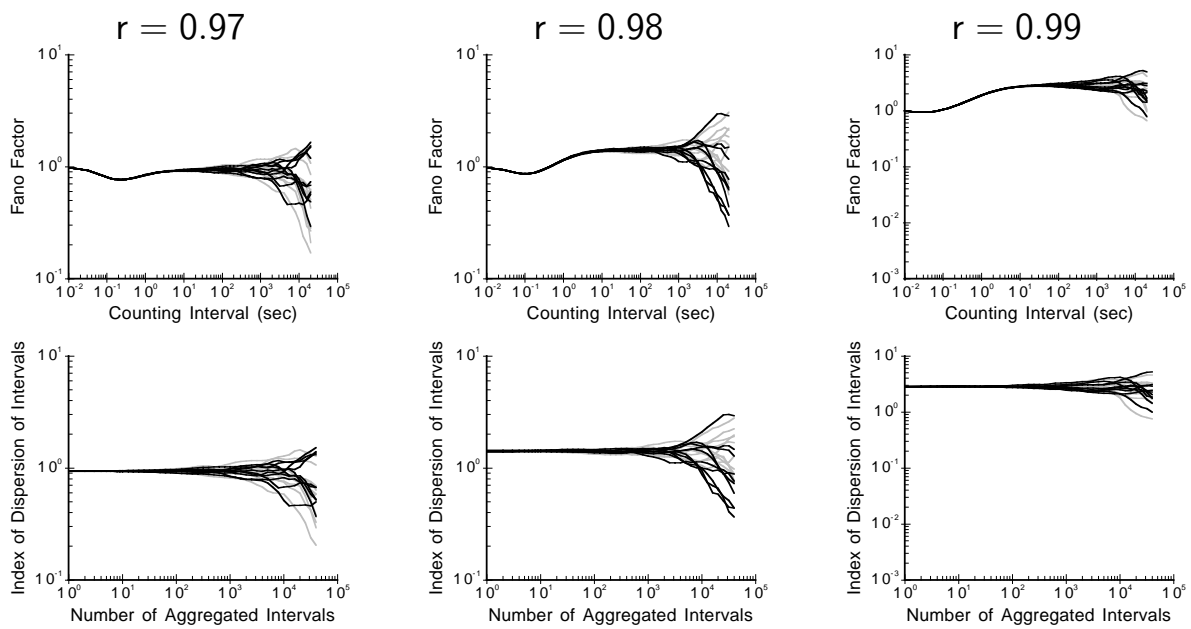
Figure B.1: FFCs and IDCs estimated from simulations of the IF model with positive-Gaussian inputs (cont.)

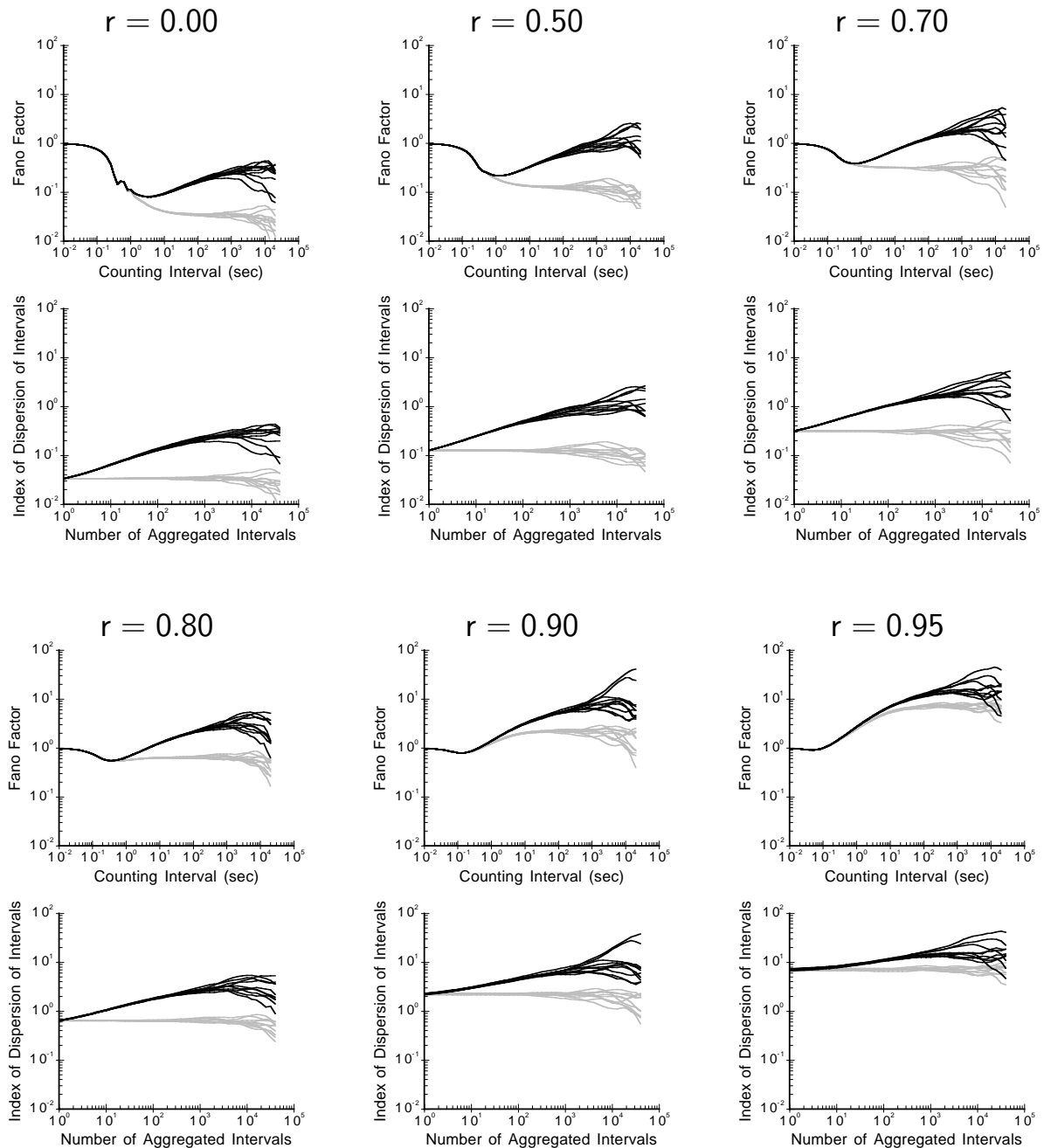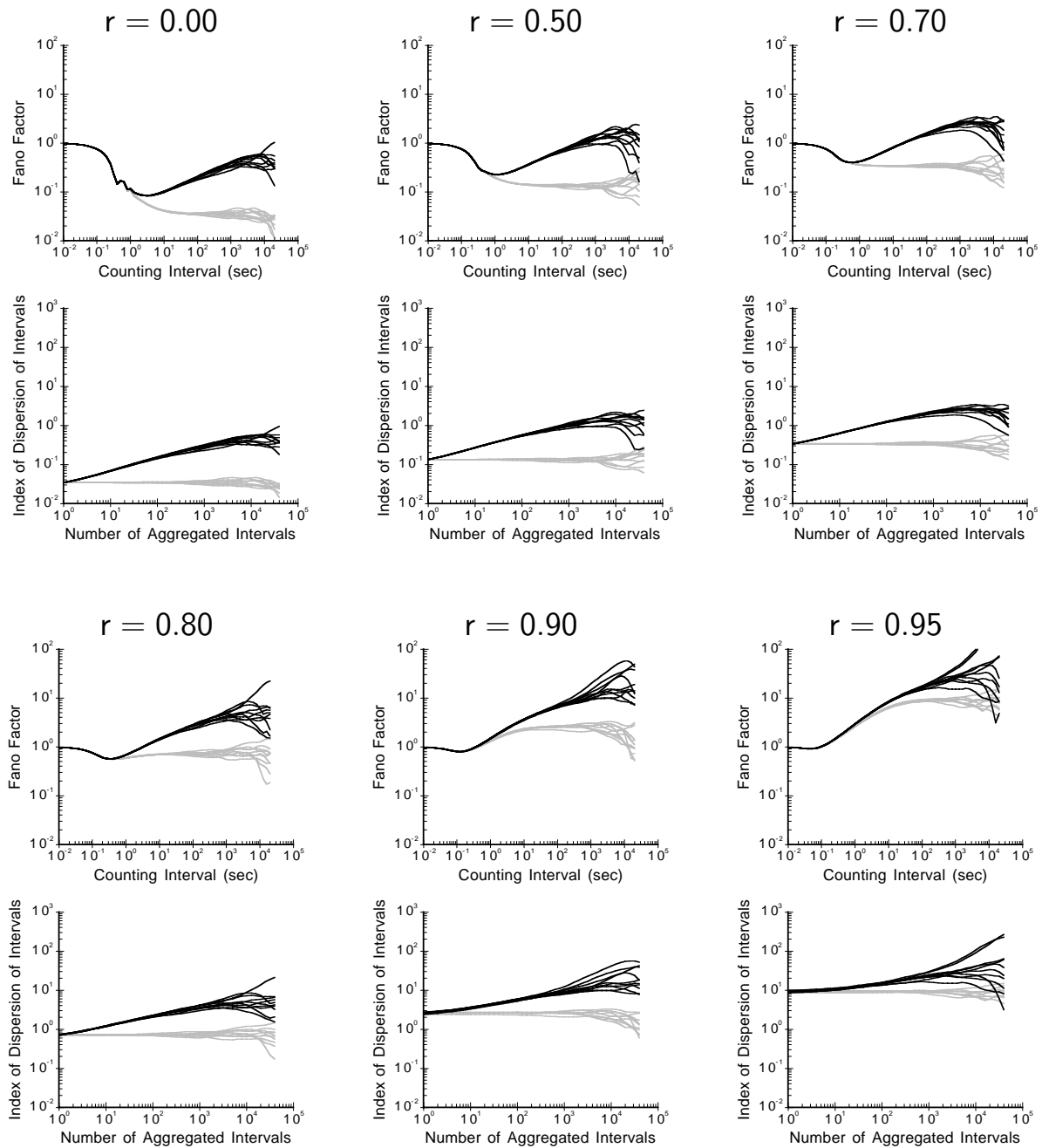## B.2 Integrate-and-Fire Model with Pareto-Distributed Inputs

$$\underline{\alpha = \mathbf{3.0}}$$
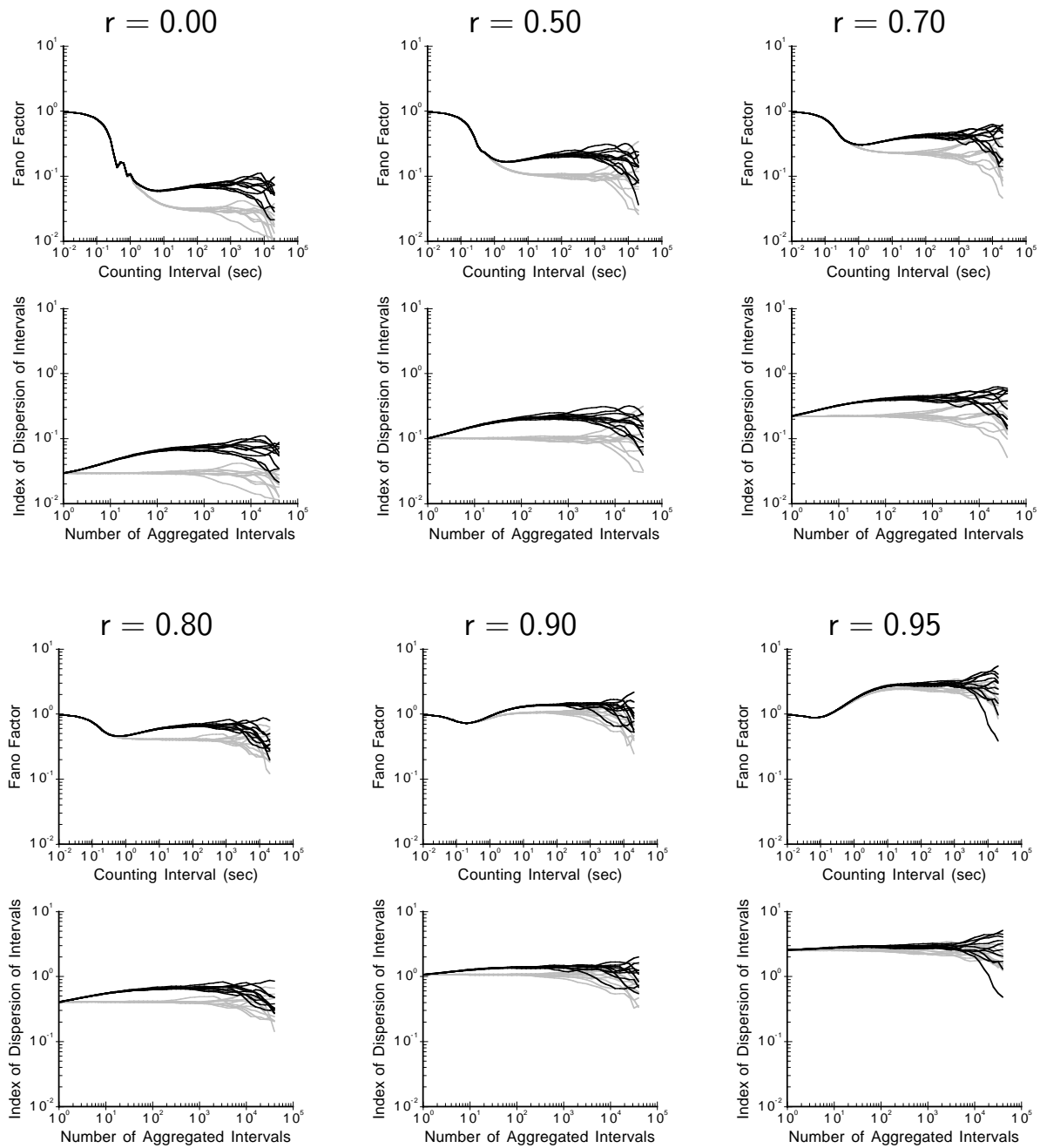
Figure B.2: FFCs and IDCs estimated from simulations of the IF model with Pareto inputs and parameter $\alpha = 3.0$. Each set of axes contains ten curves calculated from original data (black) and ten curves calculated from the corresponding shuffled surrogate data (gray). For each value of the inhibition-excitation ratio r, each inidividual FFC in the left set of axes was calculated from the same data as one of the IDCs in the right set of axes.
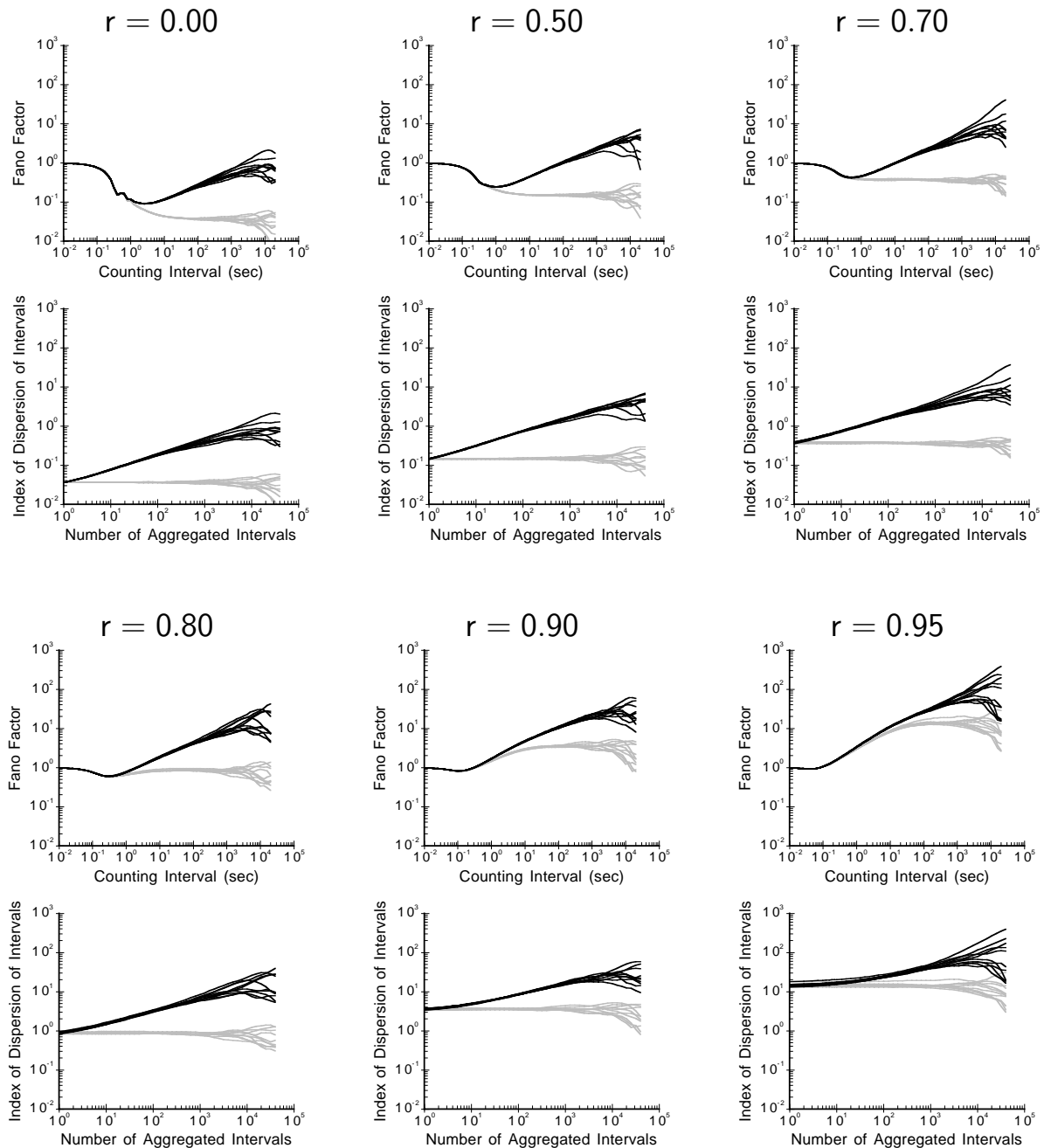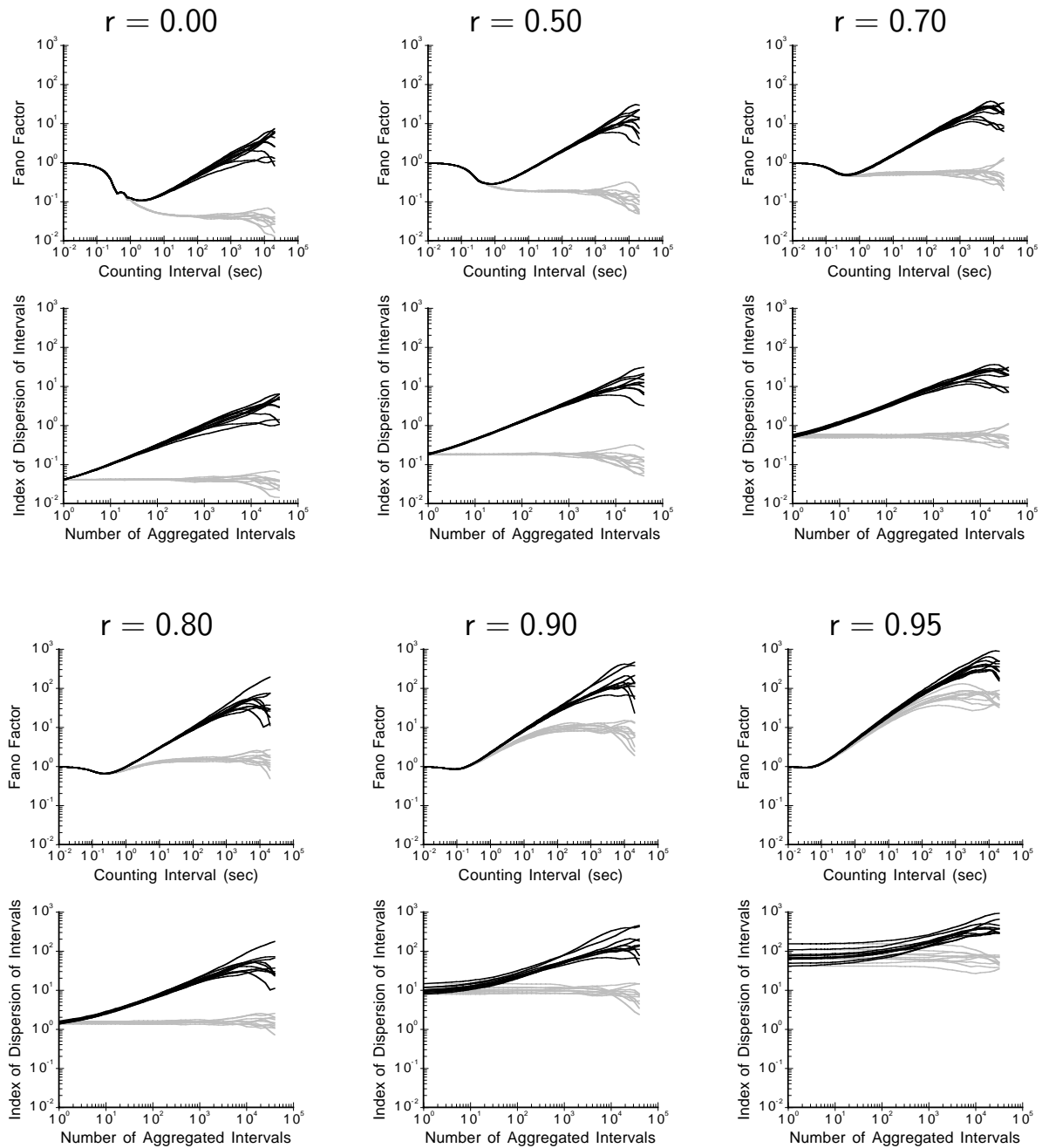
$\alpha = 2.5$
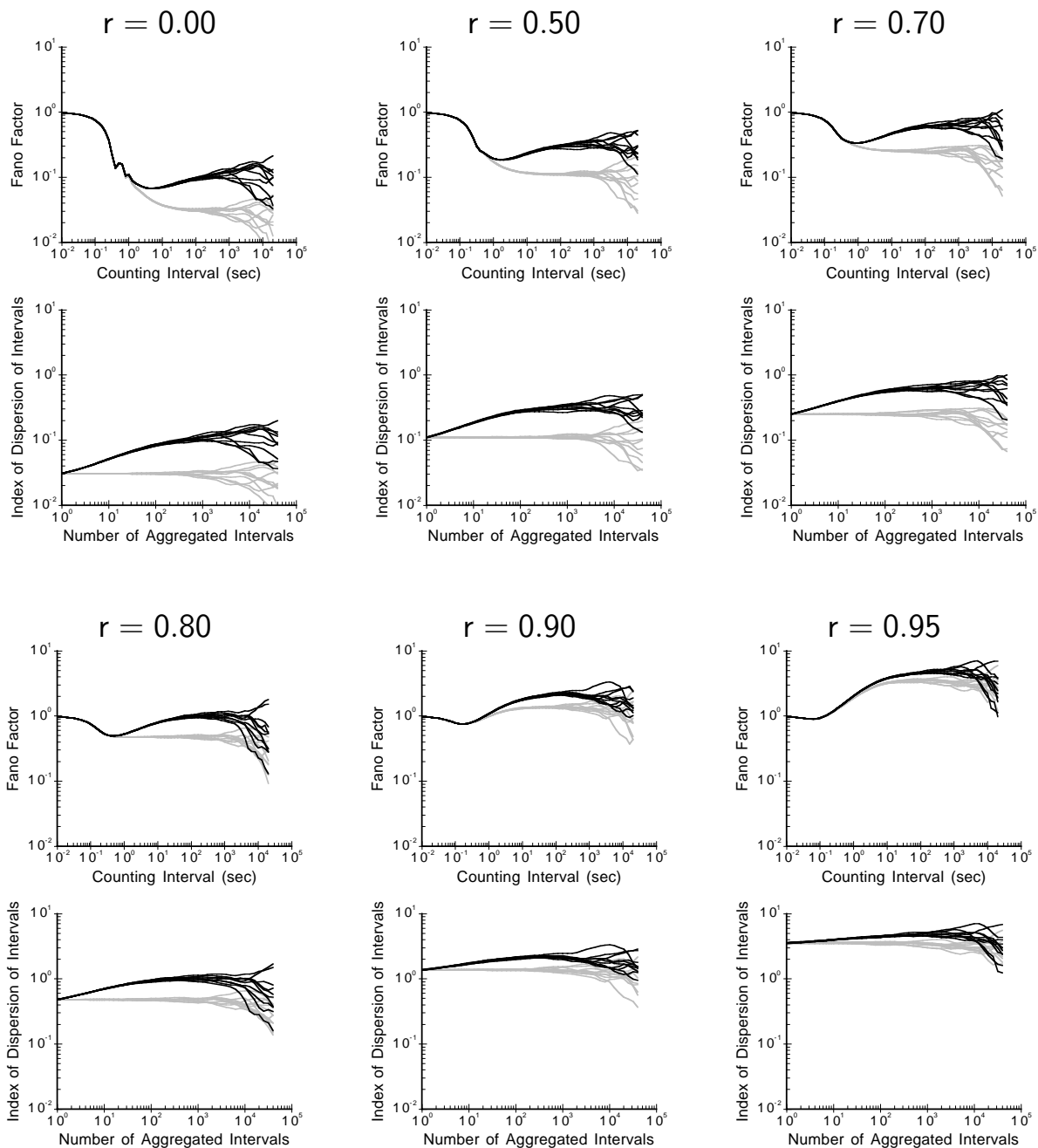
Figure B.3: FFCs and IDCs estimated from simulations of the IF model with Pareto inputs and parameter $\alpha = 2.5$. Same format as Figure B.2.

$$\underline{\alpha = 2.1}$$
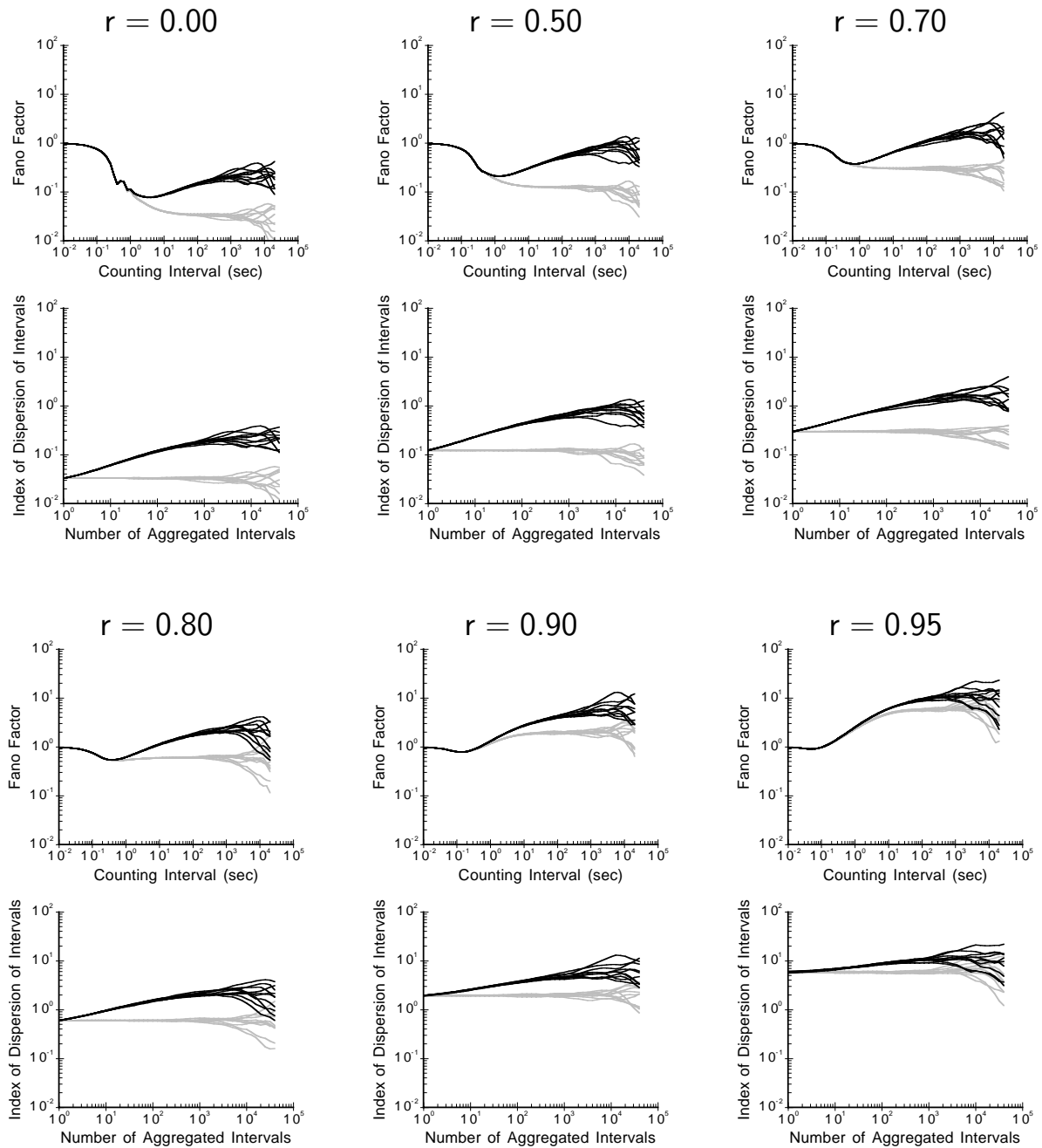
Figure B.4: FFCs and IDCs estimated from simulations of the IF model with Pareto inputs and parameter $\alpha = 2.1$. Same format as Figure B.2.

$$\underline{\alpha = \mathbf{2.0}}$$

Figure B.5: FFCs and IDCs estimated from simulations of the IF model with Pareto inputs and parameter $\alpha = 2.0$. Same format as Figure B.2.

$$\underline{\alpha = \mathbf{1.9}}$$

Figure B.6: FFCs and IDCs estimated from simulations of the IF model with Pareto inputs and parameter $\alpha = 1.9$. Same format as Figure B.2.
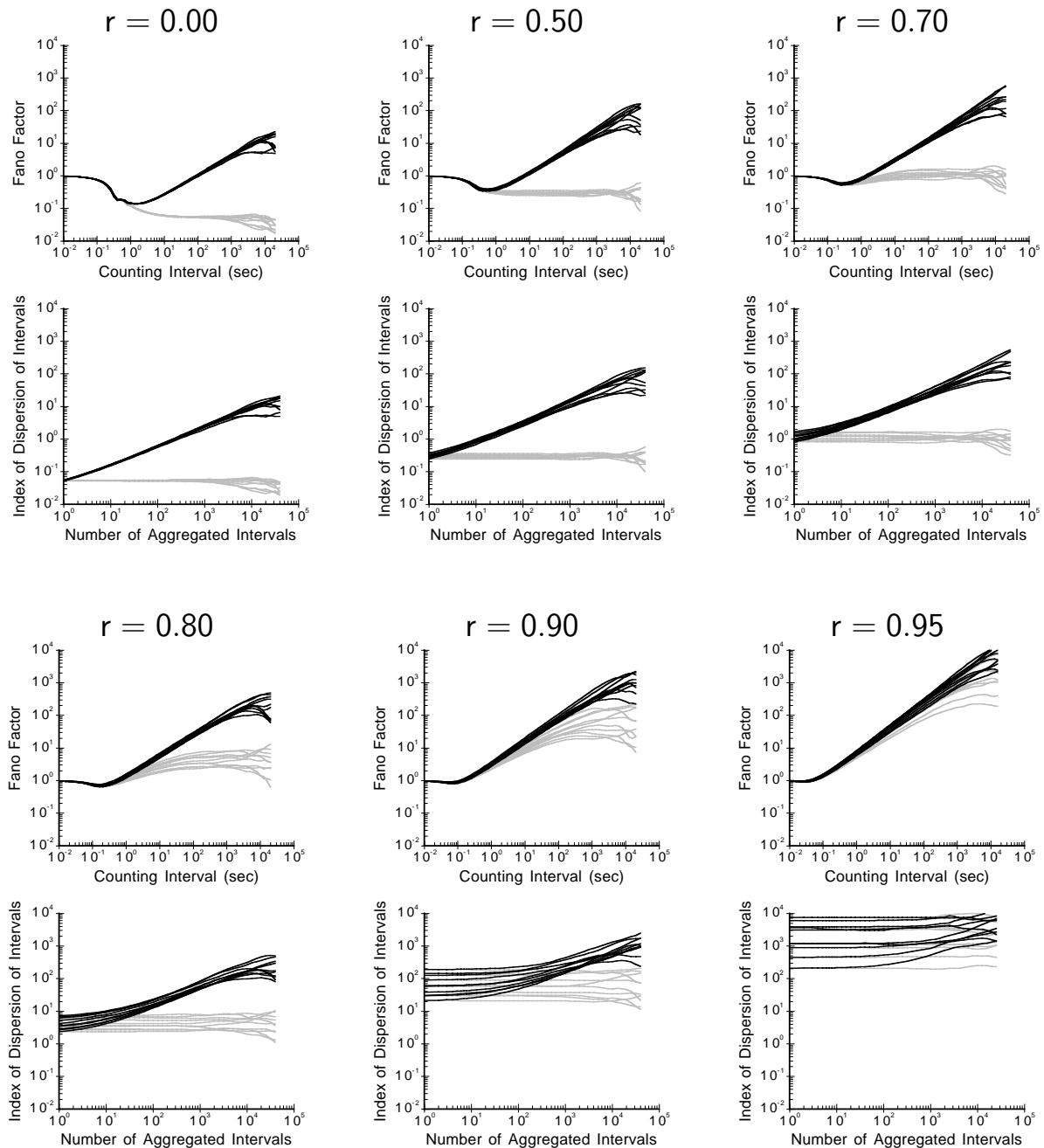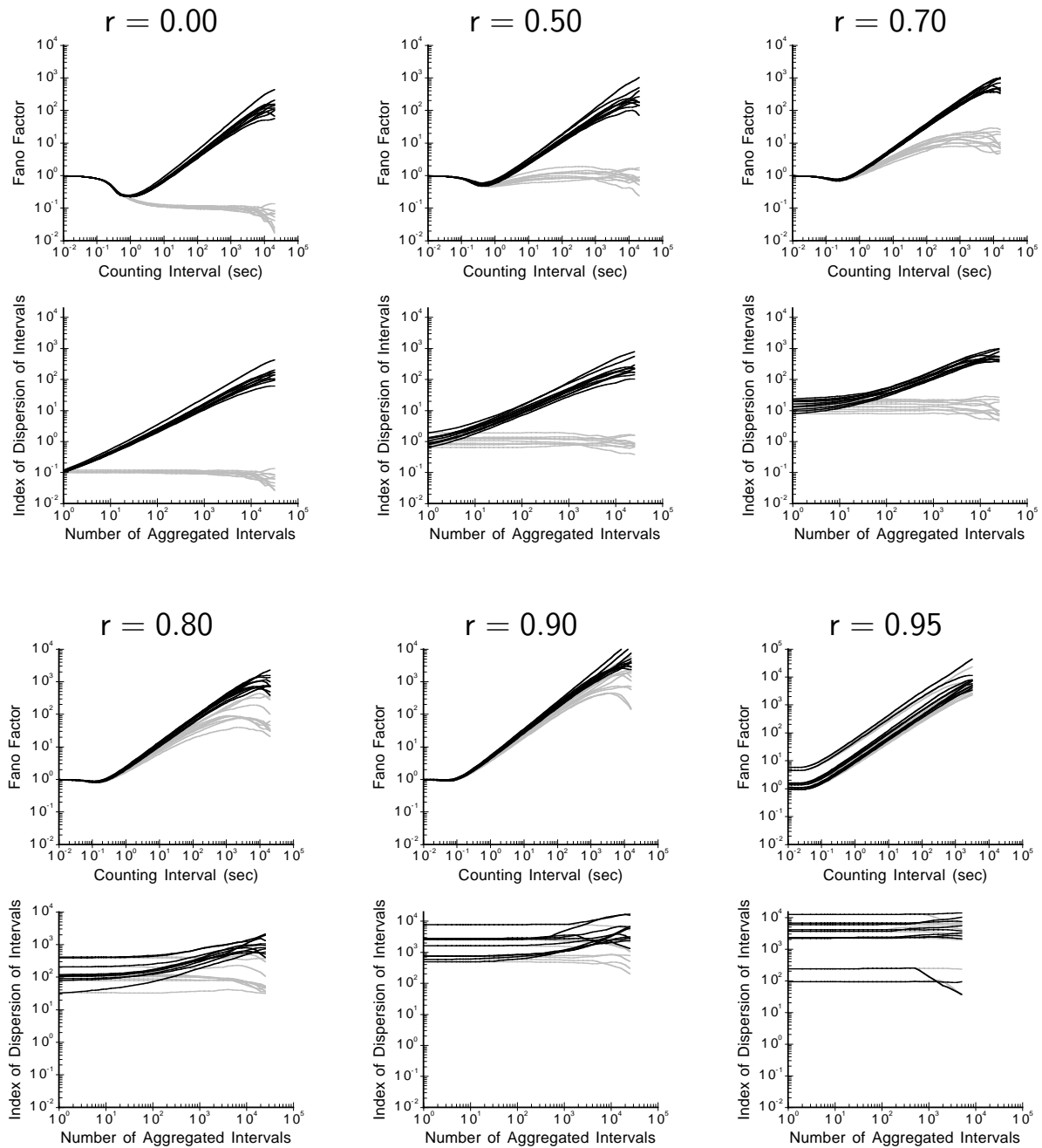
Figure B.7: FFCs and IDCs estimated from simulations of the IF model with Pareto inputs and parameter $\alpha = 1.75$. Same format as Figure B.2.

$$\underline{\alpha = \mathbf{1.5}}$$

Figure B.8: FFCs and IDCs estimated from simulations of the IF model with Pareto inputs and parameter $\alpha = 1.5$. Same format as Figure B.2.

$$\underline{\boldsymbol{\alpha = 1.25}}$$

Figure B.9: FFCs and IDCs estimated from simulations of the IF model with Pareto inputs and parameter $\alpha = 1.25$. Same format as Figure B.2.

$\alpha = 1.0$

Figure B.10: FFCs and IDCs estimated from simulations of the IF model with Pareto inputs and parameter $\alpha = 1.0$. Same format as Figure B.2.

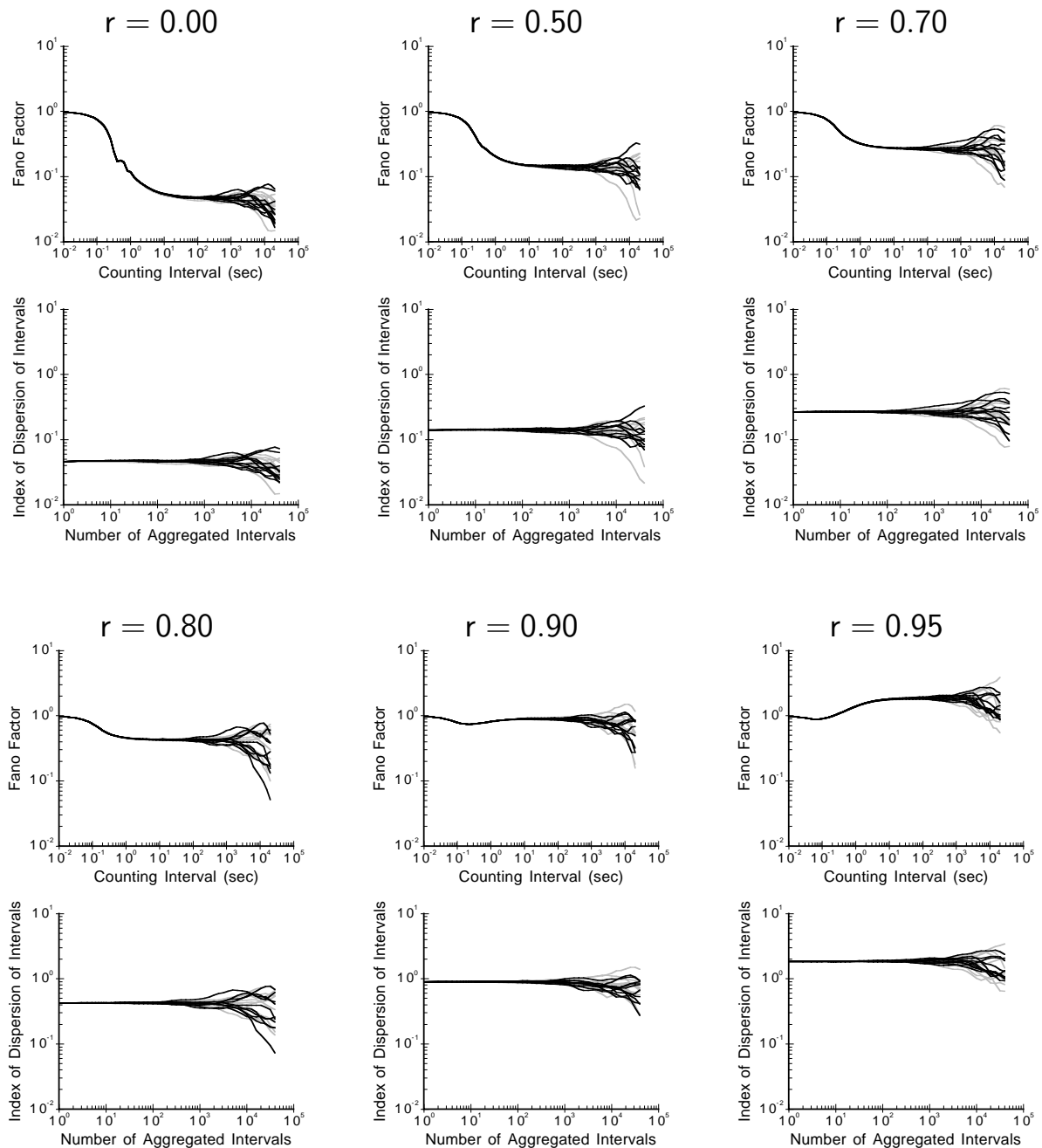## B.3 Integrate-and-Fire Model with fGnDP-Distributed Inputs

H = 0.50

Figure B.11: FFCs and IDCs estimated from simulations of the IF model with fGnDP inputs and parameter $H = 0.50$. Each set of axes contains ten curves calculated from original data (black) and ten curves calculated from the corresponding shuffled surrogate data (gray). For each value of the inhibition-excitation ratio r, each inidividual FFC in the left set of axes was calculated from the same data as one of the IDCs in the right set of axes.

$\underline{H = 0.55}$

Figure B.12: FFCs and IDCs estimated from simulations of the IF model with fGnDP inputs and parameter $H = 0.55$. Same format as Figure B.11.
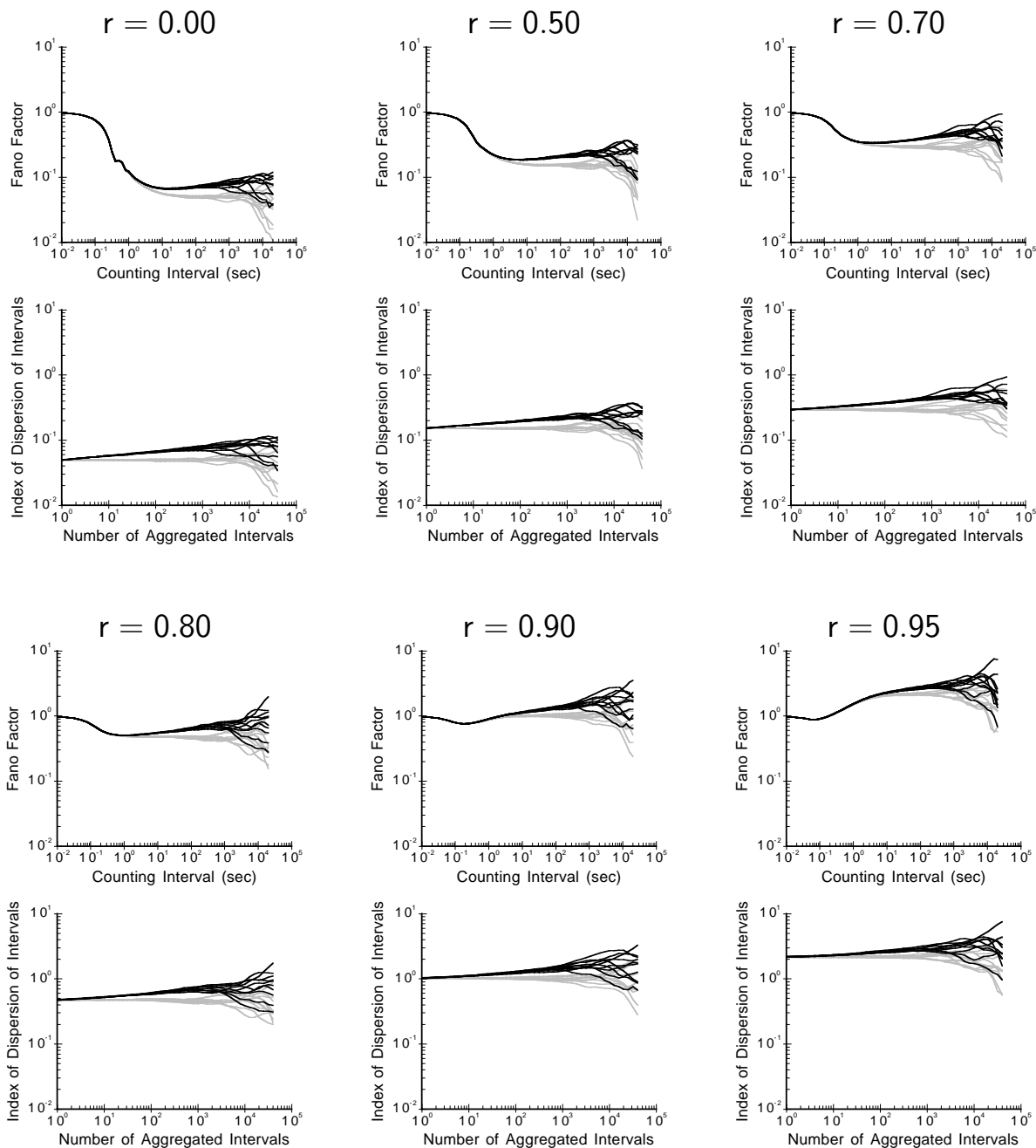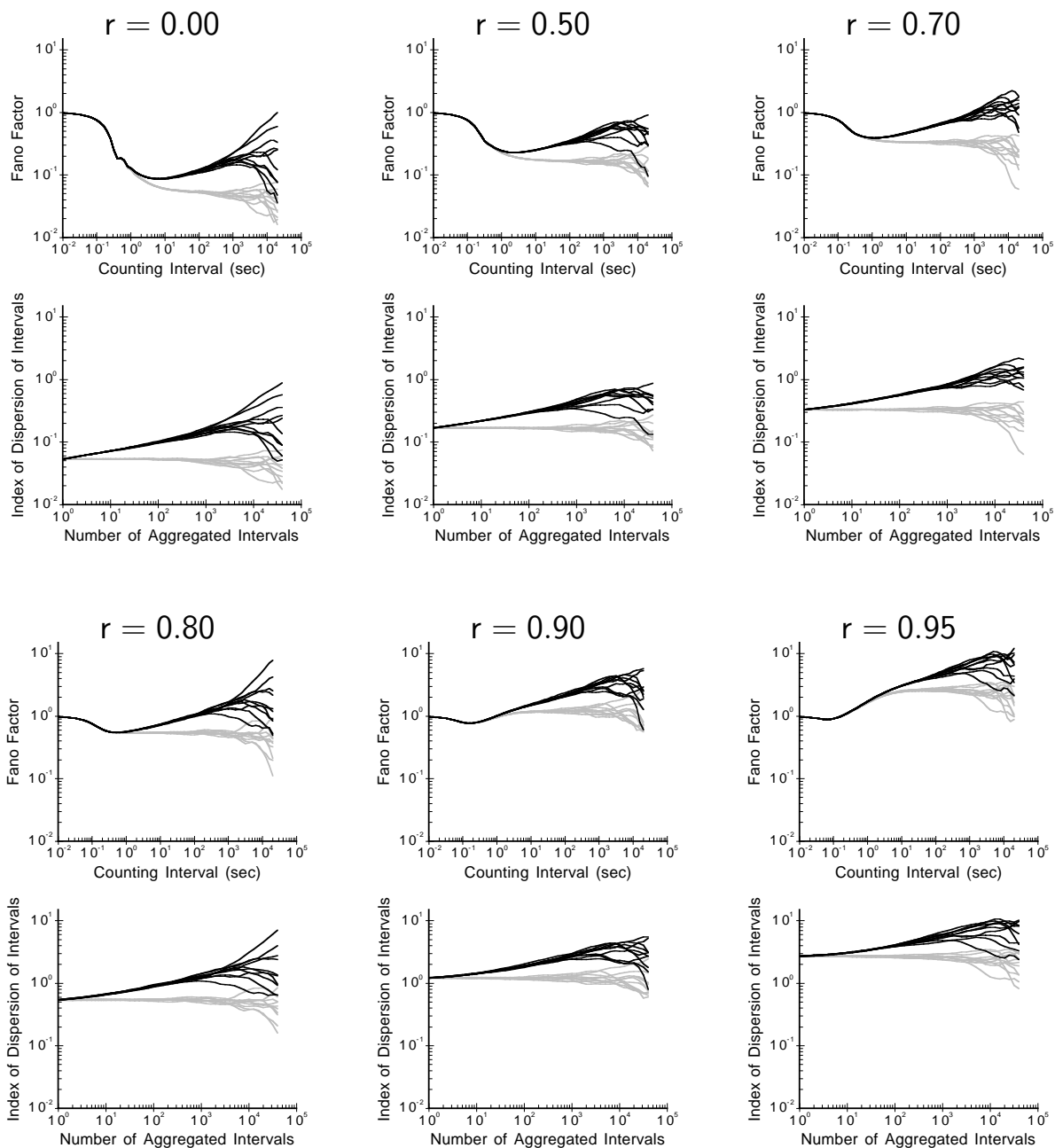
$\underline{H = 0.60}$

Figure B.13: FFCs and IDCs estimated from simulations of the IF model with fGnDP inputs and parameter $H = 0.60$. Same format as Figure B.11.

H = 0.65

Figure B.14: FFCs and IDCs estimated from simulations of the IF model with fGnDP inputs and parameter $H = 0.65$. Same format as Figure B.11.

H = 0.70

Figure B.15: FFCs and IDCs estimated from simulations of the IF model with fGnDP inputs and parameter $H = 0.70$. Same format as Figure B.11.
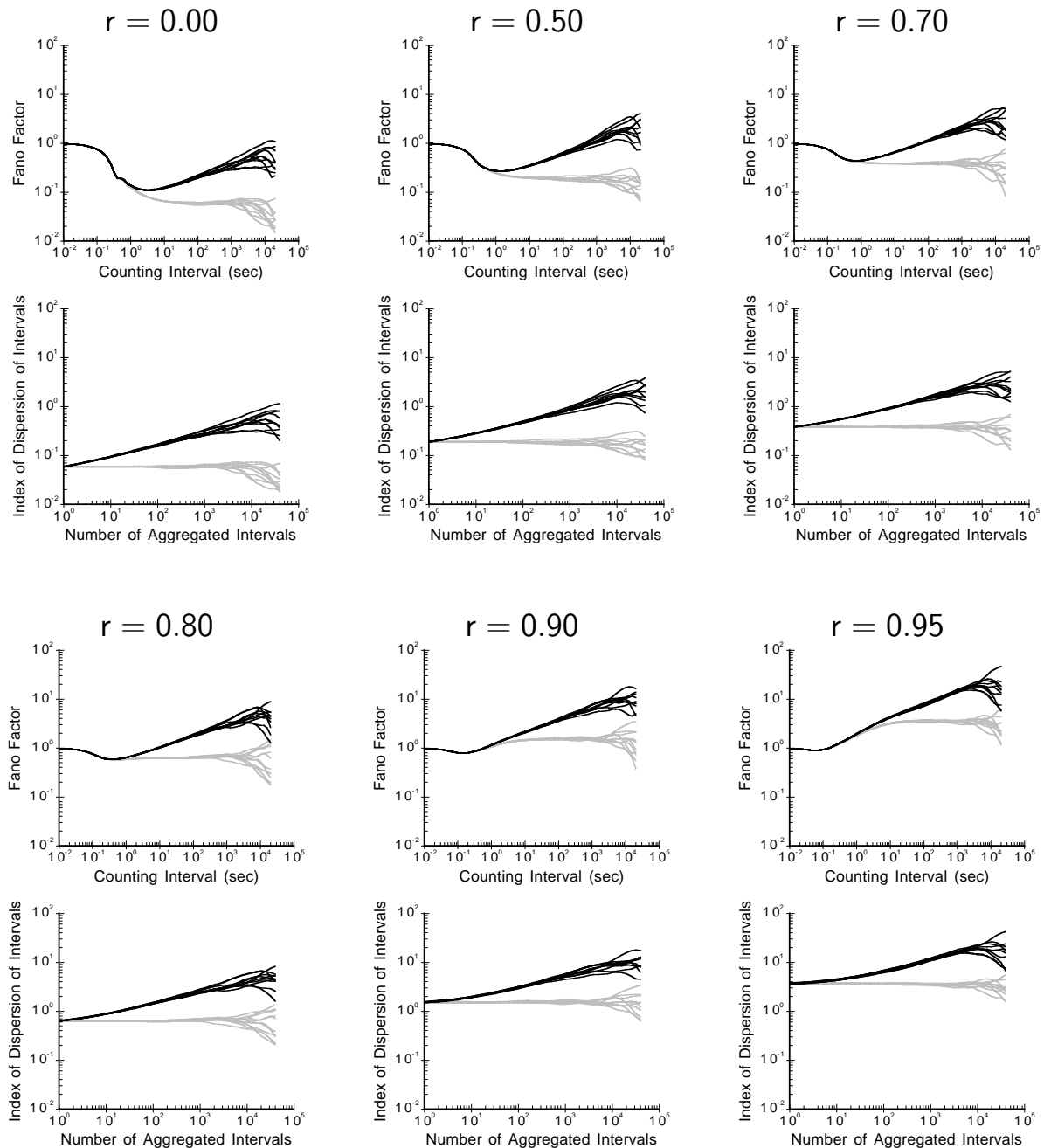
$H = 0.75$

Figure B.16: FFCs and IDCs estimated from simulations of the IF model with fGnDP inputs and parameter $H = 0.75$. Same format as Figure B.11.

H = 0.80

Figure B.17: FFCs and IDCs estimated from simulations of the IF model with fGnDP inputs and parameter $H = 0.80$. Same format as Figure B.11.

H = 0.85

Figure B.18: FFCs and IDCs estimated from simulations of the IF model with fGnDP inputs and parameter $H = 0.85$. Same format as Figure B.11.
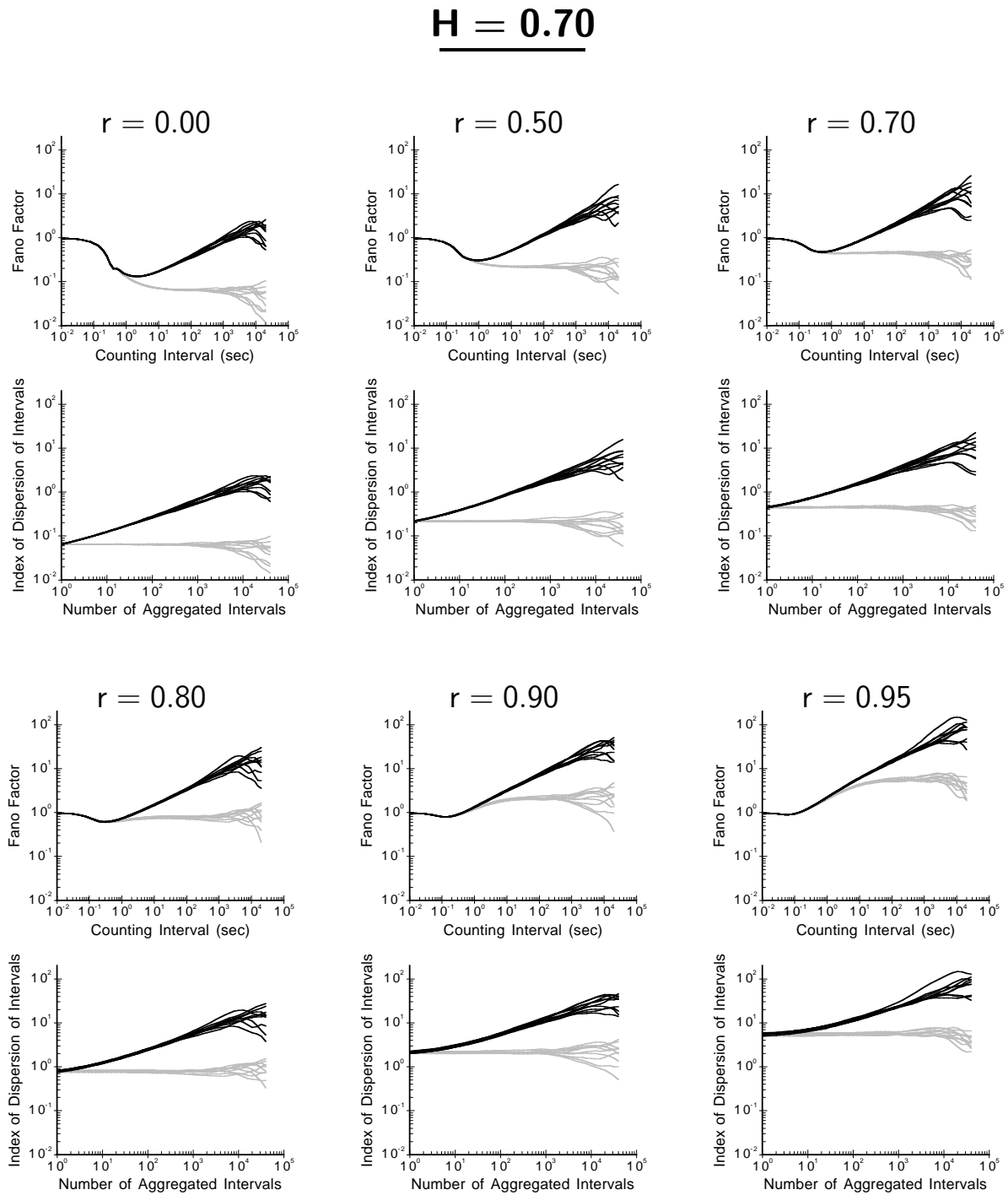
H = 0.90

r = 0.00

r = 0.50

r = 0.70
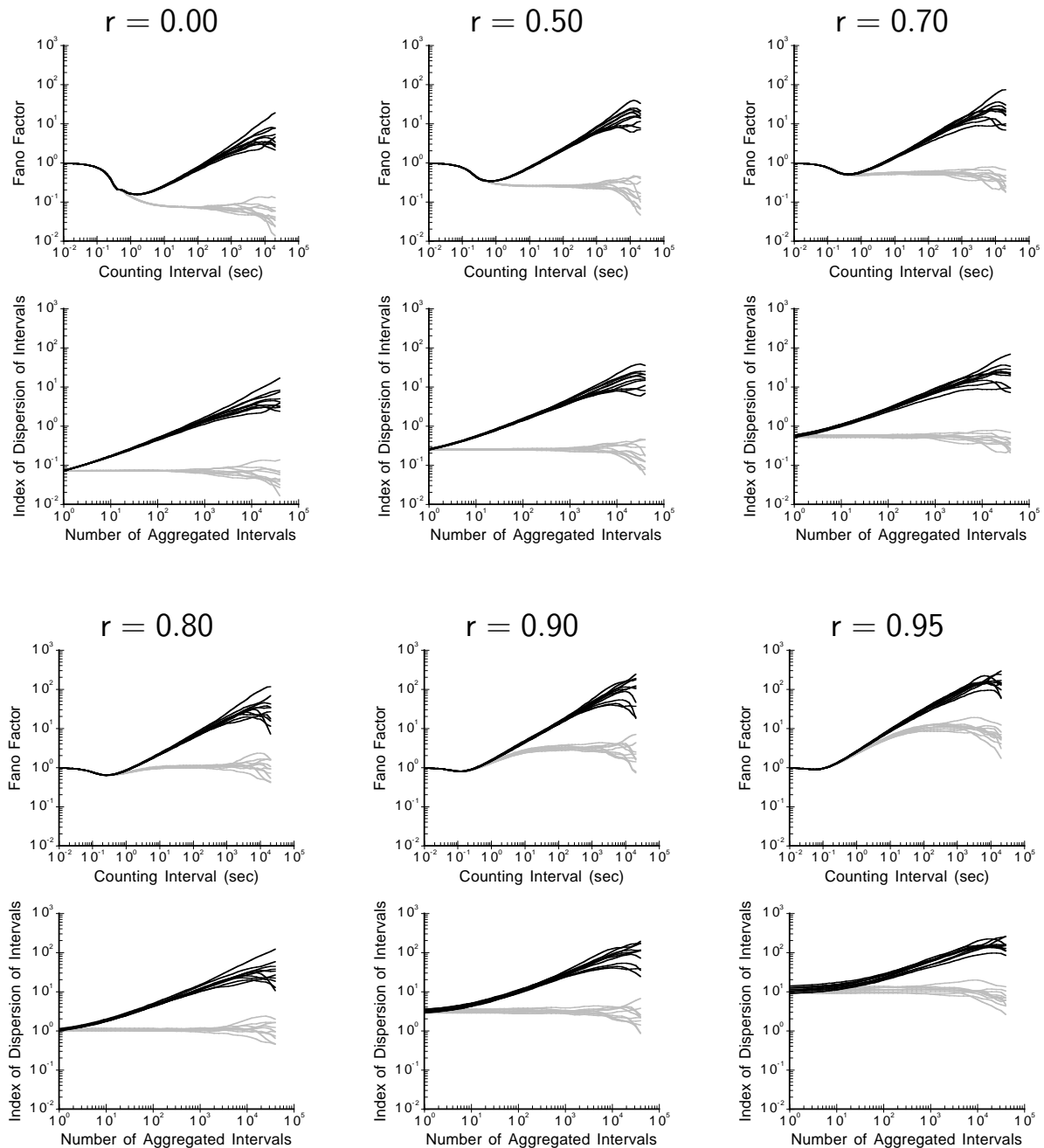

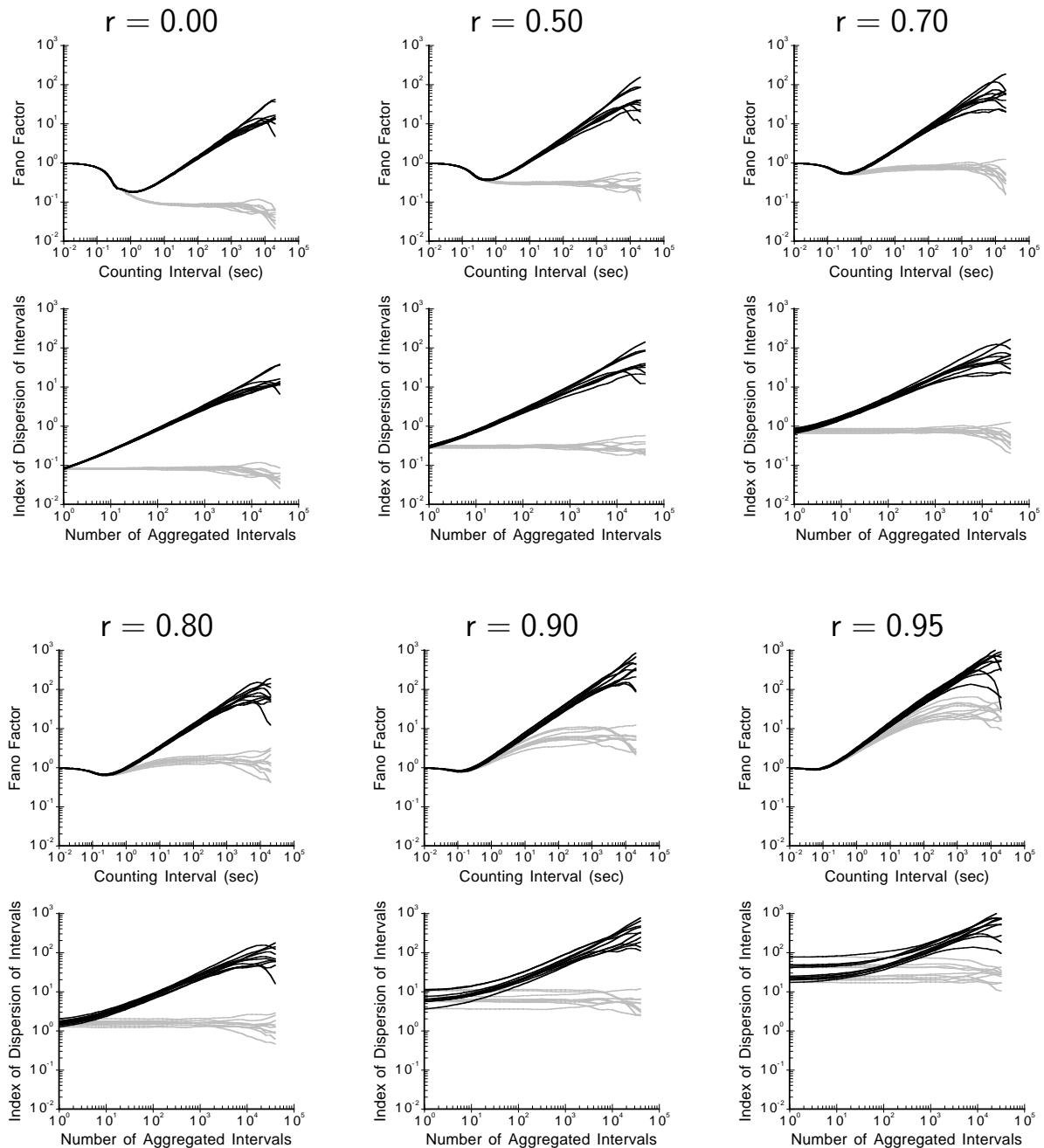
r = 0.80
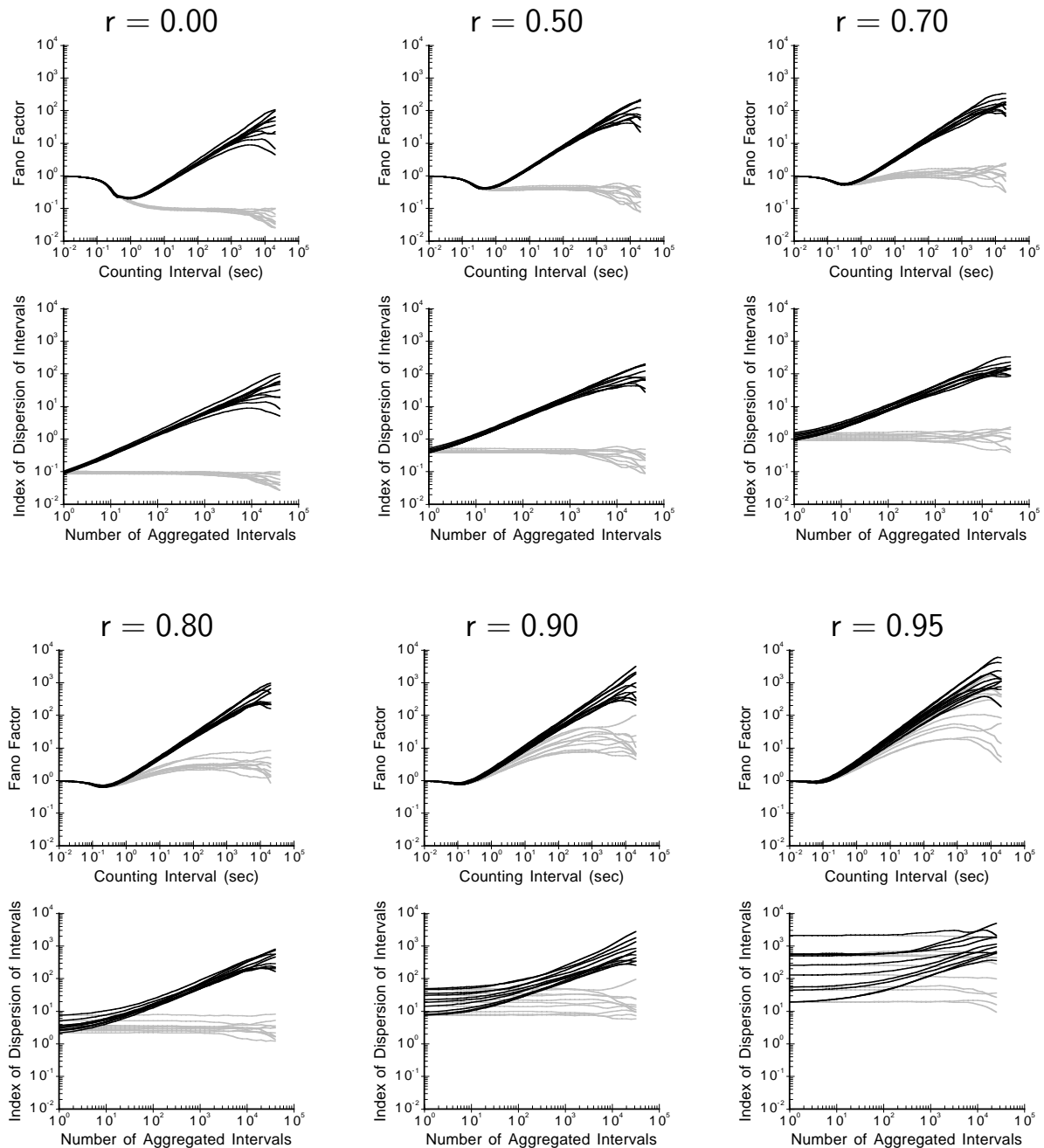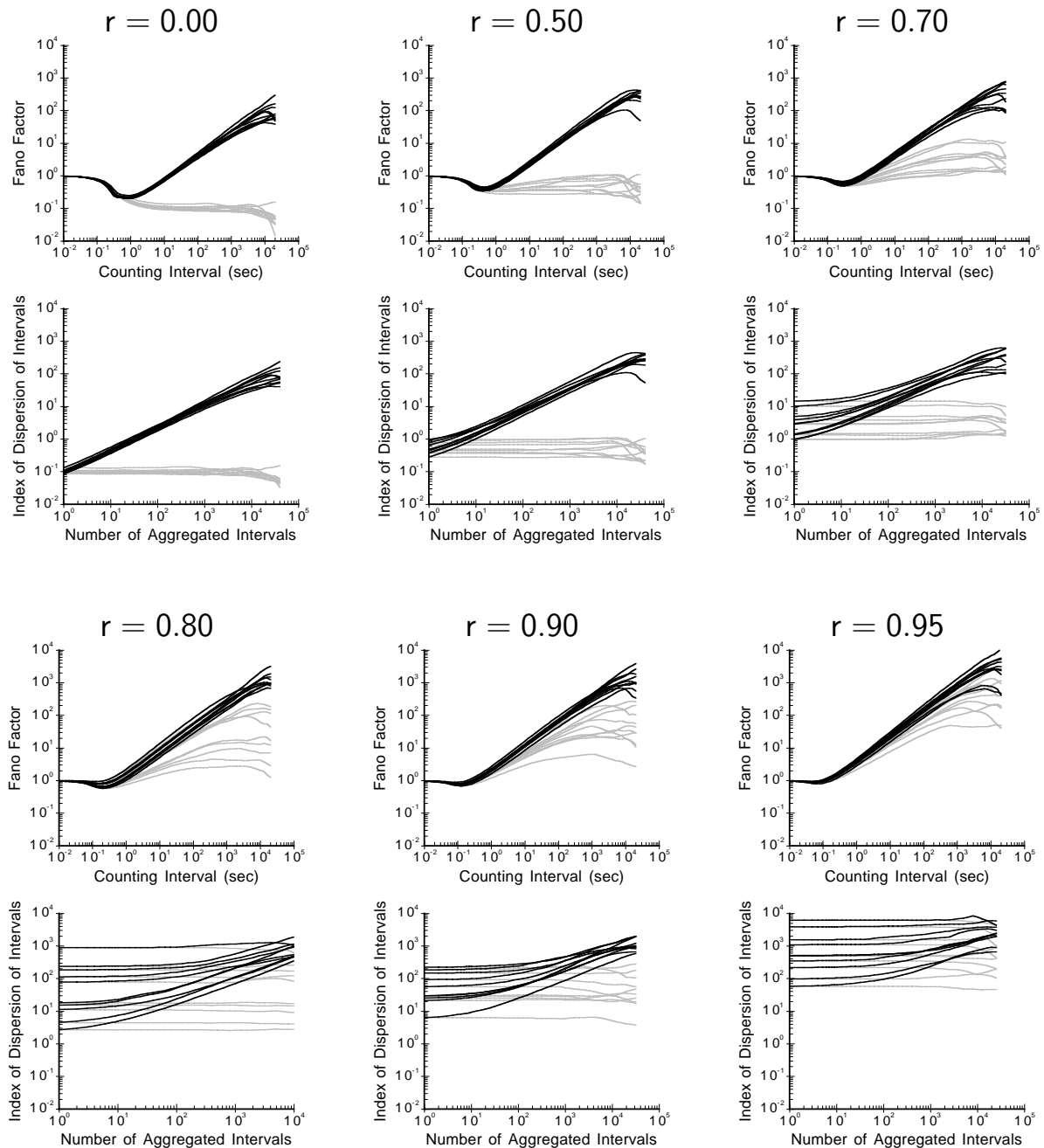

r = 0.90

r = 0.95

Figure B.19: FFCs and IDCs estimated from simulations of the IF model with fGnDP inputs and parameter $H = 0.90$. Same format as Figure B.11.
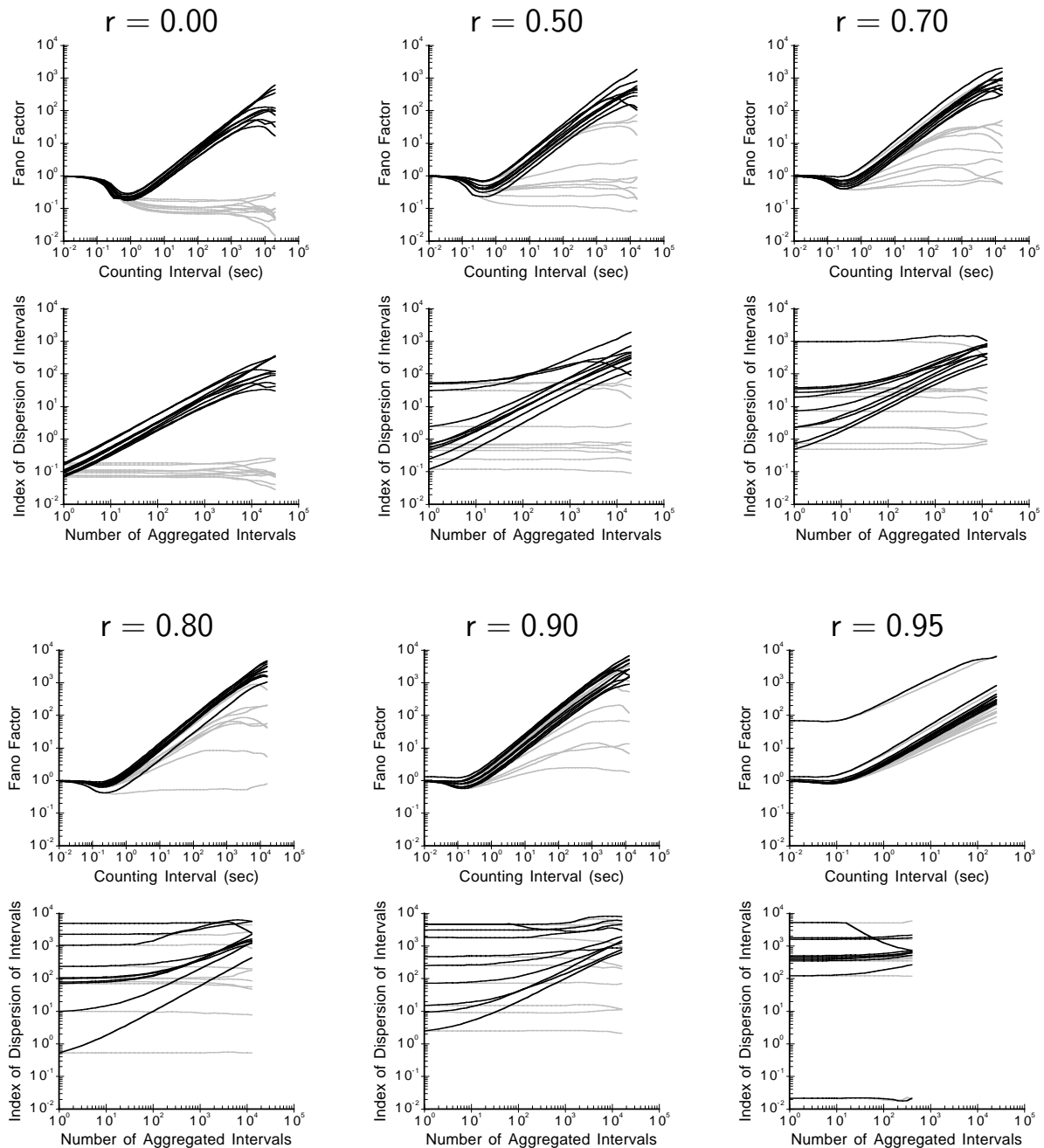
H = 0.95

Figure B.20: FFCs and IDCs estimated from simulations of the IF model with fGnDP inputs and parameter $H = 0.95$. Same format as Figure B.11.

Appendix C

Derivations of Moments and Moments of Statistics for fGnDP Counts

In Section 4.4.2, we listed a number of results for the fractional-Gaussian-noise-driven Poisson process without proof. In this appendix, we restate these results along with their proofs.

For the proof of the first theorem, we require the following two lemmas.

Lemma C.1. *For any integer $M \geq 1$,*

$$\sum_{n=1}^{M-1} \gamma(n) = \frac{1}{2}\Big(M^{2H} - (M-1)^{2H} - 1\Big), \tag{C.1}$$

where

$$\gamma(k) = \frac{1}{2}\left\{(k+1)^{2H} - 2k^{2H} + |k-1|^{2H}\right\}, \; \textit{for } k \in \mathbb{N}.$$

Proof. If $M = 1$, then both sides of (C.1) are equal to zero. If $M = 2$, then the left-hand side of (C.1) is

$$\sum_{n=1}^{1} \gamma(n) = \gamma(1) = \frac{1}{2}\left\{2^{2H} - 2 + 0\right\} = \frac{1}{2}\left\{2^{2H} - 1^{2H} - 1\right\},$$

which is equal to the right-hand side.

Now, assume that for some $k \geq 2$,

$$\sum_{n=1}^{k-1} \gamma_X(n) = \frac{1}{2}\Big(k^{2H} - (k-1)^{2H} - 1\Big).$$

Then,

$$\sum_{n=1}^{(k+1)-1} \gamma(n) = \sum_{n=1}^{k-1} \gamma(n) + \gamma(k)$$

$$= \frac{1}{2}\left(k^{2H} - (k-1)^{2H} - 1 \right) + \frac{1}{2}\left((k+1)^{2H} - 2k^{2H} + (k-1)^{2H} \right)$$

$$= \frac{1}{2}\left(k^{2H} - (k-1)^{2H} - 1 + (k+1)^{2H} - 2k^{2H} + (k-1)^{2H} \right)$$

$$= \frac{1}{2}\left((k+1)^{2H} - k^{2H} - 1 \right)$$

Thus, by induction, (C.1) is true for all integers $M \geq 1$. $\qquad\square$

Lemma C.2. *For any integer $M \geq 1$,*

$$\sum_{n=1}^{M-1} n\gamma(n) = \frac{M(M-1)}{2}\left(M^{2H-1} - (M-1)^{2H-1} \right), \qquad (C.2)$$

where

$$\gamma(k) = \frac{1}{2}\left\{ (k+1)^{2H} - 2k^{2H} + |k-1|^{2H} \right\}, \quad \textit{for } k \in \mathbb{N}.$$

Proof. If $M = 1$, then both sides of (C.2) are equal to zero. If $M = 2$, then the left-hand side of (C.2) is

$$\sum_{n=1}^{1} n\gamma(n) = \gamma(1) = \frac{1}{2}\left\{ 2^{2H} - 2 + 0 \right\} = 2^{2H-1} - 1 = \frac{2(1)}{2}\left(2^{2H-1} - 1^{2H-1} \right),$$

which is equal to the right-hand side.

Now, assume that for some $k \geq 2$,

$$\sum_{n=1}^{k-1} n\gamma(n) = \frac{k(k-1)}{2}\left(k^{2H-1} - (k-1)^{2H-1} \right).$$

Then,

$$\sum_{n=1}^{(k+1)-1} n\gamma(n) = \sum_{n=1}^{k-1} n\gamma(n) + k\gamma(k)$$

$$= \frac{k(k-1)}{2}\left(k^{2H-1} - (k-1)^{2H-1} \right) + \frac{k}{2}\left((k+1)^{2H} - 2k^{2H} + (k-1)^{2H} \right)$$

$$= \frac{(k+1)k}{2}\left(\frac{k-1}{k+1}k^{2H-1} - \frac{1}{k+1}(k-1)^{2H} + (k+1)^{2H-1} \right.$$

$$\left. - \frac{2}{k+1}k^{2H} + \frac{1}{k+1}(k-1)^{2H} \right)$$

$$= \frac{(k+1)k}{2}\left((k+1)^{2H-1} - k^{2H-1}\left[\frac{2k}{k+1} - \frac{k-1}{k+1} \right] \right)$$

$$= \frac{(k+1)k}{2}\left((k+1)^{2H-1} - k^{2H-1} \right)$$

Thus, by induction, (C.2) is true for all integers $M \geq 1$. □

Theorem C.3 (4.1). *Let $N(\cdot)$ be an fGnDP with rate process*

$$\Lambda(t) = \max\left\{0, \lambda + \sigma G_H\left(\left\lfloor \frac{t_0 + t}{\tau} \right\rfloor\right)\right\}, \tag{C.3}$$

where $G_H(k)$ is standard fractional Gaussian noise with Hurst index H. Assuming that the right term is negative with a probability of nearly zero, the mean count in an interval of length T is

$$E\{N(T)\} \approx \lambda T,$$

and the variance of the count in the same interval is given by the following:

(i) If $0 < T < \tau$, then

$$Var\{N(T)\} \approx \lambda T + \sigma^2 T^2 \left(1 + \frac{2^{2H-1} - 2}{3} \cdot \frac{T}{\tau}\right); \tag{C.4a}$$

(ii) If $T \geq \tau$ and $n = \left\lfloor \frac{T}{\tau} \right\rfloor$, then

$$Var\{N(T)\} \approx \lambda T + \frac{\sigma^2 \tau^2}{6}\left\{\left(\frac{T}{\tau} - n\right)^3 (n+2)^{2H}\right.$$

$$-\left[3\left(\frac{T}{\tau} - n\right)^2\left(\frac{T}{\tau} - n - 1\right) - \left(3\left(\frac{T}{\tau} - n\right) + 1\right)\right](n+1)^{2H}$$

$$+\left[3\left(\frac{T}{\tau} - n\right)^2\left(\frac{T}{\tau} - n - 2\right) + 4\right]n^{2H}$$

$$\left. -\left(\frac{T}{\tau} - n - 1\right)^3 (n-1)^{2H} - 2\right\}. \tag{C.4b}$$

Proof. Under the assumption that the probability of

$$\lambda + \sigma G_H\left(\left\lfloor \frac{t_0 + t}{\tau} \right\rfloor\right) < 0$$

is negligible, the mean and variance of the the rate process Λ are easily computed:

$$E\{\Lambda\} \approx E\{\lambda + \sigma G_H\} = \lambda + \sigma E\{G_H\} = \lambda$$

and

$$Var\{\Lambda\} \approx Var\{\lambda + \sigma G_H\}$$

$$= \sigma^2 Var\{G_H\} = \sigma^2\left(E\{G_H^2\} - \left(E\{G_H\}\right)^2\right) = \sigma^2\left(\gamma_G(0) - 0\right) = \sigma^2.$$

This allows computation of the mean of the fGnDP counting process:

$$E\{N(T)\} = E\{\Lambda\} \cdot T \approx \lambda T.$$

But for the variance of the counting process we need the autocovariance function of the rate process.

According to (4.4) the autocovariance function of the rate process Λ of the fGnDP is

$$\gamma_\Lambda(t) \approx \sigma^2 E_{t_0} \left\{ \gamma_G \left(\left\lfloor \frac{t_0 + t}{\tau} \right\rfloor - \left\lfloor \frac{t_0}{\tau} \right\rfloor \right) \right\}, \quad \text{for } t \geq 0,$$

where E_{t_0} is expectation with respect to the random variable t_0, which is uniformly distributed on $[0, \tau)$. Calculating this expectation simply involves integration over $[0, \tau)$ and division by τ. Thus, we get

$$\gamma_\Lambda(t) = \frac{\sigma^2}{\tau} \int_0^\tau \gamma_G \left(\left\lfloor \frac{t_0 + t}{\tau} \right\rfloor - \left\lfloor \frac{t_0}{\tau} \right\rfloor \right) dt_0.$$

Now, let $n = \left\lfloor \frac{t}{\tau} \right\rfloor$. Then, $n\tau \leq t < (n+1)\tau$, where $n \in \{0\} \cup \mathbb{N}$, and we have

$$\gamma_\Lambda(t) = \frac{\sigma^2}{\tau} \left\{ \int_0^{\tau - (t - n\tau)} \gamma_G \left(\left\lfloor \frac{t_0 + t}{\tau} \right\rfloor - \left\lfloor \frac{t_0}{\tau} \right\rfloor \right) dt_0 \right.$$

$$\left. + \int_{\tau - (t - n\tau)}^\tau \gamma_G \left(\left\lfloor \frac{t_0 + t}{\tau} \right\rfloor - \left\lfloor \frac{t_0}{\tau} \right\rfloor \right) dt_0 \right\}$$

$$= \frac{\sigma^2}{\tau} \left\{ \int_0^{(n+1)\tau - t} \gamma_G \left(\left\lfloor \frac{t_0 + t}{\tau} \right\rfloor - \left\lfloor \frac{t_0}{\tau} \right\rfloor \right) dt_0 \right.$$

$$\left. + \int_{(n+1)\tau - t}^\tau \gamma_G \left(\left\lfloor \frac{t_0 + t}{\tau} \right\rfloor - \left\lfloor \frac{t_0}{\tau} \right\rfloor \right) dt_0 \right\}$$

$$= \frac{\sigma^2}{\tau} \left\{ \int_0^{(n+1)\tau - t} \gamma_G \left(n - 0 \right) dt_0 + \int_{(n+1)\tau - t}^\tau \gamma_G \left((n+1) - 0 \right) dt_0 \right\}$$

$$= \frac{\sigma^2}{\tau} \left\{ \int_0^{(n+1)\tau - t} \gamma_G \left(n \right) dt_0 + \int_{(n+1)\tau - t}^\tau \gamma_G \left(n+1 \right) dt_0 \right\}$$

$$= \frac{\sigma^2}{\tau} \left\{ \gamma_G \left(n \right) \left[(n+1)\tau - t \right] + \gamma_G \left(n+1 \right) \left[\tau - \left((n+1)\tau - t \right) \right] \right\}$$

$$= \frac{\sigma^2}{\tau} \left\{ \gamma_G \left(n \right) \left[(n+1)\tau - t \right] + \gamma_G \left(n+1 \right) \left[t - n\tau \right] \right\}$$

$$= \sigma^2 \left\{ \gamma_G \left(n \right) \left[(n+1) - \frac{t}{\tau} \right] + \gamma_G \left(n+1 \right) \left[\frac{t}{\tau} - n \right] \right\}$$

Therefore, the variance of the counting process is

$$Var\{N(T)\} = E\{\Lambda\} \cdot T + 2\int_0^T (T-u)\,\gamma_\Lambda(u)\,du$$

$$\approx \lambda T + 2\int_0^T (T-u)\,\sigma^2 \left\{ \gamma_G\left(\left\lfloor\frac{u}{\tau}\right\rfloor\right)\left(\left\lfloor\frac{u}{\tau}\right\rfloor + 1 - \frac{u}{\tau}\right) \right.$$
$$\left. +\gamma_G\left(\left\lfloor\frac{u}{\tau}\right\rfloor + 1\right)\left(\frac{u}{\tau} - \left\lfloor\frac{u}{\tau}\right\rfloor\right) \right\} du$$

$$= \lambda T + 2\sigma^2 \int_0^T (T-u)\left\{ \gamma_G\left(\left\lfloor\frac{u}{\tau}\right\rfloor\right)\left(\left\lfloor\frac{u}{\tau}\right\rfloor + 1 - \frac{u}{\tau}\right) \right.$$
$$\left. +\gamma_G\left(\left\lfloor\frac{u}{\tau}\right\rfloor + 1\right)\left(\frac{u}{\tau} - \left\lfloor\frac{u}{\tau}\right\rfloor\right) \right\} du$$

For $0 < T < \tau$, the variance becomes

$$Var\{N(T)\} \approx \lambda T + 2\sigma^2 \int_0^T (T-u)\left\{ \gamma_G(0)\left(0 + 1 - \frac{u}{\tau}\right) + \gamma_G(0+1)\left(\frac{u}{\tau} - 0\right) \right\} du$$

$$= \lambda T + 2\sigma^2 \int_0^T (T-u)\left\{ \left(1 - \frac{u}{\tau}\right) + \left(2^{2H-1} - 1\right)\left(\frac{u}{\tau}\right) \right\} du$$

$$= \lambda T + 2\sigma^2 \int_0^T (T-u)\left\{ 1 + \left(2^{2H-1} - 2\right)\left(\frac{u}{\tau}\right) \right\} du$$

$$= \lambda T + 2\sigma^2 \left(\frac{T^2\left[3\tau + \left(2^{2H-1} - 2\right)T\right]}{6\tau} \right)$$

$$= \lambda T + \sigma^2 T^2 \left(1 + \frac{2^{2H-1} - 2}{3} \cdot \frac{T}{\tau} \right)$$

For $T \geq \tau$, the variance becomes

$$Var\{N(T)\} \approx \lambda T + 2\sigma^2 \left\{ \sum_{n=0}^{\left\lfloor\frac{T}{\tau}\right\rfloor - 1} \int_{n\tau}^{(n+1)\tau} (T-u)\left\{ \gamma_G(n)\left(n + 1 - \frac{u}{\tau}\right) \right. \right.$$

$$\left. +\gamma_G(n+1)\left(\frac{u}{\tau} - n\right) \right\} du$$

$$+ \int_{\left\lfloor\frac{T}{\tau}\right\rfloor\tau}^T (T-u)\left\{ \gamma_G\left(\left\lfloor\frac{T}{\tau}\right\rfloor\right)\left(\left\lfloor\frac{T}{\tau}\right\rfloor + 1 - \frac{u}{\tau}\right) \right.$$

$$\left. \left. +\gamma_G\left(\left\lfloor\frac{T}{\tau}\right\rfloor + 1\right)\left(\frac{u}{\tau} - \left\lfloor\frac{T}{\tau}\right\rfloor\right) \right\} du \right\}$$

Computing the first sum-integral combination, we get

$$\sum_{n=0}^{\lfloor \frac{T}{\tau} \rfloor - 1} \int_{n\tau}^{(n+1)\tau} (T-u) \left\{ \gamma_G(n) \left(n+1 - \frac{u}{\tau} \right) + \gamma_G(n+1) \left(\frac{u}{\tau} - n \right) \right\} du$$

$$= \sum_{n=0}^{\lfloor \frac{T}{\tau} \rfloor - 1} \left\{ \gamma_G(n) \int_{n\tau}^{(n+1)\tau} (T-u) \left(n+1 - \frac{u}{\tau} \right) du \right.$$

$$\left. + \gamma_G(n+1) \int_{n\tau}^{(n+1)\tau} (T-u) \left(\frac{u}{\tau} - n \right) du \right\}$$

$$= \sum_{n=0}^{\lfloor \frac{T}{\tau} \rfloor - 1} \left\{ -\frac{1}{6}\tau\, \gamma_G(n) \left(3n\tau + \tau - 3T \right) + -\frac{1}{6}\tau\, \gamma_G(n+1) \left(3n\tau + 2\tau - 3T \right) \right\}$$

$$= -\frac{\tau}{6} \sum_{n=0}^{\lfloor \frac{T}{\tau} \rfloor - 1} \left\{ \gamma_G(n) \left(3n\tau + \tau - 3T \right) + \gamma_G(n+1) \left(3n\tau + 2\tau - 3T \right) \right\}$$

$$= -\frac{\tau}{6} \left\{ \gamma_G(0) \left(\tau - 3T \right) \right.$$

$$+ \sum_{n=1}^{\lfloor \frac{T}{\tau} \rfloor - 1} \left\{ \gamma_G(n) \left(3n\tau - \tau - 3T \right) + \gamma_G(n) \left(3n\tau + \tau - 3T \right) \right\}$$

$$\left. + \gamma_G \left(\left\lfloor \frac{T}{\tau} \right\rfloor \right) \left(3\tau \left\lfloor \frac{T}{\tau} \right\rfloor - \tau - 3T \right) \right\} \qquad \text{if } T \geq \tau, \text{ else } 0$$

$$= -\frac{\tau}{6} \left\{ \gamma_G(0) \left(\tau - 3T \right) + 6 \sum_{n=1}^{\lfloor \frac{T}{\tau} \rfloor - 1} \left\{ \gamma_G(n) \left(n\tau - T \right) \right\} \right.$$

$$\left. + \gamma_G \left(\left\lfloor \frac{T}{\tau} \right\rfloor \right) \left(3\tau \left\lfloor \frac{T}{\tau} \right\rfloor - \tau - 3T \right) \right\}$$

$$= -\frac{\tau}{6} \left\{ \left(\tau - 3T \right) + 6 \sum_{n=1}^{\lfloor \frac{T}{\tau} \rfloor - 1} \left\{ \gamma_G(n) \left(n\tau - T \right) \right\} \right.$$

$$\left. + \gamma_G \left(\left\lfloor \frac{T}{\tau} \right\rfloor \right) \left(3\tau \left\lfloor \frac{T}{\tau} \right\rfloor - \tau - 3T \right) \right\}$$

Using Lemmas C.1 and C.2, and letting $n = \left\lfloor \frac{T}{\tau} \right\rfloor$, we have

$$\sum_{n=0}^{\left\lfloor \frac{T}{\tau} \right\rfloor - 1} \int_{n\tau}^{(n+1)\tau} (T - u) \left\{ \gamma_G(n) \left(n + 1 - \frac{u}{\tau} \right) + \gamma_G(n + 1) \left(\frac{u}{\tau} - n \right) \right\} du$$

$$= -\frac{\tau}{6} \left\{ (\tau - 3T) \right.$$

$$+ 6 \left(\frac{n(n-1)}{2} \left(n^{2H-1} - (n-1)^{2H-1} \right) \tau - \frac{1}{2} \left(n^{2H} - (n-1)^{2H} - 1 \right) T \right)$$

$$\left. + \frac{1}{2} \left((n+1)^{2H} - 2n^{2H} + (n-1)^{2H} \right) (3\tau n - \tau - 3T) \right\}$$

$$= -\frac{\tau}{12} \left\{ \left[(3n - 1)(n + 1)^{2H} - 4n^{2H} - (3n + 1)(n - 1)^{2H} + 2 \right] \tau \right.$$

$$\left. -3 \left[(n + 1)^{2H} - (n - 1)^{2H} \right] T \right\}$$

$$= -\frac{\tau}{12} \left\{ (3n\tau - \tau - 3T)(n + 1)^{2H} - 4\tau n^{2H} - (3n\tau + \tau - 3T)(n - 1)^{2H} + 2\tau \right\}$$

Letting $n = \left\lfloor \frac{T}{\tau} \right\rfloor$, and computing the second integral, we get

$$
\begin{aligned}
\int_{\left\lfloor \frac{T}{\tau} \right\rfloor \tau}^{T} (T-u) & \left\{ \gamma_G \left(\left\lfloor \frac{T}{\tau} \right\rfloor \right) \left(\left\lfloor \frac{T}{\tau} \right\rfloor + 1 - \frac{u}{\tau} \right) + \gamma_G \left(\left\lfloor \frac{T}{\tau} \right\rfloor + 1 \right) \left(\frac{u}{\tau} - \left\lfloor \frac{T}{\tau} \right\rfloor \right) \right\} du \\
&= \int_{n\tau}^{T} (T-u) \left\{ \gamma_G(n) \left(n + 1 - \frac{u}{\tau} \right) + \gamma_G(n+1) \left(\frac{u}{\tau} - n \right) \right\} du \\
&= \int_{n\tau}^{T} (T-u) \left\{ \left[(n+1)\gamma_G(n) - n\gamma_G(n+1) \right] + \left[\gamma_G(n+1) - \gamma_G(n) \right] \frac{u}{\tau} \right\} du \\
&= \frac{1}{2}(n\tau - T)^2 \left[(n+1)\gamma_G(n) - n\gamma_G(n+1) \right] \\
&\qquad\qquad + \frac{1}{6\tau}(n\tau - T)^2 (2n\tau + T) \left[\gamma_G(n+1) - \gamma_G(n) \right] \\
&= \frac{(n\tau - T)^2}{6\tau} \left\{ 3\tau \left[(n+1)\gamma_G(n) - n\gamma_G(n+1) \right] + (2n\tau + T) \left[\gamma_G(n+1) - \gamma_G(n) \right] \right\} \\
&= \frac{(n\tau - T)^2}{6\tau} \left\{ (3\tau + n\tau - T)\gamma_G(n) - (n\tau - T)\gamma_G(n+1) \right\} \\
&= \frac{(n\tau - T)^2}{12\tau} \left\{ (3\tau + n\tau - T) \left((n+1)^{2H} - 2n^{2H} + (n-1)^{2H} \right) \right. \\
&\qquad\qquad\qquad \left. - (n\tau - T) \left((n+2)^{2H} - 2(n+1)^{2H} + n^{2H} \right) \right\} \\
&= \frac{(n\tau - T)^2}{12\tau} \left\{ -(n\tau - T)(n+2)^{2H} + (3\tau + n\tau - T + 2n\tau - 2T)(n+1)^{2H} \right. \\
&\qquad\qquad\qquad \left. - (6\tau + 2n\tau - 2T + n\tau - T)n^{2H} + (3\tau + n\tau - T)(n-1)^{2H} \right\} \\
&= \frac{(n\tau - T)^2}{12\tau} \left\{ -(n\tau - T)(n+2)^{2H} + 3(\tau + n\tau - T)(n+1)^{2H} \right. \\
&\qquad\qquad\qquad \left. - 3(2\tau + n\tau - T)n^{2H} + (3\tau + n\tau - T)(n-1)^{2H} \right\}
\end{aligned}
$$

Thus, for $T \geq \tau$ and $n = \left\lfloor \frac{T}{\tau} \right\rfloor$, we have

$$
\begin{aligned}
Var\{N(T)\} \approx \lambda T + \frac{\sigma^2}{6\tau} & \left\{ (T - n\tau)^3 (n+2)^{2H} \right. \\
& + \left[3(T-n\tau)^2 \left((n+1)\tau - T \right) + \tau^2 \left(3(T-n\tau) + \tau \right) \right] (n+1)^{2H} \\
& \left. - \left[3(T-n\tau)^2 \left((n+2)\tau - T \right) - 4\tau^3 \right] n^{2H} + \left((n+1)\tau - T \right)^3 (n-1)^{2H} - 2\tau^3 \right\}
\end{aligned}
$$

$\square$

Theorem C.4 (4.3). *Let $N(\cdot)$ be the fGnDP in Theorem C.3. Using the approximation (4.6) to the autocovariance function of its rate process, the count variance of N can be*

approximated by

$$Var\left\{N(T)\right\} \approx \lambda T + \frac{\sigma^2\tau^2}{2(H+1)(2H+1)}$$

$$\times \left\{\left(\frac{T}{\tau}+1\right)^{2(H+1)} - 2\left(\frac{T}{\tau}\right)^{2(H+1)} + \left|\frac{T}{\tau}-1\right|^{2(H+1)} - 2\right\} \quad (C.5)$$

Proof. This approximation does not affect the mean and variance of the rate process, nor the mean of the counting process, since they do not depend on the autocovariance function. But the variance of the counting process is now

$$Var\left\{N(T)\right\} = E\left\{\Lambda\right\} \cdot T + 2\int_0^T (T-u)\,\gamma_\Lambda(u)\,du$$

$$\approx \lambda T + 2\int_0^T (T-u)\,\sigma^2\gamma_G(u)\,du$$

$$\approx \lambda T + 2\int_0^T (T-u)\,\sigma^2\left(\frac{1}{2\tau^{2H}}\right)\left\{(u+\tau)^{2H} - 2u^{2H} + |u-\tau|^{2H}\right\}\,du$$

$$= \lambda T + \frac{\sigma^2}{\tau^{2H}}\int_0^T (T-u)\left\{(u+\tau)^{2H} - 2u^{2H} + |u-\tau|^{2H}\right\}\,du.$$

For $0 < T \le \tau$, the variance is

$$Var\left[N(T)\right] = \lambda T + \frac{\sigma^2}{\tau^{2H}}\int_0^T (T-u)\left\{(u+\tau)^{2H} - 2u^{2H} + (\tau-u)^{2H}\right\}\,du$$

$$= \lambda T + \frac{\sigma^2}{\tau^{2H}}\left\{\int_0^T T\,(u+\tau)^{2H}\,du\right.$$

$$-\int_0^T u\,(u+\tau)^{2H}\,du - 2\int_0^T (T-u)\,u^{2H}\,du$$

$$\left. +\int_0^T T\,(\tau-u)^{2H}\,du - \int_0^T u\,(\tau-u)^{2H}\,du\right\}$$

Each of the integrals within the braces may be computed separately:

$$\bullet \int_0^T T\,(u+\tau)^{2H}\;du = \left.\frac{T(u+\tau)^{2H+1}}{2H+1}\right|_0^T = \frac{T\left[(T+\tau)^{2H+1}-\tau^{2H+1}\right]}{2H+1}$$

$$\bullet \int_0^T u\,(u+\tau)^{2H}\;du = \left.\frac{\left(2u^2H+2\tau Hu-\tau^2+u^2\right)(u+\tau)^{2H}}{2(H+1)(2H+1)}\right|_0^T$$
$$= \frac{\left(2T^2H+2\tau TH-\tau^2+T^2\right)(T+\tau)^{2H}+\tau^{2(H+1)}}{2(H+1)(2H+1)}$$

$$\bullet \int_0^T (T-u)\,u^{2H}\;du = \int_0^T Tu^{2H}-u^{2H+1}\;du$$
$$= \left.\left[\frac{T}{2H+1}u^{2H+1}-\frac{1}{2H+2}u^{2H+2}\right]\right|_0^T$$
$$= \frac{T^{2H+2}}{2H+1}-\frac{T^{2H+2}}{2H+2}-0$$
$$= \frac{2H+2-2H-1}{(2H+1)(2H+2)}T^{2H+2}$$
$$= \frac{T^{2(H+1)}}{2(H+1)(2H+1)}$$

$$\bullet \int_0^T T\,(\tau-u)^{2H}\;du = \left.\frac{-T(\tau-u)^{2H+1}}{2H+1}\right|_0^T = \frac{-T\left[(\tau-T)^{2H+1}-\tau^{2H+1}\right]}{2H+1}$$

$$\bullet \int_0^T u\,(\tau-u)^{2H}\;du = \left.\frac{\left(2u^2H-2\tau Hu-\tau^2+u^2\right)(\tau-u)^{2H}}{2(H+1)(2H+1)}\right|_0^T$$
$$= \frac{\left(2T^2H-2\tau TH-\tau^2+T^2\right)(\tau-T)^{2H}+\tau^{2(H+1)}}{2(H+1)(2H+1)}$$

Thus, the variance for $0 < T \le \tau$ is

$$Var\{N(T)\}$$
$$= \lambda T + \frac{\sigma^2}{\tau^{2H}}\left\{\frac{T}{2H+1}\left[(T+\tau)^{2H+1}-(\tau-T)^{2H+1}\right]-\frac{T^{2(H+1)}+\tau^{2(H+1)}}{(H+1)(2H+1)}\right.$$
$$\left.-\frac{\left(2T^2H-2\tau TH-\tau^2+T^2\right)\left[(T+\tau)^{2H}+(\tau-T)^{2H}\right]}{2(H+1)(2H+1)}\right\}$$
$$= \lambda T + \frac{\sigma^2}{\tau^{2H}}\cdot\frac{(T+\tau)^{2(H+1)}-2\left(T^{2(H+1)}+\tau^{2(H+1)}\right)+(\tau-T)^{2(H+1)}}{2(H+1)(2H+1)}$$
$$= \lambda T + \frac{\sigma^2\tau^2\left[\left(\frac{T}{\tau}+1\right)^{2(H+1)}-2\left(\left(\frac{T}{\tau}\right)^{2(H+1)}+1\right)+\left(1-\frac{T}{\tau}\right)^{2(H+1)}\right]}{2(H+1)(2H+1)}$$

For $T > \tau$, the variance is

$$
Var\left[N(T)\right] = \lambda T + \frac{\sigma^2}{\tau^{2H}} \left\{ \int_0^T T\left(u+\tau\right)^{2H} du \right.
$$

$$
- \int_0^T u\left(u+\tau\right)^{2H} du - 2\int_0^T \left(T-u\right)u^{2H} du
$$

$$
+ \int_0^\tau T\left(\tau-u\right)^{2H} du + \int_\tau^T T\left(u-\tau\right)^{2H} du
$$

$$
\left. - \int_0^\tau u\left(\tau-u\right)^{2H} du - \int_\tau^T u\left(u-\tau\right)^{2H} du \right\}
$$

The first three integrals within the braces have already been calculated above. The final four integrals are:

- $\displaystyle \int_0^\tau T\left(\tau-u\right)^{2H} du = \left. \frac{-T(\tau-u)^{2H+1}}{2H+1}\right|_0^\tau = \frac{T\tau^{2H+1}}{2H+1}$

- $\displaystyle \int_\tau^T T\left(u-\tau\right)^{2H} du = \left. \frac{T(u-\tau)^{2H+1}}{2H+1}\right|_\tau^T = \frac{T(T-\tau)^{2H+1}}{2H+1}$

- $\displaystyle \int_0^\tau u\left(\tau-u\right)^{2H} du = \left. \frac{\left(2u^2H - 2\tau Hu - \tau^2 + u^2\right)(\tau-u)^{2H}}{2(H+1)(2H+1)}\right|_0^\tau = \frac{\tau^{2(H+1)}}{2(H+1)(2H+1)}$

- $\displaystyle \int_\tau^T u\left(u-\tau\right)^{2H} du = \left. \frac{\left(2u^2H - 2\tau Hu - \tau^2 + u^2\right)(u-\tau)^{2H}}{2(H+1)(2H+1)}\right|_\tau^T$

$$
= \frac{\left[2T^2H - 2T\tau H - \tau^2 + T^2\right](T-\tau)^{2H}}{2(H+1)(2H+1)}
$$

$$
= \frac{\left[2TH(T-\tau) + (T-\tau)(T+\tau)\right](T-\tau)^{2H}}{2(H+1)(2H+1)}
$$

$$
= \frac{\left[(2H+1)T + \tau\right](T-\tau)^{2H+1}}{2(H+1)(2H+1)}
$$

Thus, the variance for $T > \tau$ is

$$Var\{N(T)\}$$

$$= \lambda T + \frac{\sigma^2}{\tau^{2H}}\left\{\frac{T}{2H+1}\left[(T+\tau)^{2H+1} + (T-\tau)^{2H+1}\right] - \frac{T^{2(H+1)} + \tau^{2(H+1)}}{(H+1)(2H+1)}\right.$$

$$\left. - \frac{\left[(2H+1)T - \tau\right](T+\tau)^{2H+1} + \left[(2H+1)T + \tau\right](T-\tau)^{2H+1}}{2(H+1)(2H+1)}\right\}$$

$$= \lambda T + \frac{\sigma^2}{\tau^{2H}} \cdot \frac{(T+\tau)^{2(H+1)} - 2\left(T^{2(H+1)} + \tau^{2(H+1)}\right) + (T-\tau)^{2(H+1)}}{2(H+1)(2H+1)}$$

$$= \lambda T + \frac{\sigma^2\tau^2\left[\left(\frac{T}{\tau}+1\right)^{2(H+1)} - 2\left(\left(\frac{T}{\tau}\right)^{2(H+1)} + 1\right) + \left(\frac{T}{\tau}-1\right)^{2(H+1)}\right]}{2(H+1)(2H+1)}$$

This differs from the expression for $0 < T \leq \tau$ only in the last term in the numerator, which is $\left(\frac{T}{\tau}-1\right)^{2(H+1)}$ for $T > \tau$ and $\left(1-\frac{T}{\tau}\right)^{2(H+1)}$ for $T \leq \tau$. Thus, for any $T > 0$, the variance is

$$Var\{N(T)\} = \lambda T + \frac{\sigma^2\tau^2\left[\left(\frac{T}{\tau}+1\right)^{2(H+1)} - 2\left(\left(\frac{T}{\tau}\right)^{2(H+1)} + 1\right) + \left|\frac{T}{\tau}-1\right|^{2(H+1)}\right]}{2(H+1)(2H+1)}$$

$\square$

Theorem C.5 (4.4). *Let $N(\cdot)$ be an fGnDP as in Theorem C.3, and define the sample mean of the counts as in Definition 4.4. Then the expected value of the sample mean, assuming only that truncation of the rate process is negligible, is*

$$E\{m_r(T)\} = E\{N(T)\} \approx \lambda T.$$

Thus, the sample mean is an unbiased estimator of the mean count.

Proof.

$$E\{m_r(T)\} = E\left\{\frac{1}{r}\sum_{i=1}^{r} N_i(T)\right\} = \frac{1}{r}\sum_{i=1}^{r} E\{N_i(T)\} = \frac{1}{r}\sum_{i=1}^{r} E\{N(T)\} = E\{N(T)\}.$$

$\square$

Theorem C.6 (4.5). *Let $N(\cdot)$ be an fGnDP as in Theorem C.3, and define the sample mean of the counts as in Definition 4.4. Assume that $S \geq \tau$. Then, using the approximation (4.6) to the autocovariance function of the rate process of N, the variance of the*

sample mean count is approximately

$$Var\{m_r(T)\} \approx \frac{\lambda T}{r} + \frac{\sigma^2}{2r\tau^{2H}(2H+1)(H+1)}$$

$$\times \left\{ \left[(T+\tau)^{2(H+1)} - 2\left(T^{2(H+1)} + \tau^{2(H+1)}\right) + |T-\tau|^{2(H+1)} \right] \right.$$

$$+ \frac{1}{r}\sum_{k=1}^{r-1}(r-k)\left[(kS+T+\tau)^{2(H+1)} - 2(kS+T)^{2(H+1)} + |kS+T-\tau|^{2(H+1)} \right.$$

$$- 2\left[(kS+\tau)^{2(H+1)} - 2(kS)^{2(H+1)} + |kS-\tau|^{2(H+1)} \right]$$

$$\left. \left. + (kS-T+\tau)^{2(H+1)} - 2(kS-T)^{2(H+1)} + |kS-T-\tau|^{2(H+1)} \right] \right\}.$$

Proof. The variance of $m_r(T)$ is

$$Var\{m_r(T)\} = \frac{1}{r^2}Var\left\{\sum_{i=1}^{r}N_i(T)\right\}$$

$$= \frac{1}{r^2}\left(\sum_{i=1}^{r}Var\{N_i(T)\} + 2\sum_{i<j}Cov\{N_i(T), N_j(T)\}\right)$$

$$= \frac{1}{r}Var\{N(T)\} + \frac{2}{r^2}\sum_{i<j}Cov\{N_i(T), N_j(T)\}$$

$$= \frac{1}{r}Var\{N(T)\} + \frac{2}{r^2}\sum_{k=1}^{r-1}\sum_{i=1}^{k}Cov\{N_i(T), N_{i+r-k}(T)\}$$

$$= \frac{1}{r}Var\{N(T)\} + \frac{2}{r^2}\sum_{k=1}^{r-1}\sum_{i=1}^{k}Cov\{N_1(T), N_{1+r-k}(T)\}$$

$$= \frac{1}{r}Var\{N(T)\} + \frac{2}{r^2}\sum_{k=1}^{r-1}k\,Cov\{N_1(T), N_{1+r-k}(T)\}$$

An approximate expression for $Var\{N(T)\}$ is given in (C.5). Since the sample counting intervals do not overlap (recall that $T \leq S$), the covariance of the counts in any two distinct intervals is (Cox & Isham, 1980, p. 34)

$$Cov\{N_1(T), N_j(T)\} = E\{\Lambda\}\int_0^T \int_{(j-1)S}^{(j-1)S+T} h(u-z)\,du\,dz - \left(E\{\Lambda\}\right)^2 T^2, \quad j > 1,$$

where h is the conditional intensity function. The conditional intensity function for a doubly stochastic Poisson process is (Cox & Isham, 1980, p. 72)

$$h(t) = E\{\Lambda\} + \gamma_\Lambda(t)/E\{\Lambda\},$$

where γ_Λ is the covariance function of Λ.

Thus, we have that

$$
Cov\Big\{N_1(T), N_j(T)\Big\}
$$

$$
= E\{\Lambda\} \int_0^T \int_{(j-1)S}^{(j-1)S+T} E\{\Lambda\} \, du \, dz
$$

$$
+ E\{\Lambda\} \int_0^T \int_{(j-1)S}^{(j-1)S+T} \Big(\gamma_\Lambda(u-z)/E\{\Lambda\}\Big) \, du \, dz - \Big(E\{\Lambda\}\Big)^2 T^2
$$

$$
= \int_0^T \int_{(j-1)S}^{(j-1)S+T} \gamma_\Lambda(u-z) \, du \, dz
$$

$$
= \frac{\sigma^2}{2\tau^{2H}} \int_0^T \int_{(j-1)S}^{(j-1)S+T} \Big\{(u-z+\tau)^{2H} - 2(u-z)^{2H} + |u-z-\tau|^{2H}\Big\} \, du \, dz
$$

$$
= \frac{\sigma^2}{2\tau^{2H}} \Bigg\{\int_0^T \int_{(j-1)S}^{(j-1)S+T} \Big[(u-z+\tau)^{2H} - 2(u-z)^{2H}\Big] \, du \, dz
$$

$$
+ \int_0^T \int_{(j-1)S}^{(j-1)S+T} |u-z-\tau|^{2H} \, du \, dz\Bigg\}
$$

Computing the first double integral, we get

$$\int_0^T \int_{(j-1)S}^{(j-1)S+T} \left[(u-z+\tau)^{2H} - 2(u-z)^{2H} \right] du\, dz$$

$$= \int_0^T \int_{(j-1)S}^{(j-1)S+T} \left\{ (u-z+\tau)^{2H} - 2(u-z)^{2H} \right\} du\, dz$$

$$= \frac{1}{2H+1} \int_0^T \left\{ (u-z+\tau)^{2H+1} - 2(u-z)^{2H+1} \right\} \Big|_{u=(j-1)S}^{u=(j-1)S+T} dz$$

$$= \frac{-1}{(2H+1)(2H+2)} \left\{ (u-z+\tau)^{2H+2} - 2(u-z)^{2H+2} \right\} \Big|_{u=(j-1)S}^{u=(j-1)S+T} \Big|_{z=0}^{z=T}$$

$$= \frac{-1}{2(2H+1)(H+1)} \Big\{ (u-T+\tau)^{2(H+1)} - 2(u-T)^{2(H+1)}$$

$$\qquad\qquad\qquad\qquad - (u+\tau)^{2(H+1)} + 2u^{2(H+1)} \Big\} \Big|_{u=(j-1)S}^{u=(j-1)S+T}$$

$$= \frac{-1}{2(2H+1)(H+1)} \Big\{ 2 \Big[\big((j-1)S+\tau \big)^{2(H+1)} - 2 \big((j-1)S \big)^{2(H+1)} \Big]$$

$$\qquad\qquad - \big((j-1)S+T+\tau \big)^{2(H+1)} + 2 \big((j-1)S+T \big)^{2(H+1)}$$

$$\qquad\qquad - \big((j-1)S-T+\tau \big)^{2(H+1)} + 2 \big((j-1)S-T \big)^{2(H+1)} \Big\}$$

$$= \frac{1}{2(2H+1)(H+1)}$$

$$\qquad \times \Big\{ \big((j-1)S+T+\tau \big)^{2(H+1)} - 2 \big((j-1)S+T \big)^{2(H+1)}$$

$$\qquad + \big((j-1)S-T+\tau \big)^{2(H+1)} - 2 \big((j-1)S-T \big)^{2(H+1)}$$

$$\qquad - 2 \Big[\big((j-1)S+\tau \big)^{2(H+1)} - 2 \big((j-1)S \big)^{2(H+1)} \Big] \Big\}$$

In order to compute the second double integral, we note that

$$\int \int |u-z-\tau|^{2H} du\, dz = \frac{-1}{(2H+1)(2H+2)} |u-z-\tau|^{2H+2}.$$

This is due to symmetry between the two integration variables: their coefficients are plus

one and minus one. Thus, the second double integral is

$$\int_0^T \int_{(j-1)S}^{(j-1)S+T} |u - z - \tau|^{2H}\, du\, dz$$

$$= \frac{-1}{(2H+1)(2H+2)} \left. |u - z - \tau|^{2H+2} \right|_{u=(j-1)S}^{u=(j-1)S+T} \Big|_{z=0}^{z=T}$$

$$= \frac{-1}{2(2H+1)(H+1)} \left\{ |u-T-\tau|^{2(H+1)} - |u-\tau|^{2(H+1)} \right\} \Big|_{u=(j-1)S}^{u=(j-1)S+T}$$

$$= \frac{-1}{2(2H+1)(H+1)} \left\{ 2\big|(j-1)S - \tau\big|^{2(H+1)} \right.$$

$$\left. - \big|(j-1)S + T - \tau\big|^{2(H+1)} - \big|(j-1)S - T - \tau\big|^{2(H+1)} \right\}$$

$$= \frac{1}{2(2H+1)(H+1)} \left\{ \big|(j-1)S + T - \tau\big|^{2(H+1)} \right.$$

$$\left. + \big|(j-1)S - T - \tau\big|^{2(H+1)} - 2\big|(j-1)S - \tau\big|^{2(H+1)} \right\}$$

Hence, the covariance is

$$
Cov\Big\{N_1(T), N_j(T)\Big\}
$$

$$
= \frac{a^2}{2\tau^{2H}} \left\{ \int_0^T \int_{(j-1)S}^{(j-1)S+T} \left[(u-z+\tau)^{2H} - 2(u-z)^{2H} \right] du\, dz \right.
$$

$$
\left. + \int_0^T \int_{(j-1)S}^{(j-1)S+T} |u-z-\tau|^{2H}\, du\, dz \right\}
$$

$$
= \frac{a^2}{4\tau^{2H}(2H+1)(H+1)}
$$

$$
\times \left\{ \Big((j-1)S+T+\tau\Big)^{2(H+1)} - 2\Big((j-1)S+T\Big)^{2(H+1)} \right.
$$

$$
+ \Big((j-1)S-T+\tau\Big)^{2(H+1)} - 2\Big((j-1)S-T\Big)^{2(H+1)}
$$

$$
- 2\left[\Big((j-1)S+\tau\Big)^{2(H+1)} - 2\Big((j-1)S\Big)^{2(H+1)} \right]
$$

$$
+ \Big|(j-1)S+T-\tau\Big|^{2(H+1)} + \Big|(j-1)S-T-\tau\Big|^{2(H+1)}
$$

$$
\left. - 2\Big|(j-1)S-\tau\Big|^{2(H+1)} \right\}
$$

$$
= \frac{a^2}{4\tau^{2H}(2H+1)(H+1)}
$$

$$
\times \left\{ \Big((j-1)S+T+\tau\Big)^{2(H+1)} - 2\Big((j-1)S+T\Big)^{2(H+1)} \right.
$$

$$
+ \Big|(j-1)S+T-\tau\Big|^{2(H+1)} - 2\left[\Big((j-1)S+\tau\Big)^{2(H+1)} \right.
$$

$$
\left. -2\Big((j-1)S\Big)^{2(H+1)} + \Big|(j-1)S-\tau\Big|^{2(H+1)} \right]
$$

$$
+ \Big((j-1)S-T+\tau\Big)^{2(H+1)} - 2\Big((j-1)S-T\Big)^{2(H+1)}
$$

$$
\left. + \Big|(j-1)S-T-\tau\Big|^{2(H+1)} \right\}
$$

Therefore, the variance of $m_r(T)$ is

$$
Var\{m_r(T)\} = \frac{\lambda T}{r} + \frac{\sigma^2}{2r\tau^{2H}(2H+1)(H+1)}
$$

$$
\times \left\{ \left[(T+\tau)^{2(H+1)} - 2\left(T^{2(H+1)} + \tau^{2(H+1)}\right) + |T-\tau|^{2(H+1)} \right] \right.
$$

$$
+ \frac{1}{r}\sum_{k=1}^{r-1} k \left[\left((r-k)S+T+\tau\right)^{2(H+1)} - 2\left((r-k)S+T\right)^{2(H+1)} \right.
$$

$$
+ \left|(r-k)S+T-\tau\right|^{2(H+1)} - 2\left\{ \left((r-k)S+\tau\right)^{2(H+1)} \right.
$$

$$
-2\left((r-k)S\right)^{2(H+1)} + \left|(r-k)S-\tau\right|^{2(H+1)}\bigg\}
$$

$$
+ \left((r-k)S-T+\tau\right)^{2(H+1)} - 2\left((r-k)S-T\right)^{2(H+1)}
$$

$$
\left. \left. + \left|(r-k)S-T-\tau\right|^{2(H+1)} \right]\right\}
$$

or

$$
Var\{m_r(T)\} = \frac{\lambda T}{r} + \frac{\sigma^2}{2r\tau^{2H}(2H+1)(H+1)}
$$

$$
\times \left\{ \left[(T+\tau)^{2(H+1)} - 2\left(T^{2(H+1)} + \tau^{2(H+1)}\right) + |T-\tau|^{2(H+1)} \right] \right.
$$

$$
+ \frac{1}{r}\sum_{k=1}^{r-1}(r-k)\left[(kS+T+\tau)^{2(H+1)} \right.
$$

$$
- 2(kS+T)^{2(H+1)} + |kS+T-\tau|^{2(H+1)}
$$

$$
- 2\left[(kS+\tau)^{2(H+1)} - 2(kS)^{2(H+1)} + |kS-\tau|^{2(H+1)}\right]
$$

$$
+ (kS-T+\tau)^{2(H+1)} - 2(kS-T)^{2(H+1)}
$$

$$
\left. \left. + |kS-T-\tau|^{2(H+1)} \right]\right\}.
$$

$\square$

Theorem C.7 (4.6). *Let $N(\cdot)$ be an fGnDP as in Theorem C.3, and define the sample mean and sample variance of the counts as in Definition 4.4. Assume that $S \geq \tau$. Then, using the approximation (4.6) to the autocovariance function of the rate process of N, the*

expected value of the sample variance is approximately

$$E\left\{s_r^2(T)\right\} \approx Var\{N(T)\} - \left(\frac{1}{r(r-1)}\right) \cdot \left(\frac{\sigma^2}{2\tau^{2H}(2H+1)(H+1)}\right)$$

$$\times \sum_{k=1}^{r-1}(r-k)\left\{(kS+T+\tau)^{2(H+1)} - 2(kS+T)^{2(H+1)} + |kS+T-\tau|^{2(H+1)}\right.$$

$$- 2\left[(kS+\tau)^{2(H+1)} - 2(kS)^{2(H+1)} + |kS-\tau|^{2(H+1)}\right]$$

$$\left. + (kS-T+\tau)^{2(H+1)} - 2(kS-T)^{2(H+1)} + |kS-T-\tau|^{2(H+1)}\right\}$$

Proof. The expected value of the sample variance is

$$E\left\{\frac{1}{r-1}\sum_{i=1}^{r}\Big(N_i(T)-m_r(T)\Big)^2\right\}$$

$$=\frac{1}{r-1}\sum_{i=1}^{r}E\left\{\Big(N_i(T)-m_r(T)\Big)^2\right\}$$

$$=\frac{1}{r-1}\sum_{i=1}^{r}E\left\{N_i^2(T)-2N_i(T)\,m_r(T)+m_r^2(T)\right\}$$

$$=\frac{1}{r-1}\sum_{i=1}^{r}\left(E\left\{N_i^2(T)\right\}-2E\left\{N_i(T)\,m_r(T)\right\}+E\left\{m_r^2(T)\right\}\right)$$

$$=\frac{1}{r-1}\sum_{i=1}^{r}\left(E\left\{N^2(T)\right\}-2E\left\{N_i(T)\,m_r(T)\right\}+E\left\{m_r^2(T)\right\}\right)$$

$$=\frac{1}{r-1}\left(nE\left\{N^2(T)\right\}-2E\left\{\Big(\sum_{i=1}^{r}N_i(T)\Big)m_r(T)\right\}+nE\left\{m_r^2(T)\right\}\right)$$

$$=\frac{1}{r-1}\left(nE\left\{N^2(T)\right\}-2nE\left\{m_r^2(T)\right\}+nE\left\{m_r^2(T)\right\}\right)$$

$$=\frac{r}{r-1}\left(E\left\{N^2(T)\right\}-E\left\{m_r^2(T)\right\}\right)$$

$$=\frac{r}{r-1}\left(Var\{N(T)\}+\Big(E\{N(T)\}\Big)^2-E\left\{m_r^2(T)\right\}\right),$$

$$\text{since } Var\{N(T)\}=E\left\{N^2(T)\right\}-\Big(E\{N(T)\}\Big)^2$$

$$=\frac{r}{r-1}\left(Var\{N(T)\}+\Big(E\{N(T)\}\Big)^2-Var\left\{m_r(T)\right\}-\Big(E\left\{m_r(T)\right\}\Big)^2\right),$$

$$\text{since } Var\left\{m_r(T)\right\}=E\left\{m_r^2(T)\right\}-\Big(E\left\{m_r(T)\right\}\Big)^2$$

$$=\frac{r}{r-1}\left(Var\{N(T)\}+\Big(E\{N(T)\}\Big)^2-Var\left\{m_r(T)\right\}-\Big(E\{N(T)\}\Big)^2\right)$$

$$=\frac{r}{r-1}\left(Var\{N(T)\}-Var\left\{m_r(T)\right\}\right)$$

$$=\frac{r}{r-1}\left(Var\{N(T)\}-\frac{1}{r}Var\{N(T)\}-\frac{2}{r^2}\sum_{1\le i<j\le r}Cov\Big\{N_i(T),N_j(T)\Big\}\right)$$

$$=Var\{N(T)\}-\frac{2}{r(r-1)}\sum_{i<j}Cov\Big\{N_i(T),N_j(T)\Big\}$$

From the proof of Theorem C.6, we know that

$$
2\sum_{i<j} Cov\Big\{N_i(T), N_j(T)\Big\} \approx \frac{\sigma^2}{2\tau^{2H}(2H+1)(H+1)}
$$
$$
\times \sum_{k=1}^{r-1}(r-k)\Big\{(kS+T+\tau)^{2(H+1)} - 2(kS+T)^{2(H+1)} + |kS+T-\tau|^{2(H+1)}
$$
$$
- 2\Big[(kS+\tau)^{2(H+1)} - 2(kS)^{2(H+1)} + |kS-\tau|^{2(H+1)}\Big]
$$
$$
+ (kS-T+\tau)^{2(H+1)} - 2(kS-T)^{2(H+1)} + |kS-T-\tau|^{2(H+1)}\Big\}
$$

Thus, the expected value of the sample variance is approximately

$$
E\Big\{\frac{1}{r-1}\sum_{i=1}^{r}(N_i(T) - m_r(T))^2\Big\} \approx Var\{N(T)\}
$$
$$
- \Big(\frac{1}{r(r-1)}\Big) \cdot \Big(\frac{\sigma^2}{2\tau^{2H}(2H+1)(H+1)}\Big)
$$
$$
\times \sum_{k=1}^{r-1}(r-k)\Big\{(kS+T+\tau)^{2(H+1)} - 2(kS+T)^{2(H+1)} + |kS+T-\tau|^{2(H+1)}
$$
$$
- 2\Big[(kS+\tau)^{2(H+1)} - 2(kS)^{2(H+1)} + |kS-\tau|^{2(H+1)}\Big]
$$
$$
+ (kS-T+\tau)^{2(H+1)} - 2(kS-T)^{2(H+1)} + |kS-T-\tau|^{2(H+1)}\Big\}
$$

$\square$

Bibliography

Abry, P., Flandrin, P., Taqqu, M. S., & Veitch, D. (2000). Wavelets for the analysis, estimation and synthesis of scaling data. In K. Park & W. Willinger (Eds.), *Self-similar network traffic and performance evaluation* (pp. 39–88). New York: Wiley.

Bardet, J.-M., Lang, G., Oppenheim, G., Philippe, A., & Taqqu, M. S. (2002). Generators of long-range dependent processes : A survey. In *Long-range dependence : Theory and applications.* Boston: Birkhauser.

Bell, A. J., Mainen, Z. F., Tsodyks, M., & Sejnowski, T. J. (1995). *"Balancing" of conductances may explain irregular cortical spiking* (Technical Report No. INC-9502). Institute for Neural Computation, University of California at San Diego.

Beran, J. (1994). *Statistics for long-memory processes* (Vol. 61). Boca Raton, FL: Chapman & Hall/CRC Press.

Bishop, P. O., Levick, W. R., & Williams, W. O. (1964). Statistical analysis of the dark discharges of lateral geniculate neurones. *Journal of Physiology, 170,* 598-612.

Brockwell, P. J., & Davis, R. A. (1991). *Time series: Theory and methods* (2nd ed.). New York: Springer-Verlag.

Brown, D., & Feng, J. (1999). Is there a problem matching real and model CV(ISI)? *Neurocomputing, 26–27,* 87–91.

Bugmann, G., Christodoulou, C., & Taylor, J. G. (1997). Role of temporal integration and fluctuation detection in the highly irregular firing of a leaky integrator neuron model with partial reset. *Neural Computation, 9*(5), 985–1000.

Burkitt, A. N. (2000). Interspike interval variability for balanced networks with reversal potentials for large numbers of inputs. *Neurocomputing, 32–33,* 313–321.

Burkitt, A. N. (2001). Balanced neurons: analysis of leaky integrate-and-fire neurons with reversal potentials. *Biological Cybernetics, 85*(4), 247–255.

Burkitt, A. N., & Clark, G. M. (2000). Calculation of interspike intervals for integrate-and-fire neurons with poisson distribution of synaptic inputs. *Neural Computation, 12*(8), 1789-1820.

Burns, B. D., & Webb, A. C. (1976). The spontaneous activity of neurones in the cat's cerebral cortex. *Proceedings of the Royal Society of London B, 194*(1115), 211-223.

Christodoulou, C., & Bugmann, G. (2000). Near poisson-type firing produced by concurrent excitation and inhibition. *Biosystems, 58*(1-3), 41-48.

Christodoulou, C., Clarkson, T. G., Bugmann, G., & Taylor, J. G. (1994). Modelling of the high firing variability of real cortical neurons with the temporal noisy-leaky integrator neuron model. In *Proceedings of 1994 IEEE international conference on neural networks (ICNN'94)* (Vol. 4, p. 2239-2244). New York, NY, USA: IEEE.

Cooper, N. P., & Yates, G. K. (1994). Nonlinear input-output functions derived from the responses of guinea-pig cochlear nerve fibres: variations with characteristic frequency. *Hearing Research, 78*(2), 221–234.

Cox, D. R. (1955). Some statistical methods connected with series of events. *Journal of the Royal Statistical Society, Series B, 17*(2), 129–157.

Cox, D. R. (1967). *Renewal theory.* London: Chapman and Hall.

Cox, D. R. (1984). Long-range dependence: a review. In H. A. David & H. T. David (Eds.), *Statistics: An appraisal* (pp. 55–74). Ames, IA: Iowa State University Press.

Cox, D. R., & Isham, V. (1980). *Point processes.* London: Chapman and Hall.

Cox, D. R., & Lewis, P. A. W. (1966). *The statistical analysis of series of events.* London: Chapman and Hall.

Cox, D. R., & Smith, W. L. (1954). On the superposition of renewal processes. *Biometrika, 41*(1/2), 91–99.

Daley, D. J. (1999). The Hurst index of long-range dependent renewal processes. *Annals of Probability, 27*(4), 2035–2041.

Daley, D. J., Rolski, T., & Vesilo, R. (2000). Long-range dependent point processes and their palm-khinchin distributions. *Advances in Applied Probability, 32*(4), 1051–1063.

Daley, D. J., & Vere-Jones, D. (1988). *An introduction to the theory of point processes.* New York: Springer-Verlag.

Daley, D. J., & Vesilo, R. (1997). Long range dependence of point processes, with queueing examples. *Stochastic Processes and their Applications, 70*(2), 265–282.

Davies, R. B., & Harte, D. S. (1987). Tests for hurst effect. *Biometrika, 74*(1), 95–101.

Davis, R., & Resnick, S. (1986). Limit theory for the sample covariance and correlation functions of moving averages. *Annals of Statistics, 14*(2), 533–558.

Dean, A. F. (1981). The variability of discharge of simple cells in the cat striate cortex. *Experimental Brain Research, 44*(4), 437-440.

DeFelice, L. J., & Isaac, A. (1993). Chaotic states in a random world: relationship between the nonlinear differential equations of excitability and the stochastic properties of ion channels. *Journal of Statistical Physics*, *70*(1–2), 339–354.

Delgutte, B. (1987). Peripheral auditory processing of speech information: Implications from a physiological study of intensity discrimination. In M. Schouten (Ed.), *The psychophysics of speech perception* (p. 333-353). Dordrecht: Nijhof.

Delgutte, B. (1990a). Physiological mechanisms of psychophysical masking: observations from auditory-nerve fibers. *Journal of the Acoustical Society of America*, *87*(2), 791-809.

Delgutte, B. (1990b). Two-tone rate suppression in auditory-nerve fibers: dependence on suppressor frequency and level. *Hearing Research*, *49*(1-3), 225-246.

Delgutte, B. (1996). Physiological models for basic auditory percepts. In H. L. Hawkins, T. A. McMullen, A. N. Popper, & R. R. Fay (Eds.), *Auditory computation* (Vol. 6, p. 157-220). New York: Springer.

Destexhe, A., & Pare, D. (1999). Impact of network activity on the integrative properties of neocortical pyramidal neurons in vivo. *Journal of Neurophysiology*, *81*(4), 1531–1547.

Destexhe, A., & Pare, D. (2000). A combined computational and intracellular study of correlated synaptic bombardment in neocortical pyramidal neurons in vivo. *Neurocomputing*, *32–33*, 113–119.

Doukhan, P., Oppenheim, G., & Taqqu, M. S. (Eds.). (2003). *Theory and applications of long-range dependence.* Boston: Birkhauser.

Enns, E. G. (1970). A stochastic superposition process and an integral inequality for distributions with monotone hazard rates. *Australian Journal of Statistics*, *12*, 44–49.

Fano, U. (1947). Ionization yield of radiations. II. the fluctuations of the number of ions. *Physical Review*, *72*, 26–29.

Feller, W. (1971). *An introduction to probability theory and its applications, Vol. 2* (Second ed.). New York: John Wiley & Sons.

Feng, J. (1997). Behaviours of spike output jitter in the integrate-and-fire model. *Physical Review Letters*, *79*(22), 4505–4508.

Feng, J. (1999). Origin of firing variability of the integrate-and-fire model. *Neurocomputing*, *26–27*, 117–122.

Feng, J. (2001). Is the integrate-and-fire model good enough?—a review. *Neural Networks*, *14*(6–7), 955–975.

Feng, J., & Brown, D. (1998a). Impact of temporal variation and the balance between excitation and inhibition on the output of the perfect integrate-and-fire model. *Biological Cybernetics*, *78*(5), 369–376.

Feng, J., & Brown, D. (1998b). Spike output jitter, mean firing time and coefficient of variation. *Journal of Physics A*, *31*(4), 1239–1252.

Feng, J., & Brown, D. (1999). Coefficient of variation of interspike intervals greater than 0.5. How and when? *Biological Cybernetics*, *80*(5), 291–297.

Feng, J., & Brown, D. (2000a). Impact of correlated inputs on the output of the integrate-and-fire model. *Neural Computation*, *12*(3), 671–692.

Feng, J., & Brown, D. (2000b). Integrate-and-fire models with nonlinear leakage. *Bulletin of Mathematical Biology*, *62*(3), 467–481.

Feng, J., Tirozzi, B., & Brown, D. (1998). Output jitter diverges to infinity, converges to zero or remains constant. In M. Verleysen (Ed.), *6th European symposium on artificial neural networks. ESANN'98* (pp. 39–45). Brussels, Belgium: D-Facto.

Feng, J., & Zhang, P. (2001). Behavior of integrate-and-fire and Hodgkin-Huxley models with correlated inputs. *Physical Review E*, *63*(5, pt.1–2), 051902/1–11.

Gath, I. (1974). Analysis of point process signals applied to motor unit firing patterns I. superposition of independent spike trains. *Mathematical Biosciences*, *22*, 211–222.

Gautrais, J., & Thorpe, S. (1998). Rate coding versus temporal order coding: a theoretical approach. *Biosystems*, *48*(1-3), 57-65.

Geisler, C. D., Deng, L., & Greenberg, S. R. (1985). Thresholds for primary auditory fibers using statistically defined criteria. *Journal of the Acoustical Society of America*, *77*(3), 1102–1109.

Geisler, C. D., Rhode, W. S., & Kennedy, D. T. (1974). Responses to tonal stimuli of single auditory nerve fibers and their relationship to basilar membrane motion in the squirrel monkey. *Journal of Neurophysiology*, *37*(6), 1156–1172.

Gerstein, G., & Mandelbrot, B. (1964). Random walk models for the spike activity of a single neuron. *Biophysical Journal*, *4*, 41–68.

Gruneis, F., Nakao, M., Mizutani, Y., Yamamoto, M., Meesmann, M., & Musha, T. (1993). Further study on 1/f fluctuations observed in central single neurons during REM sleep. *Biological Cybernetics*, *68*(3), 193–198.

Gruneis, F., Nakao, M., & Yamamoto, M. (1990). Counting statistics of 1/f fluctuations in neuronal spike trains. *Biological Cybernetics*, *62*(5), 407–413.

Gruneis, F., Nakao, M., Yamamoto, M., Musha, T., & Nakahama, H. (1989). An interpretation of 1/f fluctuations in neuronal spike trains during dream sleep. *Biological Cybernetics*, *60*(3), 161–169.

Heath, D., Resnick, S., & Samorodnitsky, G. (1998). Heavy tails and long range dependence in on/off processes and associated fluid models. *Mathematics of Operations Research*, *23*(1), 145–165.

Hurst, H. E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, *116*, 770–799.

Jackson, B. S., & Relkin, E. M. (1998). A frequency-dependent saturation evident in rate-intensity functions of the chinchilla auditory nerve. *Hearing Research*, *126*(1–2), 75–83.

Kawase, T., & Liberman, M. C. (1992). Spatial organization of the auditory nerve according to spontaneous discharge rate. *The Journal of Comparative Neurology*, *319*(2), 312–318.

Kelly, O. E., Johnson, D. H., Delgutte, B., & Cariani, P. (1996). Fractal noise strength in auditory-nerve fiber recordings. *Journal of the Acoustical Society of America*, *99*(4 Pt 1), 2210–2220.

Khintchine, A. Y. (1960). *Mathematical methods in the theory of queueing* (Vol. 7). New York: Hafner Publishing Co.

Kiang, N. Y., Watanabe, T., Thomas, C., & Clark, L. F. (1965). *Discharge patterns of single fibers in the cat's auditory nerve.* Cambridge, MA: MIT Press.

Kodama, T., Mushiake, H., Shima, K., Nakahama, H., & Yamamoto, M. (1989). Slow fluctuations of single unit activities of hippocampal and thalamic neurons in cats. I. relation to natural sleep and alert states. *Brain Research*, *487*(1), 26–34.

Kokoszka, P. S., & Taqqu, M. (1995). Fractional arima with stable innovations. *Stochastic Processes and their Applications*, *60*(1), 19–47.

Kokoszka, P. S., & Taqqu, M. S. (1993). Asymptotic dependence of moving average type self-similar stable random fields. *Nagoya Mathematical Journal*, *130*, 85–100.

Kokoszka, P. S., & Taqqu, M. S. (1994). Infinite variance stable arma processes. *Journal of Time Series Analysis*, *15*(2), 203–220.

Kokoszka, P. S., & Taqqu, M. S. (1996). Infinite variance stable moving averages with long memory. *Journal of Econometrics*, *73*(1), 79–99.

Konig, P., Engel, A. K., & Singer, W. (1996). Integrator or coincidence detector? The role of the cortical neuron revisited. *Trends in Neurosciences*, *19*(4), 130–137.

Kuffler, S. W., FitzHugh, R., & Barlow, H. B. (1957). Maintained activity in the cat's retina in light and darkness. *Journal of General Physiology*, *40*, 683-702.

Kulik, R., & Szekli, R. (2001). Sufficient conditions for long-range count dependence of stationary point processes on the real line. *Journal of Applied Probability*, *38*(2), 570–581.

Kumar, A. R., & Johnson, D. H. (1993). Analyzing and modeling fractal intensity point processes. *Journal of the Acoustical Society of America*, *93*(6), 3365–3373.

Lansky, P., & Radil, T. (1987). Statistical inference on spontaneous neuronal discharge patterns. I. Single neuron. *Biological Cybernetics*, *55*(5), 299–311.

Lansky, P., & Smith, C. E. (1989). The effect of a random initial value in neural first-passage-time models. *Mathematical Biosciences*, *93*(2), 191–215.

Lapicque, L. (1907). Recherches quantitatifs sur l'excitation electrique des nerfs traitée comme une polarisation. *Journal of Physiology (Paris)*, *9*, 620-635.

Lapicque, L. (1926). *L'excitabilité en fonction du temps.* Paris, France: Presses Universitaires de France.

Lau, W.-C., Erramilli, A., Wang, J. L., & Willinger, W. (1995). Self-similar traffic generation: the random midpoint displacement algorithm and its properties. In *Icc '95 seattle. communications—gateway to globalization. 1995 ieee international conference on communications* (Vol. 1, pp. 466–472). New York, NY, USA: IEEE.

Lawrance, A. J. (1973). Dependency of intervals between events in superposition processes. *Journal of the Royal Statistical Society. Series B*, *35*(2), 306–315.

Leake, P. A., & Snyder, R. L. (1989). Topographic organization of the central projections of the spiral ganglion in cats. *The Journal of Comparative Neurology*, *281*(4), 612–629.

Leake, P. A., Snyder, R. L., & Hradek, G. T. (1993). Spatial organization of inner hair cell synapses and cochlear spiral ganglion neurons. *The Journal of Comparative Neurology*, *333*(2), 257–270.

Leake, P. A., Snyder, R. L., & Merzenich, M. M. (1992). Topographic organization of the cochlear spiral ganglion demonstrated by restricted lesions of the anteroventral cochlear nucleus. *The Journal of Comparative Neurology*, *320*(4), 468–478.

Ledesma, S., & Derong, L. (2000). Synthesis of fractional gaussian noise using linear approximation for generating self-similar network traffic. *Computer Communication Review*, *30*(2), 4–17.

Liberman, M. C. (1978). Auditory-nerve response from cats raised in a low-noise chamber. *Journal of the Acoustical Society of America*, *63*(2), 442–455.

Liberman, M. C. (1982). Single-neuron labeling in the cat auditory nerve. *Science*, *216*(4551), 1239–1241.

Liberman, M. C. (1991). Central projections of auditory-nerve fibers of differing spontaneous rate. I. anteroventral cochlear nucleus. *The Journal of Comparative Neurology*, *313*(2), 240–258.

Liberman, M. C. (1993). Central projections of auditory nerve fibers of differing spontaneous rate, II: Posteroventral and dorsal cochlear nuclei. *The Journal of Comparative Neurology*, *327*(1), 17–36.

Liebovitch, L. S. (1996). Ion channel kinetics. In P. M. Iannaccone & M. Khokha (Eds.), *Fractal geometry in biological systems* (pp. 31–56). Boca Raton, FL: CRC Press.

Liebovitch, L. S., & Koniarek, J. P. (1992). Ion channel kinetics. *IEEE Engineering in Medicine and Biology Magazine*, *11*(2), 53–56.

Liebovitch, L. S., Scheurle, D., Rusek, M., & Zochowski, M. (2001). Fractal methods to analyze ion channel kinetics. *Methods*, *24*(4), 359–375.

Liebovitch, L. S., & Sullivan, J. M. (1987). Fractal analysis of a voltage-dependent potassium channel from cultured mouse hippocampal neurons. *Biophysical Journal*, *52*(6), 979–988.

Liebovitch, L. S., & Todorov, A. T. (1996). Using fractals and nonlinear dynamics to determine the physical properties of ion channel proteins. *Critical Reviews in Neurobiology*, *10*(2), 169–187.

Liebovitch, L. S., & Toth, T. I. (1990a). Fractal activity in cell membrane ion channels. *Annals of the New York Academy of Sciences*, *591*, 375–391.

Liebovitch, L. S., & Toth, T. I. (1990b). Using fractals to understand the opening and closing of ion channels. *Annals of Biomedical Engineering*, *18*(2), 177–194.

Linebarger, D. A., & Johnson, D. H. (1986). Superposition models of the discharge patterns of units in the lower auditory system. *Hearing Research*, *23*(2), 185–198.

Lowen, S. B., Liebovitch, L. S., & White, J. A. (1999). Fractal ion-channel behavior generates fractal firing patterns in neuronal models. *Physical Review E*, *59*(5), 5970–5980.

Lowen, S. B., Ozaki, T., Kaplan, E., Saleh, B. E. A., & Teich, M. C. (2001). Fractal features of dark, maintained, and driven neural discharges in the cat visual system. *Methods*, *24*(4), 377–394.

Lowen, S. B., & Teich, M. C. (1991). Doubly stochastic poisson point process driven by fractal shot noise. *Physical Review A*, *43*(8), 4192–4215.

Lowen, S. B., & Teich, M. C. (1992). Auditory-nerve action potentials form a nonrenewal point process over short as well as long time scales. *Journal of the Acoustical Society of America*, *92*(2 Pt 1), 803–806.

Lowen, S. B., & Teich, M. C. (1993a). Estimating the dimension of a fractal point process. *Proceedings of the SPIE*, *2036*, 64–76.

Lowen, S. B., & Teich, M. C. (1993b). Fractal auditory-nerve firing patterns may derive from fractal switching in sensory hair-cell ion channels. *AIP Conference Proceedings: Noise in Physical Systems and 1/f Fluctuations*, *285*, 745–748.

Lowen, S. B., & Teich, M. C. (1993c). Fractal renewal processes. *IEEE Transactions on Information Theory*, *39*(5), 1669–1671.

Lowen, S. B., & Teich, M. C. (1993d). Fractal renewal processes generate 1/f noise. *Physical Review E*, *47*(2), 992–1001.

Lowen, S. B., & Teich, M. C. (1995). Estimation and simulation of fractal stochastic point processes. *Fractals*, *3*(1), 183–210.

Lowen, S. B., & Teich, M. C. (1996a). Refractoriness-modified fractal stochastic point processes for modeling sensory-system spike trains. In J. M. Bower (Ed.), *Computational neuroscience* (pp. 447–452). San Diego: Academic.

Lowen, S. B., & Teich, M. C. (1996b). The periodogram and Allan variance reveal fractal exponents greater than unity in auditory-nerve spike trains. *Journal of the Acoustical Society of America*, *99*(6), 3585–3591.

Lowen, S. B., & Teich, M. C. (1997). Estimating scaling exponents in auditory-nerve spike trains using fractal models incorporating refractoriness. In E. R. Lewis, G. R. Long, R. F. Lyon, P. M. Narins, C. R. Steele, & E. Hecht-Pointar (Eds.), *Diversity in auditory mechanics* (pp. 197–204). Singapore: World Scientific.

Mandelbrot, B. B. (1965). Une classe de processus stochastiques homothetiques a soi: application a loi climatologique de h. e. hurst. *Comptes Rendus de l'Acadèmie des Sciences de Paris*, *240*, 3274–3277.

Mandelbrot, B. B. (1969). Long-run linearity, locally gaussian process, h-spectra and infinite variances. *International Economic Review*, *10*(1), 82–111.

Mandelbrot, B. B. (1971). A fast fractional gaussian noise generator. *Water Resources Research*, *7*(3), 543–553.

Mandelbrot, B. B., & van Ness, J. W. (1968). Fractional brownian motions, fractional noises and applications. *SIAM Review*, *10*(4), 422–437.

Mandelbrot, B. B., & Wallis, J. R. (1968). Noah, joseph and operational hydrology. *Water Resources Research*, *4*(5), 909–918.

Mandelbrot, B. B., & Wallis, J. R. (1969a). Computer experiments with fractional gaussian noises. *Water Resources Research*, *5*(1), 228–267.

Mandelbrot, B. B., & Wallis, J. R. (1969b). Robustness of the rescaled range r/s in the measurement of noncyclic long run statistical dependence. *Water Resources Research*, *5*, 967–988.

Mandelbrot, B. B., & Wallis, J. R. (1969c). Some long-run properties of geophysical records. *Water Resources Research, 5*, 321–340.

Mueller, C. G. (1954). A quantitative theory of visiual excitation for the single photoreceptor. *Proceedings of the National Academy of Science USA, 40*, 853-863.

Mushiake, H., Kodama, T., Shima, K., Yamamoto, M., & Nakahama, H. (1988). Fluctuations in spontaneous discharge of hippocampal theta cells during sleep-waking states and pcpa-induced insomnia. *Journal of Neurophysiology, 60*(3), 925–939.

Noda, H., & Adey, W. R. (1970). Firing variability in cat association cortex during sleep and wakefulness. *Brain Research, 18*(3), 513-526.

Ohlemiller, K. K., Echteler, S. M., & Siegel, J. H. (1991). Factors that influence rate-versus-intensity relations in single cochlear nerve fibers of the gerbil. *Journal of the Acoustical Society of America, 90*(1), 274–287.

Palmer, A. R., & Evans, E. F. (1980). Cochlear fibre rate-intensity functions: no evidence for basilar membrane nonlinearities. *Hearing Research, 2*(3–4), 319–326.

Paxson, V. (1997). Fast, approximate synthesis of fractional gaussian noise for generating self-similar network traffic. *Computer Communication Review, 27*(5), 5–18.

Relkin, E. M., & Doucet, J. R. (1991). Recovery from prior stimulation. I: Relationship to spontaneous firing rates of primary auditory neurons. *Hearing Research, 55*(2), 215–222.

Relkin, E. M., & Pelli, D. G. (1987). Probe tone thresholds in the auditory nerve measured by two-interval forced-choice procedures. *Journal of the Acoustical Society of America, 82*(5), 1679–1691.

Relkin, E. M., & Turner, C. W. (1988). A reexamination of forward masking in the auditory nerve. *Journal of the Acoustical Society of America, 84*(2), 584–591.

Resnick, S., & Samorodnitsky, G. (1997). Performance decay in a single server exponential queueing model with long range dependence. *Operations Research, 45*(2), 235–243.

Ryu, B. K. (1997). Fractal network traffic modeling: past, present, and future. In *Proc. 35th allerton conference on communication, control, and computing.* Urbana, IL.

Ryu, B. K., & Lowen, S. B. (1996). Point process approaches to the modeling and analysis of self-similar traffic: Part I—Model construction. In *Proceedings ieee infocom '96. the conference on computer communications. fifteenth annual joint conference of the ieee computer societies. networking the next generation* (Vol. 3, pp. 1468–1475). Los Alamitos, CA, USA: IEEE Comput. Soc. Press.

Ryu, B. K., & Lowen, S. B. (1997). Point process approaches to the modeling and analysis of self-similar traffic: Part II—Applications. In *Proc. 5th int'l conf. telecommunications systems, modeling, and analysis* (pp. 62–70). Nashville, TN.

Ryu, B. K., & Lowen, S. B. (1998). Point process models for self-similar network traffic, with applications. *Communications in Statistics. Stochastic Models*, *14*(3), 735–761.

Ryu, B. K., & Lowen, S. B. (2000). Fractal traffic models for internet simulation. In S. Tohme & M. Ulema (Eds.), *Proceedings iscc 2000. fifth ieee symposium on computers and communications* (pp. 200–206). Los Alamitos, CA, USA: IEEE Comput. Soc.

Ryugo, D. K., & May, S. K. (1993). The projections of intracellularly labeled auditory nerve fibers to the dorsal cochlear nucleus of cats. *The Journal of Comparative Neurology*, *329*(1), 20–35.

Ryugo, D. K., & Rouiller, E. M. (1988). Central projections of intracellularly labeled auditory nerve fibers in cats: morphometric correlations with physiological properties. *The Journal of Comparative Neurology*, *271*(1), 130–142.

Sachs, M. B., & Abbas, P. J. (1974). Rate versus level functions for auditory-nerve fibers in cats: tone-burst stimuli. *Journal of the Acoustical Society of America*, *56*(6), 1835–1847.

Sachs, M. B., Winslow, R. L., & Sokolowski, B. H. A. (1989). A computational model for rate-level functions from cat auditory-nerve fibers. *Hearing Research*, *41*(1), 61–69.

Sakai, Y., Funahashi, S., & Shinomoto, S. (1999). Temporally correlated inputs to leaky integrate-and-fire models can reproduce spiking statistics of cortical neurons. *Neural Networks*, *12*(7–8), 1181–1190.

Salinas, E., & Sejnowski, T. J. (2000). Impact of correlated synaptic input on output firing rate and variability in simple neuronal models. *Journal of Neuroscience*, *20*(16), 6193–6209.

Samorodnitsky, G., & Taqqu, M. S. (1994). *Stable non-gaussian random processes: Stochastic models with infinite variance.* Boca Raton: Chapman & Hall/CRC.

Shadlen, M. N., & Newsome, W. T. (1994). Noise, neural codes and cortical organization. *Current Opinion in Neurobiology*, *4*(4), 569–579.

Shadlen, M. N., & Newsome, W. T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *Journal of Neuroscience*, *18*(10), 3870–3896.

Shinomoto, S., & Sakai, Y. (1998). Spiking mechanisms of cortical neurons. *Philosophical Magazine B*, *77*(5), 1549–1555.

Shinomoto, S., Sakai, Y., & Funahashi, S. (1999). The Ornstein-Uhlenbeck process does not reproduce spiking statistics of neurons in prefrontal cortex. *Neural Computation*, *11*(4), 935–951.

Shinomoto, S., & Tsubo, Y. (2001). Modeling spiking behavior of neurons with time-dependent Poisson processes. *Physical Review E, 64*(4–1), 041910.

Siebert, W. M. (1965). Some implications of the stochastic behavior of primary auditory neurons. *Kybernetik, 2*(5), 206-215.

Smith, D. R., & Smith, G. K. (1965). A statistical analysis of the continual activity of single cortical neurones in the cat unanaesthetized isolated forebrain. *Biophysical Journal, 5*, 47-74.

Softky, W. R. (1994). Sub-millisecond coincidence detection in active dendritic trees. *Neuroscience, 58*(1), 13–41.

Softky, W. R. (1995). Simple codes versus efficient codes. *Current Opinion in Neurobiology, 5*(2), 239–247.

Softky, W. R., & Koch, C. (1992). Cortical cells should fire regularly, but do not. *Neural Computation, 4*(5), 643–646.

Softky, W. R., & Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *Journal of Neuroscience, 13*(1), 334–350.

Stevens, C. F., & Zador, A. M. (1998). Input synchrony and the irregular firing of cortical neurons. *Nature Neuroscience, 1*(3), 210–217.

Taqqu, M. S., & Levy, J. B. (1986). Using renewal processes to generate long-range dependence and high variability. In E. Eberlein & M. S. Taqqu (Eds.), *Dependence in probability and statistics: A survey of recent results* (Vol. 11, pp. 73–89). Boston: Birkhauser.

Taqqu, M. S., Teverovsky, V., & Willinger, W. (1995). Estimators for long-range dependence: an empirical study. *Fractals, 3*(4), 785–798.

Taqqu, M. S., Willinger, W., & Sherman, R. (1997). Proof of a fundamental result in self-similar traffic modeling. *Computer Communication Review, 27*(2), 5–23.

Teich, M. C. (1989). Fractal character of the auditory neural spike train. *IEEE Transactions on Biomedical Engineering, 36*(1), 150–160.

Teich, M. C. (1992). Fractal neuronal firing patterns. In T. McKenna, J. Davis, & S. F. Zornetzer (Eds.), *Single neuron computation* (pp. 589–625). Boston: Academic Press.

Teich, M. C. (1996). Self-similarity in sensory neural signals. In J. K. J. Li & S. S. Reisman (Eds.), *Proceedings of the IEEE 22nd annual northeast bioengineering conference* (pp. 36–37). New Brunswick, NJ, USA: IEEE; New York, NY, USA.

Teich, M. C., Heneghan, C., Lowen, S. B., Ozaki, T., & Kaplan, E. (1997). Fractal character of the neural spike train in the visual system of the cat. *Journal of the Optical Society of America A, 14*(3), 529–546.

Teich, M. C., Johnson, D. H., Kumar, A. R., & Turcott, R. G. (1990). Rate fluctuations and fractional power-law noise recorded from cells in the lower auditory pathway of the cat. *Hearing Research*, *46*(1–2), 41–52.

Teich, M. C., & Khanna, S. M. (1985). Pulse-number distribution for the neural spike train in the cat's auditory nerve. *Journal of the Acoustical Society of America*, *77*(3), 1110-1128.

Teich, M. C., & Lowen, S. B. (1994). Fractal patterns in auditory nerve-spike trains. *IEEE Engineering in Medicine and Biology Magazine*, *13*(2), 197–202.

Teich, M. C., Turcott, R. G., & Lowen, S. B. (1990). The fractal doubly stochastic poisson point process as a model for the cochlear neural spike train. In P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, & C. R. Steele (Eds.), *The mechanics and biophysics of hearing* (Vol. 87, pp. 354–361). New York: Springer-Verlag.

Teich, M. C., Turcott, R. G., & Siegel, R. M. (1996). Temporal correlation in cat striate-cortex neural spike trains. *IEEE Engineering in Medicine and Biology Magazine*, *15*(5), 79–87.

Thorpe, S., Delorme, A., & Van Rullen, R. (2001). Spike-based strategies for rapid processing. *Neural Networks*, *14*(6-7), 715-725.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520-522.

Thorpe, S., & Gautrais, J. (1997). How can the visual system process a natural scene in under 150 ms? on the role of asynchronous spike propagation. In M. Verleysen (Ed.), *5th european symposium on artificial neural networks esann '97* (p. 79-84). Brussels, Belgium: D facto.

Thurner, S., Lowen, S. B., Feurstein, M. C., Heneghan, C., Feichtinger, H. G., & Teich, M. C. (1997). Analysis, synthesis, and estimation of fractal-rate stochastic point processes. *Fractals*, *5*, 565–596.

Troyer, T. W., & Miller, K. D. (1997). Physiological gain leads to high ISI variability in a simple model of a cortical regular spiking cell. *Neural Computation*, *9*(5), 971–983.

Tsodyks, M. V., & Sejnowski, T. (1995). Rapid state switching in balanced cortical network models. *Network*, *6*(2), 111–124.

Tuckwell, H. C. (1988). *Introduction to theoretical neurobiology Vol. 2* (Vol. 2). Cambridge, U.K.: Cambridge University Press.

Turcott, R. G., Barker, D. R., & Teich, M. C. (1995). Long-duration correlation in the sequence of action potentials in an insect visual interneuron. *Journal of Statistical Computation and Simulation*, *52*(3), 253–271.

Turcott, R. G., Lowen, S. B., Li, E., Johnson, D. H., Tsuchitani, C., & Teich, M. C. (1994). A nonstationary poisson point process describes the sequence of action potentials over long time scales in lateral-superior-olive auditory neurons. *Biological Cybernetics*, *70*(3), 209–217.

Turcott, R. G., & Teich, M. C. (1996). Fractal character of the electrocardiogram: distinguishing heart-failure and normal patients. *Annals of Biomedical Engineering*, *24*(2), 269–293.

Usher, M., Stemmler, M., Koch, C., & Olami, Z. (1994). Network amplification of local fluctuations causes high spike rate variability, fractal firing patterns and oscillatory local field potentials. *Neural Computation*, *6*(5), 795–836.

Usher, M., Stemmler, M., & Olami, Z. (1995). Dynamic pattern formation leads to 1/f noise in neural populations. *Physical Review Letters*, *74*(2), 326–329.

Van Rullen, R., & Thorpe, S. J. (2001). Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex. *Neural Computation*, *13*(6), 1255-1283.

Viemeister, N. F. (1988). Psychophysical aspects of auditory intensity coding. In G. M. Edelman, W. E. Gall, & W. M. Cowan (Eds.), *Auditory function: Neurobiological bases of hearing* (p. 213-241). New York: John Wiley & Sons.

Vreeswijk, C. van, & Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*, *274*(5293), 1724–1726.

Vreeswijk, C. van, & Sompolinsky, H. (1998). Chaotic balanced state in a model of cortical circuits. *Neural Computation*, *10*(6), 1321–1371.

Wang, X., & Sachs, M. B. (1993). Neural encoding of single-formant stimuli in the cat. I. responses of auditory nerve fibers. *Journal of Neurophysiology*, *70*(3), 1054–1075.

Wilbur, W. J., & Rinzel, J. (1983). A theoretical basis for large coefficient of variation and bimodality in neuronal interspike interval distributions. *Journal of Theoretical Biology*, *105*(2), 345–368.

Willinger, W., Taqqu, M. S., Sherman, R., & Wilson, D. V. (1995). Self-similarity through high-variability: statistical analysis of ethernet lan traffic at the source level. *Computer Communication Review*, *25*(4), 100–113.

Willinger, W., Taqqu, M. S., Sherman, R., & Wilson, D. V. (1997). Self-similarity through high-variability: statistical analysis of ethernet lan traffic at the source level. *IEEE/ACM Transactions on Networking*, *5*(1), 71–86.

Winslow, R. L., & Sachs, M. B. (1988). Single-tone intensity discrimination based on auditory-nerve rate responses in backgrounds of quiet, noise, and with stimulation of the crossed olivocochlear bundle. *Hearing Research*, *35*(2-3), 165-189.

Winter, I. M., & Palmer, A. R. (1991). Intensity coding in low-frequency auditory-nerve fibers of the guinea pig. *Journal of the Acoustical Society of America, 90*(4 Pt 1), 1958-1967.

Winter, I. M., Robertson, D., & Yates, G. K. (1990). Diversity of characteristic frequency rate-intensity functions in guinea pig auditory nerve fibres. *Hearing Research, 45*(3), 191–202.

Wise, M. E. (1981). Spike interval distributions for neurons and random walks with drift to a fluctuating threshold. In C. Taillie, G. Patil, & B. Baldessari (Eds.), *Statistical distributions in scientific work* (Vol. 6, pp. 211–231). Boston, MA: Reidel.

Yamamoto, M., Nakahama, H., Shima, K., Kodama, T., & Mushiake, H. (1986). Markov-dependency and spectral analyses on spike-counts in mesencephalic reticular neurons during sleep and attentive states. *Brain Research, 366*(1–2), 279–289.

Young, E. D., & Barta, P. E. (1986). Rate responses of auditory nerve fibers to tones in noise near masked threshold. *Journal of the Acoustical Society of America, 79*(2), 426–442.

VITA

NAME OF AUTHOR: Brian Scott Jackson

PLACE OF BIRTH: Charleston, South Carolina

DATE OF BIRTH: February 7, 1969

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:
    Syracuse University, Syracuse, New York
    University of Toledo, Toledo, Ohio

DEGREES AWARDED:

    Master of Science in Mathematics, 1996, Syracuse University

    Bachelor of Science in Electrical Engineering, 1992, University of Toledo

AWARDS AND HONORS:

    University Fellow, Syracuse University, 1993–1996

    Computational Neuroscience Fellow, Syracuse University, 1997

    University Fellow, Syracuse University, 1997–1998

PROFESSIONAL EXPERIENCE:

Teaching Assistant, Department of Bioengineering, Syracuse University, 1994–1995

Teaching Assistant, Department of Mathematics, Syracuse University, 1996–1997

Teaching Assistant, Department of Mathematics, Syracuse University, 2001–2002

Research Assistant, Institute for Sensory Research, Syracuse University
    Auditory Research Laboratory, 1994
    Vision Research Laboratory, 1995
    Auditory Research Laboratory, 1996
    Auditory Research Laboratory, 1998–2001
    Auditory Research Laboratory, 2002–2003