Fractals, Vol. 3, No. 1 (1995) 183–210

ESTIMATION AND SIMULATION OF FRACTAL STOCHASTIC POINT PROCESSES

STEVEN B. LOWEN and MALVIN C. TEICH
Columbia University, Department of Electrical Engineering, New York, NY 10027

Received November 21, 1994; Revised Manuscript Accepted December 2, 1994

Abstract

We investigate the properties of fractal stochastic point processes (FSPPs). First, we define FSPPs and develop several mathematical formulations for these processes, showing that over a broad range of conditions they converge to a particular form of FSPP. We then provide examples of a wide variety of phenomena for which they serve as suitable models. We proceed to examine the analytical properties of two useful fractal dimension estimators for FSPPs, based on the second-order properties of the points. Finally, we simulate several FSPPs, each with three specified values of the fractal dimension. Analysis and simulation reveal that a variety of factors confound the estimate of the fractal dimension, including the finite length of the simulation, structure or type of FSPP employed, and fluctuations inherent in any FSPP. We conclude that for segments of FSPPs with as many as 10^6 points, the fractal dimension can be estimated only to within ± 0.1.

1. DEFINITION AND PROPERTIES OF FRACTAL STOCHASTIC POINT PROCESSES

1.1 Fractal Stochastic Processes

The definition of a stochastic process involves a complete description of all possible joint probabilities of the various events occurring in the process. Different statistics provide complementary windows into the process; no single statistic can in general describe a stochastic process completely. One definition of a *fractal* stochastic process is a stochastic process in which the sample paths of the processes have non-integral dimensions; the expected measure of the sample path included within some radius scales with the size of the radius.[1] Since this

is but one statistic of the process, we call a stochastic process fractal if several of the relevant statistics exhibit scaling.[2-5]

Such scaling leads mathematically to power-law dependencies in the scaled quantities. Consider a statistic f which depends continuously on the scale x over which measurements are taken. Suppose changing the scale by a factor a effectively scales the statistic by some other factor $g(a)$, related to the factor but independent of the argument:

$$f(ax) = g(a)f(x)\,. \tag{1}$$

The only non-trivial solution of this scaling equation is

$$f(x) = cg(x) \quad \text{with } g(x) = x^D\,, \tag{2}$$

for some constants c and D.[6,7] Thus fractals and power-law forms of their statistics are closely related. Each statistic which scales will therefore provide a dimension; for a (mono-) fractal process all are simply related, yielding a single dimension for the process.

In general, fractal scaling in one statistic does not necessarily imply fractal scaling in other statistics. Consider a stationary, continuous-time process with a power spectral density which scales with frequency, so that $S(\omega) \propto \omega^{-D}$ for some range of frequencies (or, equivalently, times) and some positive exponent D. For $0 < D < 1$, the autocorrelation function will indeed scale with delay time over the same range of times, yielding $R(\tau) \propto \tau^{D-1}$.[6] However, for $D > 1$, the autocorrelation approaches a constant value in this same range of times, and in fact does not scale with the delay time τ. Therefore we do not call this process fractal.[6]

1.2 Fractal Stochastic Point Processes (FSPPs)

Some random phenomena occur at discrete times (or locations), with the individual events largely identical, such as the detection of particles from radioactive decay. A stochastic point process is a mathematical construction which represents these events as random points in a space. Such a process may be called fractal when a number of the relevant statistics exhibit scaling with related scaling exponents, indicating that the represented phenomenon contains clusters of points over all (or a relatively large set of) time or length scales.[1,3-5] In this paper we consider point processes on a line, so that the associated dimension must lie between a lower limit of zero (the dimension of a point), and an upper limit of unity (the dimension of a line).[1]

Relevant statistics for an FSPP are the power spectral density (PSD), coincidence rate (CR), Fano factor (FF), and interevent-time survivor function (ISF). The PSD $S(\omega)$ is much the same as for continuous-time processes; it provides a measure of how the power in a process is concentrated in various frequency bands.[8] The CR measures the correlation between pairs of events with a specified time delay between them, regardless of intervening events, and is related to the autocorrelation function used with continuous processes.[9] It is defined as

$$G(\tau) \equiv \lim_{\Delta \to 0} \frac{\Pr\{\mathcal{E}(0,\,\Delta) \text{ and } \mathcal{E}(\tau,\,\tau+\Delta)\}}{\Delta^2}\,, \tag{3}$$

where $\mathcal{E}(s,\,t)$ denotes the occurrence of at least one event of the point process in the interval $[s,\,t)$. A particularly useful statistic is the Fano factor $F(T)$, denoted FF, which is defined as the variance of the number of counts in a specified time window T divided by the mean number of counts. Finally, the ISF highlights the behavior of the times between adjacent events, but reveals none of the information contained in relationships among these times, such as correlation between adjacent time intervals.

For relevant time and frequency ranges, and with $0 < D < 1$, any one of the following relationships

$$\begin{aligned} &\text{PSD:} && S(\omega)/S_0 \approx 1 + (\omega/\omega_0)^{-D} \\ &\text{CR:} && G(\tau)/G_0 \approx 1 + (\tau/\tau_0)^{D-1} \\ &\text{FF:} && F(T) \approx 1 + (T/T_0)^{D}\,, \end{aligned} \tag{4}$$

implies the other two, with

$$\begin{aligned} G_0 &= S_0^2 \\ \omega_0^D T_0^D &= \cos(\pi D/2)\Gamma(D+2) \\ S_0 \tau_0^{1-D} T_0^D &= D(D+1)/2\,. \end{aligned} \tag{5}$$

In principle, any of the three statistics in Eq. (4) could be solved for the dimension D of an FSPP, and in some cases the form of the ISF also permits the determination of a value of the fractal dimension. When applied to a real data set, each would yield a different number, thus giving rise to a family of dimensions. For practical estimation of the dimension, however, we focus on the FF and the PSD, rather than on the CR and the ISF. We do this

because the CR involves essentially a double derivative in its definition, and therefore its estimates are particularly noisy. Furthermore, for typical FSPPs the CR exceeds its asymptotic value at $\tau \to \infty$ by only a small fraction at any practical value of τ, so that determining the fractal dimension with this small excess presents serious difficulties. We reject the ISF because most FSPPs contain interleaved fractal clusters, so that the fractal structure is not evident in the first-order interevent-time statistics. Thus for general FSPPs, the ISF does not scale over a significant range of interevent times, the salient exception being the fractal renewal point process (see Sec. 2.2).

For this particular case only, another family of dimensions becomes relevant: the well-known generalized fractal dimension.[10–12] If a data segment of length L is divided into intervals of length T, with Z_n representing the number of points in the nth interval, then the generalized dimension D_q of a point process is defined as

$$D_q \equiv \frac{1}{q-1} \lim_{T \to 0} \frac{\log(\sum Z_n^q)}{\log(T)}, \tag{6}$$

where the sum extends over all non-empty intervals. Particular cases are the capacity or box-counting dimension $\lim_{q \to 0+} D_q$, the information dimension $\lim_{q \to 1} D_q$, and the correlation dimension D_2. The sum is a form of time averaging; for a stochastic point process, it is convenient to replace the sum by the product of L/T and the expected value of Z^q. For a general fractal point process, analytical values of the D_q will not equal analytical values of the dimensions obtained from the PSD or FF, and in fact the D_q will often assume integer values.

2. MATHEMATICAL FORMULATIONS OF FRACTAL STOCHASTIC POINT PROCESSES

In previous work we defined several FSPPs and derived a number of their statistics.[3,5] We briefly summarize these here, together with the clustered Poisson point process model of Grüneis and colleagues.[13] We then proceed to demonstrate that under a broad range of conditions, superpositions of arbitrary FSPPs converge to a particular process, the fractal-Gaussian-noise-driven Poisson process. For all of the processes considered, the scaling relations in Eqs. (4) and (5) hold.

2.1 Fractal-Shot-Noise-Driven Poisson Point Process

The one-dimensional homogeneous Poisson point process (HPP) is perhaps the simplest stochastic point process.[14] The HPP is characterized by a single constant quantity, its rate, which is the number of events expected to occur in a unit interval. A fundamental property of the HPP is that it is memoryless; given this rate, knowledge of the entire history and future of a given realization of an HPP yields no information about the behavior of the process at the present.

Other point processes do not share this memoryless property and therefore cannot be described in terms of a constant rate. An important example of a nonhomogeneous point process is the doubly stochastic Poisson point process (DSPP).[15] For this point process, the rate itself varies stochastically. Thus the DSPP displays two forms of randomness: that associated with the stochastically varying rate, and that associated with the underlying Poisson nature of the process even if its rate were constant.

The fractal-shot-noise-driven Poisson point process (FSNDP)[3] is a special case of the DSPP. The rate of the Poisson process is determined by fractal shot noise,[2,16,17] which is itself a filtered version of another homogeneous Poisson point process. Figure 1 schematically illustrates the FSNDP as a two-stage stochastic process.[18] The first stage is an HPP with constant rate μ. Its output $M(t)$ becomes the input to a linear filter with a power-law decaying impulse response function

$$h(t) = \begin{cases} k/t^{1-D/2} & \text{for } A < t < B, \\ 0 & \text{otherwise}, \end{cases} \tag{7}$$

where D is the dimension, A and B are cutoff parameters, and k is a normalization constant. This filter produces fractal shot noise $I(t)$ at its output, which then becomes the time-varying rate for the last stage, a second Poisson point process. The resulting point process $N(t)$ is not homogeneous, but rather reflects the variations of the fractal-shot-noise driving process. Thus the two forms of randomness inherent in the DSPP are, in the particular case of the FSNDP, two separate Poisson processes, linked by a power-law-decaying linear filter.

2.2 Fractal Renewal Point Process

Perhaps the simplest FSPP is the standard fractal renewal process (FRP).[5,19–22] For the standard

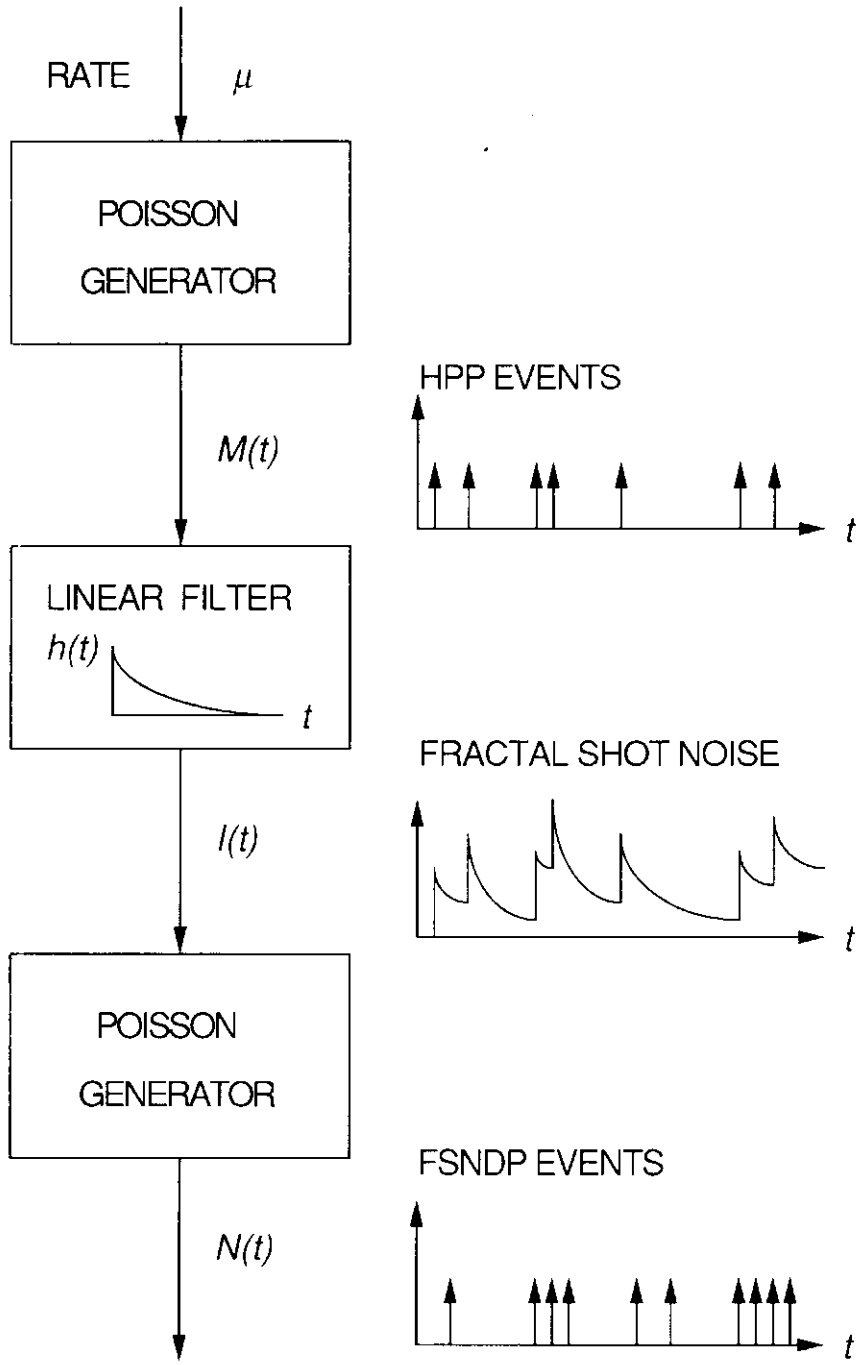

Fig. 1 A primary homogeneous Poisson point process $M(t)$ with constant rate μ serves as the input to a linear filter with impulse response function $h(t)$. The continuous-time stochastic process $I(t)$ at the output of this filter is shot noise, which serves as the random rate for another Poisson point process, whose output is $N(t)$. $N(t)$ is a special doubly stochastic Poisson point process (DSPP), known as a shot-noise driven Poisson point process (SNDP). If $h(t)$ decays in a power-law fashion, then $I(t)$ is fractal shot noise and $N(t)$ is a fractal SNDP or FSNDP.

FRP, the times between adjacent events are independent random variables T drawn from the same fractal probability distribution. Figure 2(a) provides a schematic representation of this point process. In particular, the survivor function $\mathcal{S}(t)$ decays essentially as a power law

$$\mathcal{S}(t) = \Pr\{T > t\} = \begin{cases} k/t^D - k/B^D & \text{for } A < t < B\,, \\ 0 & \text{otherwise}\,, \end{cases} \tag{8}$$

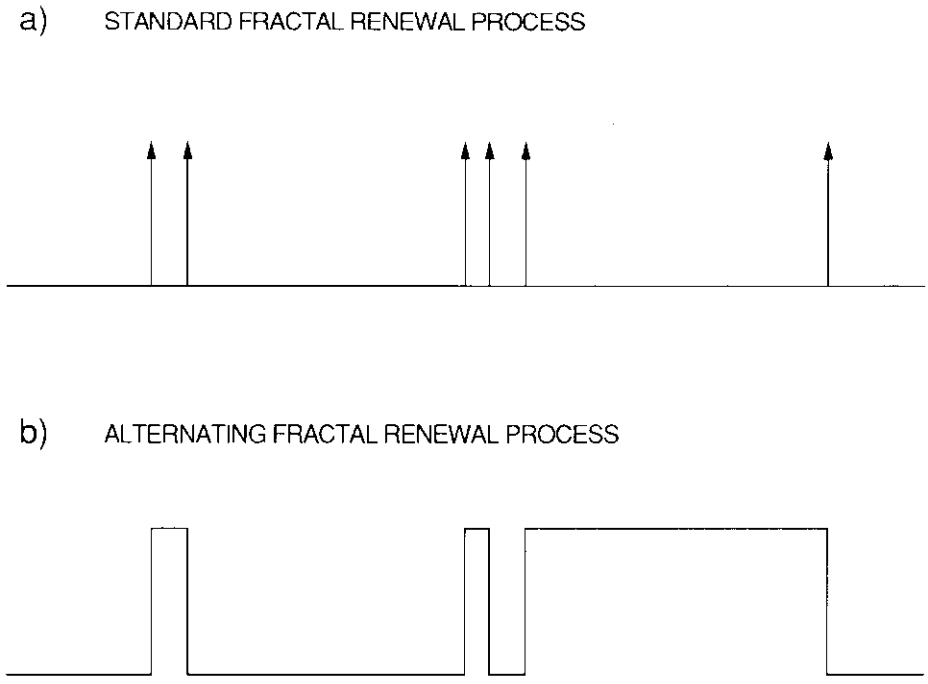

Fig. 2 Sample functions of fractal renewal processes. Interevent times are power-law distributed. **(a)** The standard fractal renewal process (FRP) consists of Dirac δ functions and is zero-valued elsewhere. **(b)** The alternating FRP switches between values of zero and unity.

with D, A, B and k defined as following Eq. (7). The FRP exhibits fractal behavior over time scales lying between A and B. This process is fully fractal: the power spectral density, coincidence rate, Fano factor, and even the interevent-time survivor function all exhibit scaling as in Eq. (1) with the same power-law exponent D (or $D-1$ for the coincidence rate). Further, for this process the capacity or box-counting dimension D_0 assumes the value D;[5] since the FRP is ergodic, the generalized fractal dimension D_q becomes independent of the index q,[12] so $D_q = D_0 = D$ for all q, and all fractal dimensions coincide.

A different point process results from the superposition of a number of independent FRPs; however, for this combined process, and indeed for any FSPP besides the FRP, the interevent-time survivor function no longer scales, and the generalized dimensions D_q no longer equal D, although the PSD, CR, and FF retain their scaling behavior.

The standard FRP described above is a point process, consisting of a set of points or marks on the time axis; however, it may be recast as a real-valued process which alternates between two values, for example zero and unity. This alternating FRP would then start at a value of zero (for example), and then switch to a value of unity at a time corresponding to the first event in the standard FRP. At the second such event in the standard FRP, the alternating FRP would switch back to zero, and would proceed to switch back and forth at every successive event of the standard FRP. Thus the alternating FRP is a Bernoulli process, with times between transitions given by the same interevent distribution as in the standard FRP, as portrayed in Fig. 2(b).

2.3 Fractal-Binomial-Noise-Driven Poisson Point Process

A number of independent, identical alternating fractal renewal processes may be added together, yielding a binomial process with the same fractal dimension as the single alternating FRP.[5] This binomial process can serve as a rate function for a Poisson process; the fractal-binomial-noise-driven Poisson point process (FBNDP) results. It is schematized in Fig. 3. The FSNDP (see Sec. 2.1) and the FBNDP thus both belong to the fractal DSPP family, comprising a real-valued fractal rate function driving a Poisson point process. They differ only in how the fractal rate functions are constructed.

2.4 Cluster Poisson Point Process

Other important formulations for FSPPs exist. Grüneis and colleagues defined a clustered Poisson

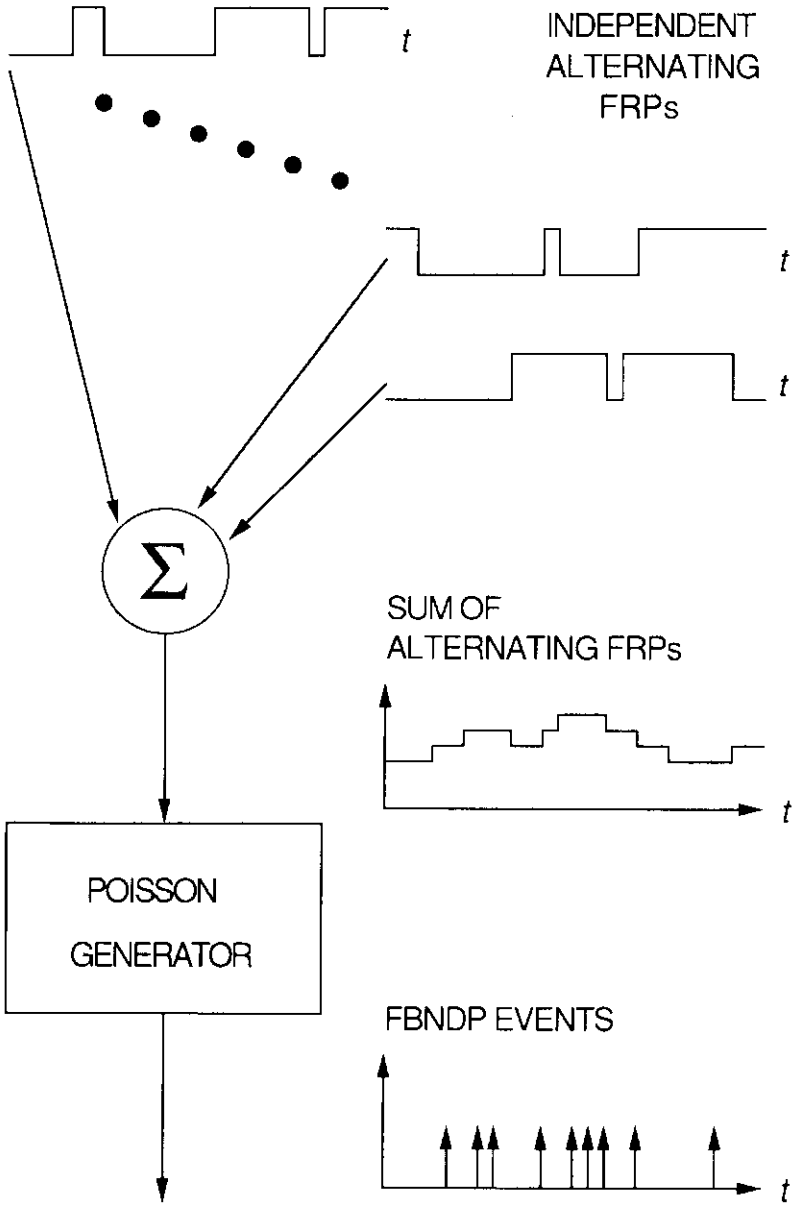

Fig. 3 A sum of several independent and identical alternating FRPs (top) are added (center) to produce a fractal binomial noise process which serves as the rate function for a Poisson point process (bottom). The result is the fractal-binomial-noise-driven Poisson point process (FBNDP), a fractal DSPP.

point process in which each member in a sequence of primary events from an HPP gives rise to a train of secondary events (as in the FSNDP), but where the events in the train form a segment of a renewal process (usually also an HPP, but often with a different rate), with a fractal (power-law distributed) duration.[13] The resulting process indeed exhibits power-law scaling in the same statistics as other FSPPs.[23] The cluster Poisson point process is therefore a Bartlett-Lewis-type process, whereas the FSNDP and FBNDP are processes of the Neyman-Scott type.[24]

2.5 Fractal-Gaussian-Noise-Driven Poisson Point Process

Under suitable conditions, the underlying fractal shot noise of the FSNDP converges to a Gaussian probability density, as provided by the central limit theorem, and therefore becomes fractal Gaussian noise.[2] The resulting point process then becomes a fractal-Gaussian-noise-driven Poisson process (FGNDP).[3] The superposition of many independent standard FRPs, mentioned above, also converges to a FGNDP for certain ranges of

parameters, as does the FBNDP constructed from many alternating FRPs.[5] Finally, the Poisson transform of fractal Gaussian noise (FGN) provides a direct route to the FGNDP. FGN, in turn, may be obtained by any number of methods.[25–28]

We proceed to demonstrate that this convergence to the FGNDP applies for superpositions of large numbers of arbitrary FSPPs as long as they are well behaved. Consider a set of independent point processes $\{N_n(t)\}$, with identical statistics, and define their superposition $N_T(t) \equiv \sum_n N_n(t)$. Let the power spectral density of these processes $S_n(\omega)$ scale with ω as in Eq. (4), so that the constituent point processes $N_n(t)$ are fractal. Then the superposition $N_T(t)$ is also fractal. Finally, we require that the individual point processes have finite second counting moments. For the FSNDP this means that the impulse-response function $h(t)$ has finite area, whereas for the clustered Poisson process the clusters must have a finite expected length. Under these conditions, for an arbitrary set of point processes $\{N_n(t)\}$, the superposition process $N_T(t)$ converges to a DSPP in the limit of a large number of component processes.[29] Since $N_T(t)$ may be represented as a DSPP, then it must have a stochastic rate function $\lambda(t)$ with the same scaling power spectral density as in the component processes (except for a constant term). Since the second counting moments of the component processes are finite, the first and second counting moments of $N_T(t)$ must also be finite, and therefore so are those of $\lambda(t)$. Therefore the Central Limit Theorem applies, and $\lambda(t)$ must converge to a Gaussian distribution in the limit of a large number of component processes. Thus $\lambda(t)$ converges to a Gaussian process with a fractal power spectral density: fractal Gaussian noise. Since $\lambda(t)$ serves as the rate for the DSPP $N_T(t)$, then $N_T(t)$ converges to the FGNDP.

The FGNDP is important because Gaussian processes are ubiquitous, well understood, and may be completely described by their means and autocovariance functions; thus comparison with experiment is facilitated. In particular, the dimension of an FGNDP is identical to the dimension of the FGN process which serves as the rate function of the Poisson process.

3. APPLICATIONS OF FRACTAL STOCHASTIC POINT PROCESSES

Many phenomena are readily represented by FSPPs. We provide several examples drawn from the physical and biological sciences.

3.1 Trapping Times in Amorphous Semiconductors

A multiple trapping model has been used to show how traps which are exponentially distributed over a large range of energies lead to a power-law decay of current in an amorphous semiconductor.[30–33] If a pulse of light strikes such a semiconductor, the many carriers excited out of their traps will be available to carry current until they are recaptured by a trap, which happens relatively quickly. At some point each carrier will be released from its trap by thermal excitation and become mobile for a time, and then be recaptured by another trap. For exponentially distributed energy states with identical capture cross sections, the electrons tend to be trapped in shallow states at first, but the probability of being caught in a deep trap increases as time progresses. This leads to a current that decreases as a power-law function of time.

The multiple trapping model may be usefully recast in terms of a standard FRP.[5,20] Consider an amorphous semiconductor with localized states (traps) that are exponentially distributed with parameter E_0 between a minimum energy E_L of the order of $\kappa\mathcal{T}$, where κ is Boltzmann's constant and $\mathcal{T}$ is the absolute temperature; and a maximum energy E_H determined by the bandgap of the material. For a particular trap with energy $\mathcal{E}$, the corresponding mean waiting time is

$$\tau = \tau_0 \exp(\mathcal{E}/\kappa\mathcal{T}), \tag{9}$$

where τ_0 is the average vibrational period of the atoms in the semiconductor. If we define characteristic time cutoffs $A \equiv \tau_0 \exp(E_L/\kappa\mathcal{T})$ and $B \equiv \tau_0 \exp(E_H/\kappa\mathcal{T})$, and the power-law exponent $D \equiv \kappa\mathcal{T}/E_0$, then the mean waiting time τ has a density which decays as a power law between these two cutoffs. Each trap holds carriers for times that are exponentially distributed given the conditional parameter τ, and averaging this exponential density over all possible values of τ yields the unconditional trapping time density, which is itself approximately power law:

$$p(t) \approx D\Gamma(D+1)A^D t^{-(D+1)}, \tag{10}$$

for $A \ll t \ll B$. Thus each carrier will be trapped for a period that is essentially power-law distributed.

Upon escaping from a trap, the carrier can conduct current for a short time until it is again captured by another trap. Thus each carrier executes a series of current-carrying jumps well described by a standard FRP. Assuming that each carrier acts independently of the others, the action of the carriers as a whole can be modeled as the superposition of a collection of such component processes, which converges to the FGNDP in the limit of a large number of carriers. Again, both experimental[34] and theoretical[35,36] results point to a power-law or fractal decay in the power spectral density, while other statistics also show scaling behavior.

3.2 Noise and Traffic in Communications Systems

Burst noise occurs in many communications systems and is characterized by relatively brief noise events which cluster together, separated by relatively longer periods of quiet. Mandelbrot[21,22] long ago showed that burst errors in communication systems are well modeled by a version of a fractal renewal process, and in particular that the interevent times were essentially independent of each other for time scales determined by the resolution and the duration of the observation.

Furthermore, the rate of traffic flow itself displays fractal fluctuations on a variety of high-speed packet-switching networks conducting different types of traffic.[37–39] This has been demonstrated for time scales greater than about one second in both the power spectral density and the Fano factor. Over these time scales, the FSNDP has been successfully employed to model the traffic,[40] as well as a broad range of other phenomena.[3]

3.3 Biological Ion-Channel Openings

Ion channels are embedded in cell membranes, permitting ions to diffuse in or out.[41] These channels are usually specific to a particular ion, or group of related ions, and block the passage of other kinds of ions. Further, most channels have gates, and thus the channels may be either open or closed. In many instances, intermediate conduction states are not observed. Some ion channels may be modeled by a two-state Markov process,[42] with one state representing the open channel, and the other representing the closed channel. This model generates exponentially distributed dwell times in both states, which are, in fact, sometimes observed. However, many ion channels exhibit power-law distributed closed times between open times of negligible duration,[43,44] and are well described by a standard FRP.[19] Count moments of all orders, coincidence rates, and power spectral densities then all vary as power laws, indicating fractal behavior. In other cases, both open and closed times have distributions following power-law forms, so that one such channel has a conduction pattern well modeled by the alternating FRP.[19] Superpositions of independent versions of these elemental processes give rise to fractal binomial noise (see Fig. 3), which can in turn be approximated by fractal Gaussian noise. Even for dependent ion channels, in fact, evidence exists that the overall effect will be the same, although with a higher variance than for the independent channel case.[45]

3.4 Auditory-Nerve-Fiber Action Potentials

Many biological neurons transmit information by means of action potentials, which are localized regions of depolarization traveling down the length of an axon. Action potentials on a given axon are brief and largely identical events, so their reception at another neuron (or at a recording electrode) may be well represented by a point process. FSPPs have been shown to describe the action potentials in primary auditory-nerve fibers in a number of species.[4,44,46–48] Over short time scales, nonfractal stochastic point processes prove adequate for representing such nerve spikes, but over long time scales (typically greater than one second) fractal behavior becomes evident. In particular, both the Fano factor and the power spectral density vary in a power-law fashion over these long time scales. Furthermore, estimators of the rate of the process converge more slowly than for nonfractal processes, displaying fluctuations which are self-affine over varying time scales greater than one second.[49] With the inclusion of the refractory effects of nerve fibers, the FGNDP provides an excellent approximation for modeling the behavior of nerve spikes in auditory fibers in several species over all time scales and for a broad variety of statistical measures[4,46–48,50]; only three parameters are required. This process may well arise from superpositions of fractal ion-channel transitions in inner-ear sensory cells (see Sec. 3.3).[19,50]

3.5 Human Heartbeat Times

The sequence of human heartbeats exhibits considerable variability over time and between individuals, both in the short-term patterns of the individual beats (contractions) and in the long-term patterns of the times between beats. To study these long-term effects we focus on the times of maximum contraction, thus forming a point process of heartbeats. A particular FSPP, similar to the FGNDP but with an integrate-and-reset (rather than a Poisson) substrate, has been constructed and shown to successfully describe these events.[51,52] In many respects the heartbeat process resembles the process formed by auditory-nerve action potentials. Over short time scales nonfractal point processes model the pattern of times between contractions; for times longer than roughly ten seconds, only fractal models suffice. Both the power spectral density and the Fano factor exhibit scaling, and estimates of the rate exhibit variances which decay only slowly with the averaging time. Further, parameters of the FSPP used to model the data may have applicability for the diagnosis of various disease states.

4. ESTIMATION OF FRACTAL DIMENSION: THEORY

By their nature, fractal processes display fluctuations over a broad range of time scales, including long ones, so that estimating the properties of a segment of data presents more difficulty than estimating those of a nonfractal process. For example, the estimate of the mean rate of a fractal process has a variance which decreases only slowly with the length of the data segment.[4,47] For a nonfractal stochastic point process, such as the HPP, the variance decreases as T^{-1}, where T is the length of the segment; for an FSPP, such as the FGNDP, the variance decreases more slowly, as T^{D-1}, where D is the fractal dimension of the point process.[4,49] This slow convergence derives from the fractal nature of the process, implying long-range correlations which do not average out as quickly as independent fluctuations. As a result of these long-range variations, detecting a change in a property of a fractal process often requires a prohibitively large quantity of data, and in fact often cannot be practically achieved. Therefore, in this paper we only consider the case in which the data segment under study derives from a stationary process, whose properties therefore do not change over time.

Given a segment of an FSPP, we wish to estimate the fractal dimension, D, of the entire process from which the segment was derived. Several effects contribute to estimation error for finite-length segments, regardless of the method used. First, for an FSPP with a relatively large fractal dimension, the energy is concentrated more in longer time scale fluctuations than for an FSPP with a smaller fractal dimension. Any data segment obtained from this first FSPP, however, has a non-zero probability of lacking many of these fluctuations and therefore appearing to be from a process with a smaller fractal dimension.[53] This phenomenon provides an appreciable bias for shorter data segments, and often a surprisingly long set of data is required before reliable estimates of the fractal dimension may be obtained. Similarly, in other cases a data segment will contain a surplus of fluctuations, leading to an inflated estimate of the fractal dimension. Second, finite data lengths introduce windowing errors, as will be considered more fully in the next two subsections. It is generally appreciated that cutoffs in the time domain give rise to oscillations in the frequency domain, confounding the pure power-law behavior of the fractal PSD and introducing bias; a similar effect occurs with the FF. Third, cutoffs in the definitions of the FSPPs themselves result in the fractal-dimension estimates being biased towards the median value of 1/2, and away from the extreme values of zero and unity in some cases, as has been shown for continuous processes.[54] Explicit or implicit cutoffs occur in any practical FSPP, since otherwise the process would be unnormalizable. For example, integration of the form for the power spectral density in Eq. (4) over all time scales would lead to infinite power, while a fractal renewal process without cutoffs would exhibit a survivor function [Eq. (8)] with infinite probability. Finally, the physical limitations of the equipment employed in measuring a process impose practical limits on the range of time scales available. Although algorithms exist for accurately compensating for these cutoffs, this presupposes a detailed knowledge of the process *a priori*, which is not in the spirit of estimating a completely unknown signal. Consequently, we do not attempt to compensate for these cutoffs in this manner.

Many statistical measures may be applied to a point process, but some, such as the interevent-time histogram (closely related to the survivor function), often only provide information over short time scales,[55] and are therefore not useful

in determining the character of long-term fluctuations except in the special case of the fractal renewal point process. We consider in more detail the power spectral density (PSD) and the Fano factor (FF) as a function of the counting time, because they elucidate the clustering in FSPPs particularly well. The coincidence rate theoretically also exhibits fractal scaling, but in practice noise obscures this scaling behavior much more than for the FF and PSD, as mentioned previously. (See Sec. 1.2.)

4.1 Power Spectral Density

For a fractal process, the PSD decreases as an inverse power-law function of frequency, with the exponent equal to the fractal dimension. Often the PSD of an FSPP will approach an asymptotic value at high frequencies, and assume the form of Eq. (4):

$$S(\omega) = S_0[1 + (\omega/\omega_0)^{-D}], \qquad (11)$$

where ω_0 represents the white-noise cutoff frequency. The PSD for a representative recording of auditory-nerve action potentials is provided in Fig. 4. [Throughout the text of this paper we employ radian frequency ω (radians per time unit) to simplify the analysis, while figures are plotted in common frequency $f = \omega/2\pi$ (cycles per time unit) in accordance with common usage.]

To estimate the fractal dimension D, we begin by partitioning the data segment into N adjacent windows of equal length, and counting the number of points that falls in each window, resulting in a sequence of counts $\{Z_n\}_{n=0}^{N-1}$. For least error, we choose N large enough so that the bias introduced by the finite data is small; and small enough so that the counting windows are large compared to the average time between points, reducing the variance of the estimate. Then by the Central Limit Theorem the number of counts in each window approximates a Gaussian distribution. We then compute the discrete Fourier transform of this sequence, yielding the (unsmoothed) estimate of the power spectral density $\hat{S}(k)$. Each of the values of $\hat{S}(k)$ provides information independent of the other values (for large N), and has an error which is also independent of the other errors.[56]

A typical estimate will have the form

$$\hat{S}(k) = S_0[1 + (k/k_0)^{-D}] \exp[n(k)] \qquad (12)$$

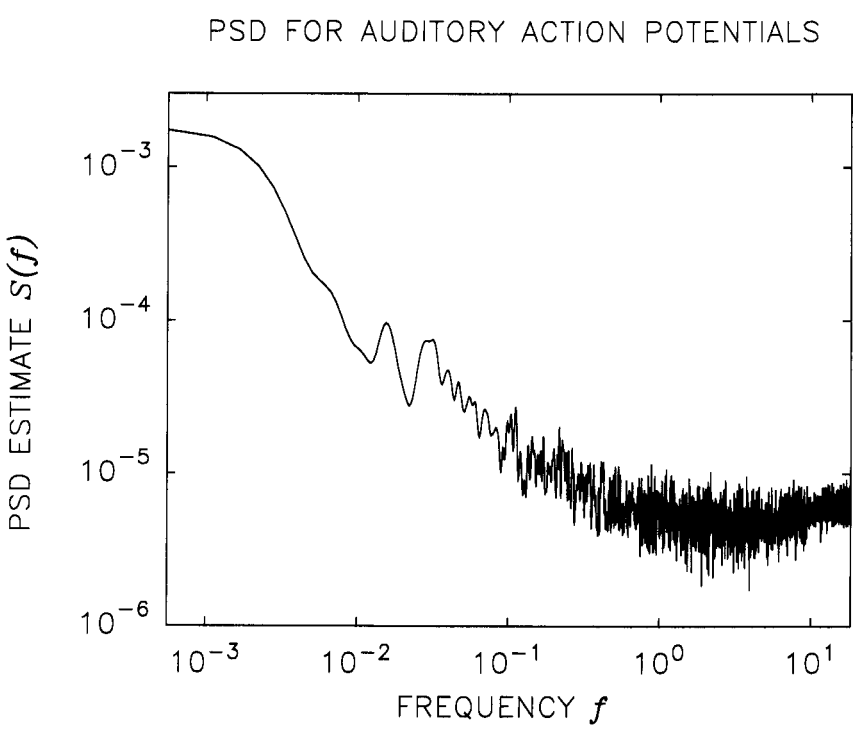

Fig. 4 Doubly logarithmic plot of power spectral density estimate of the point process recorded from a primary cat auditory-nerve fiber under sinusoidal stimulation. The data segment has a duration of 1800 s, with an average time between events of 14.11 ms. The estimate was smoothed with a triangular window in the time domain (autocorrelation) of length equal to 1/8 the duration of the data segment. Over long time scales (low frequencies) the curve follows a straight line of slope −0.72 indicating fractal behavior. The Fano factor for this same data set is displayed in Fig. 5. After Ref. 57.

for $k > 0$, where D is the dimension of the FSPP, and k_0 represents the white-noise cutoff frequency in the integer-frequency space indexed by k. To obtain the best estimate of D, we further require that the number of windows N be chosen to be greater than k_0. The random variable $n(k)$ represents the error in the kth value in the power spectral density estimate, and does not depend on $\hat{S}(k)$.[56] We then have

$$\begin{aligned} \mathrm{E}\{\exp[n(k)]\} &= 1 \\ \mathrm{E}\{[\exp[n(k)] - 1][\exp[n(l)] - 1]\} &= \delta_{kl}\,, \end{aligned} \tag{13}$$

where δ_{kl} is the discrete (Kronecker) delta function. Given these two relations, we expect an approximate value of unity for the variance of $n(k)$. The exact value of the variance depends on the probability density function of $n(k)$, but only weakly, and does indeed lie near unity for three representative two-parameter densities, the results of which are presented in Table 1. To estimate D, we perform a least-squares fit on the logarithm of the power spectral density estimate (less the asymptotic value for high frequencies $\hat{S}_0$) vs. the logarithm of the index k, using the first k_0 values. For $k \geq k_0$ there is essentially no additional information, so we ignore these values. Since each value in $\hat{S}(k)$ for $1 \leq k < k_0$ provides essentially the same amount of information, and all are independent, there is no reason to weigh any one of them more than the others in estimating D. Define

$$\begin{aligned} x_k &\equiv \ln(k)\,, \\ y_k &\equiv \ln[\hat{S}(k) - \hat{S}_0] \\ &= \ln(S_0) + D\,\ln(k_0) - D\,\ln(k) + n(k)\,. \end{aligned} \tag{14}$$

Then the estimate of D is simply the covariance of $\{x\}$ and $\{y\}$ divided by the variance of $\{x\}$:[57]

$$\begin{aligned} \hat{D} &= -\frac{(k_0-1)^{-1}\sum_{k=1}^{k_0} x_k y_k - k_0^{-1}(k_0-1)^{-1}\left(\sum_{k=1}^{k_0} x_k\right)\left(\sum_{l=1}^{k_0} y_l\right)}{(k_0-1)^{-1}\sum_{k=1}^{k_0} x_k^2 - k_0^{-1}(k_0-1)^{-1}\left(\sum_{k=1}^{k_0} x_k\right)^2} \\ &= D + \frac{k_0^{-2}\sum_{k=1}^{k_0}[n(k)-\bar{n}]\sum_{l=1}^{k_0}\ln(l) - k_0^{-1}\sum_{k=1}^{k_0}[n(k)-\bar{n}]\ln(k)}{f(k_0)}\,, \end{aligned} \tag{15}$$

where

$$f(l) \equiv \frac{1}{l}\sum_{k=1}^{l}\ln^2(k) - \left[\frac{1}{l}\sum_{k=1}^{l}\ln(k)\right]^2 \tag{16}$$

is the variance of $\ln(l)$. The function $f(l)$ rises monotonically to an asymptotic value of unity as $l \to \infty$.

The variance of $\hat{D}$ is given by[57]

$$\begin{aligned} \mathrm{Var}[\hat{D}] &= \mathrm{E}\left\{\left[\frac{1}{k_0^2}\sum_{k=1}^{k_0}[n(k)-\bar{n}]\sum_{l=1}^{k_0}\ln(l) - \frac{1}{k_0}\sum_{k=1}^{k_0}[n(k)-\bar{n}]\ln(k)\right]^2\right\}\frac{1}{f^2(k_0)} \\ &= \frac{\mathrm{Var}[n(k)]}{k_0 f(k_0)}\,. \end{aligned} \tag{17}$$

As the length L of the data segment increases, the effective white-noise cutoff frequency k_0 will increase proportionately, and the variance of the estimate $\hat{D}$ will decrease as L^{-1}. This contrasts with the variance of the estimate of the mean, which decays as L^{D-1} for an FSPP.[4,49]

The second term on the rhs of Eq. (15) contains sums of zero-mean random variables, so that $E\{\hat{D}\} = D$, and $\hat{D}$ appears to be an unbiased estimator.[57] However, this calculation for the statistics of the power-spectral-density-based dimension estimate does not explicitly account for the finite length L, of a practical data set. Consider a fully fractal stochastic point process with a CR which

Table 1 Variance of $n(k)$ for three representative two-parameter densities, where a and b are the two parameters, adjusted so that Eq. 13 holds.

Probability Density Function Name	Formula	Variance of $n(k)$
Noncentral Laplacian	$(2a)^{-1}\exp[-\lvert n-b\rvert/a]$	$\sqrt{40}-6=0.325$
Gaussian	$(2\pi)^{-1/2}a^{-1}\exp[(n-b)^2/2a^2]$	$\ln(2)=0.693$
Uniform	$(b-a)^{-1}$, with $a<n<b$	1.222

assumes the form

$$G(\tau)=S_0\delta(\tau)+G_0[1+(\tau/\tau_0)^{D-1}], \tag{18}$$

similar to that of Eq. (4). To illustrate the effects of finite data lengths, we use a rectangular window

$$W(t)=\begin{cases}1 & \text{for } 0\le t<L,\\ 0 & \text{otherwise}.\end{cases} \tag{19}$$

The estimated CR $G_W(\tau)$ will be the product of the true CR and the window convolved with itself

$$G_W(\tau)=S_0\delta(\tau)+G_0(1-|\tau|/L)[1+(\tau/\tau_0)^{D-1}] \tag{20}$$

in the range $|\tau|<L$, and zero otherwise. We obtain the estimated PSD by Fourier transform:

$$\begin{aligned}S_W(\omega)&=\int_{-\infty}^{\infty}G_W(\tau)\exp(-j\omega\tau)d\tau\\&=S_0+2G_0L[1-\cos(\omega L)]/(\omega L)^2+2G_0\tau_0^{1-D}\omega^{-D}\mathrm{Re}\{j^{-D}\Gamma(D,j\omega L)\}\\&\quad-2G_0\tau_0^{1-D}\omega^{-D}\int_0^{\omega L}x^D\cos(x)dx/\omega L\\&\approx 2G_0\tau_0^{1-D}\omega^{-D}[\cos(\pi D/2)\Gamma(D)-(\omega L)^{D-1}\sin(\omega L)]\end{aligned} \tag{21}$$

so that

$$\begin{aligned}&\ln[S_W(\omega)]\\&\quad\approx\ln[2\cos(\pi D/2)\Gamma(D)G_0\tau_0^{1-D}]-D\ln(\omega)\\&\qquad-[\cos(\pi D/2)\Gamma(D)]^{-1}(\omega L)^{D-1}\sin(\omega L)\end{aligned} \tag{22}$$

where Re{} refers to the real part of the argument, $\Gamma(,)$ is the incomplete gamma function

$$\Gamma(a,x)\equiv\int_x^{\infty}t^{a-1}e^{-t}dt, \tag{23}$$

and we employ approximations valid for $1/L\ll\omega\ll\omega_0$.

To estimate the fractal dimension, a straight line is fitted to a plot of the logarithm of the PSD in Eq. (22) vs. the logarithm of the frequency; the slope becomes this estimate. The last term will contribute a bias, and for a least-squares fit the Appendix provides a method for estimating this bias. Equation (62) returns the estimate

$$\begin{aligned}\hat{D}=D-6[\cos(\pi D/2)\Gamma(D)]^{-1}(\omega_L L)^{D-2}\\\times\cos(\omega_L L)\ln^{-2}(\omega_U L),\end{aligned} \tag{24}$$

where ω_L and ω_U are the lower and upper limits of the range of ω used to estimate D. As expected, the bias vanishes as the data length L approaches infinity.

For a fixed L, a *larger* value of ω_L (and therefore a smaller range of ω over which the estimate is computed) yields a smaller bias, concurrent with an increase in the variance. As is often the case, this estimation problem exhibits a bias-variance tradeoff, and some intermediate value of ω_L will yield the best overall performance. However, if

$\omega_L L = (n + 1/2)\pi$, where n is any positive integer, then the bias [as shown in Eq. (24)] vanishes for finite L. For this frequency range the smaller terms in Eq. (22), which are otherwise unimportant and were therefore neglected, will dominate, and the actual bias will be small but non-zero. Other windows will yield other biases, but all will decrease with increasing L. For example, an exponential window

$$W(t) = \exp(-|t|/L) \tag{25}$$

yields the biased estimate

$$\hat{D} = D - 6D \tan(\pi D/2)(\omega_L L)^{-1} \ln^{-3}(\omega_U L) \times \ln(\omega_U / e^2 \omega_L^2 L) \tag{26}$$

which also decreases with increasing L, although more slowly than for the rectangular window.

4.2 Fano Factor

For an FSPP, the FF has the functional form of Eq. (4):

$$F(T) = 1 + (T/T_0)^D , \tag{27}$$

where D is again the dimension of the FSPP, and T_0 is an intercept time. Thus another estimate of D may be obtained by performing linear regression on $\ln[F(T) - 1]$ vs. $\ln(T)$, as in Eq. (15). An FF constructed from the same recording of auditory-nerve action potentials as in Fig. 4 is provided in Fig. 5.

The properties of this estimator are considerably more difficult to derive. Since the estimate of the mean will have the same proportional error for all counting times T, this will have no effect on the estimate of D and can be ignored. However, the estimation of the variance, particularly for long counting times, will suffer from the same slow convergence problems as the determination of the mean. Heuristically, estimates of the power-law slope are restricted to counting times less than one tenth of the length of the data segment.

The FF-based estimate of D also differs from that of the PSD in having a maximum possible value of unity, while the PSD may have any slope. This may be seen by considering a data segment divided into $2N$ intervals of length T, with z_m representing the number of events falling within the mth interval. The data segment may also be divided into N intervals of length $2T$, with Z_n events in the nth interval. Then estimates of the mean values of z and Z are related simply by $\langle Z \rangle = 2\langle z \rangle$ by construction,

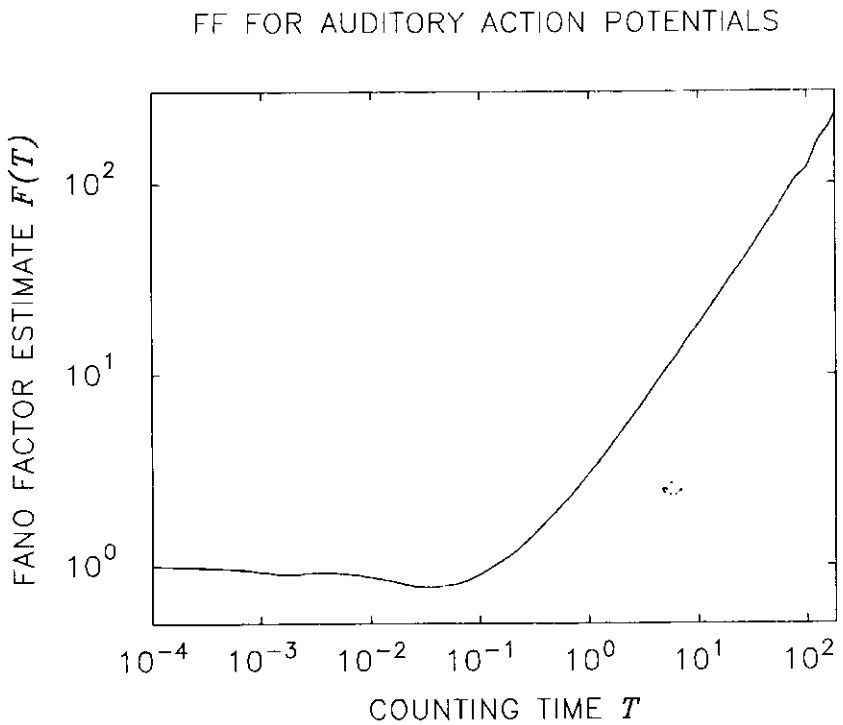

Fig. 5 Fano factor estimate vs. counting time for the same spike train as used for the PSD in Fig. 4. Over long time scales the shape of the curve follows a straight line of slope 0.84 indicating fractal behavior with a similar dimension to that of the PSD. The dip near $T = 30$ ms derives from refractory effects in the neuron. After Ref. 57.

and for the mean square

$$\begin{aligned}\langle Z^2\rangle &= N^{-1}\sum_{n=0}^{N-1} Z_n^2 = N^{-1}\sum_{n=0}^{N-1}(z_{2n}+z_{2n+1})^2\\ &= N^{-1}\sum_{n=0}^{N-1}(2z_{2n}^2+2z_{2n+1}^2-z_{2n}^2\\ &\quad -z_{2n+1}^2+2z_{2n}z_{2n+1})\\ &= 2N^{-1}\sum_{m=0}^{2N-1} z_m^2 - N^{-1}\sum_{m=0}^{N-1}(z_{2m}-z_{2m+1})^2\\ &\le 2N^{-1}\sum_{m=0}^{2N-1} z_m^2 = 4\langle z^2\rangle\,. \end{aligned} \tag{28}$$

For the estimates of the variances of z and Z, and assuming a large value of N, then

$$\mathrm{Var}[Z] = \langle Z^2\rangle - \langle Z\rangle^2 \le 4\langle z^2\rangle - (2\langle z\rangle)^2 = 4\mathrm{Var}[z]\,, \tag{29}$$

and for the estimates of the Fano factor

$$F(2T) = \frac{\mathrm{Var}[Z]}{\langle Z\rangle} \le \frac{4\mathrm{Var}[z]}{2\langle z\rangle} = 2F(T)\,. \tag{30}$$

Therefore, the estimate of D obtained from the FF must be less than unity.[57]

Indeed, in practice the PSD yields an estimate greater than that of the FF.[58] This maximum value of the FF-based estimate effectively skews the estimate towards lower values when D is close to unity. The skew and the slow convergence of the variance estimate yield qualitative information about the statistics of the FF-based estimate, but more quantitative information requires the use of numerical simulations.

The FF and the PSD are uniquely determined by each other through the coincidence rate $G(\tau)$, defined in Eq. (3). The coincidence rate and PSD are Fourier transforms of each other, and the FF may be obtained from the coincidence rate by an integral transform:

$$F(T) = \frac{2}{\lambda T}\int_0^T (T-\tau)[G(\tau)-\lambda^2]d\tau\,, \tag{31}$$

where λ is the average rate of events of the point process. These relationships assure us that the fractal dimensions obtained from the PSD and the FF must indeed be the same number,[4] and they provide a relationship between the intercept time T_0 and the white-noise cutoff frequency ω_0 [Eq. (5)].

Since the two estimators of D employ information over roughly the same time scales, and derive from the same underlying value, we expect their statistics to be similar.

Indeed, the Fano-factor-based dimension estimate also yields a bias for data of finite length. To see this, divide the data segment of total length L into N intervals of length T each, with Z_m representing the number of events falling within the mth interval. The estimate of the mean of Z becomes

$$\hat{\mathrm{E}}[Z] = N^{-1}\sum_{n=0}^{N-1} Z_n\,, \tag{32}$$

and that of the variance is given by

$$\begin{aligned}\widehat{\mathrm{Var}}[Z] &= (N-1)^{-1}\sum_{n=0}^{N-1}\{Z_n-\hat{\mathrm{E}}[Z]\}^2\\ &= (N-1)^{-1}\sum_{n=0}^{N-1} Z_n^2\\ &\quad - N^{-1}(N-1)^{-1}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} Z_m Z_n\\ &= N^{-1}\sum_{n=0}^{N-1} Z_n^2 - 2N^{-1}(N-1)^{-1}\\ &\quad \times\sum_{m=0}^{N-1}\sum_{n>m}^{N-1} Z_m Z_n\,, \end{aligned} \tag{33}$$

with expected value

$$\begin{aligned}\mathrm{E}\{\widehat{\mathrm{Var}}[Z]\} &= N^{-1}\sum_{n=0}^{N-1}\mathrm{E}[Z_0^2] - 2N^{-1}(N-1)^{-1}\\ &\quad\times\sum_{m=0}^{N-1}\sum_{n>m}^{N-1}\mathrm{E}[Z_0 Z_{n-m}]\\ &= N^{-1}\sum_{n=0}^{N-1}\mathrm{E}[Z_0^2] - 2N^{-1}(N-1)^{-1}\\ &\quad\times\sum_{k=1}^{N}(N-k)\mathrm{E}[Z_0 Z_k]\,, \end{aligned} \tag{34}$$

where we have used the stationarity of the point process. For a general FSPP, the expected values above become[40]

$$\begin{aligned}\mathrm{E}[Z_0 Z_k] &= S_0 T\delta_k + S_0^2 T^2\\ &\quad + 2^{-1}S_0 T_0^{-D}T^{D+1}[(k+1)^{D+1}\\ &\quad - 2k^{D+1} + (k-1)^{D+1}]\,, \end{aligned} \tag{35}$$

where δ_k is the Kronecker delta function. Dividing by the mean produces the Fano factor.

The estimate of the Fano factor is simply the estimate of the variance divided by the estimate of the mean; however, computing the expected value of this estimate presents analytical problems. We instead employ the true mean rather than its estimate, which for two reasons does not appreciably affect the result. First, the error so introduced remains a constant factor for all counting times, and so cancels in power-law slope calculations where logarithms are used. Second, the estimate of the variance exhibits much larger variations than the estimate of the mean, so the fluctuations in the Fano factor estimate are dominated by the estimate of the variance. Since the variations in the estimate of the mean therefore play only a minor role in the resulting Fano factor calculations, we may set them equal to zero and simply use the true mean. Therefore, for the expected value of the Fano factor estimate we obtain

$$\begin{aligned}
\mathrm{E}[\hat{F}(T)] &= \mathrm{E}\{\widehat{\mathrm{Var}}[Z]/\hat{\mathrm{E}}[Z]\} \\
&\approx \mathrm{E}\{\widehat{\mathrm{Var}}[Z]\}/\mathrm{E}[Z] \\
&= 1+(T/T_0)^D \left\{ 1-\sum_{k=1}^{N} \frac{N-k}{N(N-1)} \right. \\
&\quad \left. \times[(k+1)^{D+1}-2k^{D+1}+(k-1)^{D+1}] \right\} \\
&\approx 1+(T/T_0)^D \left\{ 1-\int_{z=1}^{N} \frac{N-x}{N(N-1)} \right. \\
&\quad \left. \times[(x+1)^{D+1}-2x^{D+1}+(x-1)^{D+1}]dx \right\} \\
&\approx 1+(T/T_0)^D[1-N^{D-1}] \\
&= 1+(T/T_0)^D[1-(T/L)^{1-D}] \qquad (36)
\end{aligned}$$

where we have employed the relation $N = L/T$. The method used in the Appendix now provides

$$\hat{D} \approx D-6(1-D)^{-1}(T_U/L)^{1-D} \ln^{-2}(T_L/L)\,, \quad (37)$$

where T_L and T_U are the lower and upper limits of the range of T used to estimate D and $T_0 \ll T_L < T_U \ll L$. Again, the bias vanishes as the data length L approaches infinity, and again for a fixed L, a *smaller* value of T_U (and therefore a smaller range of T over which the estimate is computed) yields a smaller bias and an increased variance. Thus this estimate also exhibits a bias-variance tradeoff.

4.3 Example: Cantor Set

Despite the close relationship between the PSD and the FF, the FF sometimes reveals fractal behavior where the PSD does not. An example is the Cantor set $\mathcal{C}$, which, although deterministic, nevertheless illustrates the utility of the FF. We construct a specific Cantor set from the unit interval by removing the middle third, and then removing the middle thirds of the two remaining segments, and continuing this an infinite number of times. Consider the sequence of sets

$$\begin{aligned}
x_0(t) &= \delta(t) \\
x_1(t) &= [\delta(t)+\delta(t-2/3)]/2 \\
x_2(t) &= [\delta(t)+\delta(t-2/9) \\
&\quad +\delta(t-2/3)+\delta(t-8/9)]/4\,,
\end{aligned} \quad (38)$$

which may be generated by the rule

$$x_n(t) = h_n(t) \star x_{n-1}(t)\,, \quad (39)$$

where $\delta()$ is the Dirac delta function, $\star$ represents the convolution operation, and the filter function $h_n(t)$ is defined by

$$h_n(t) = [\delta(t)+\delta(t-2/3^n)]/2\,. \quad (40)$$

If we define

$$g_n(t) \equiv \begin{cases} 3^n & \text{for } 0 < t < 1/3^n\,, \\ 0 & \text{otherwise}\,, \end{cases} \quad (41)$$

then we may represent the (unit measure) Cantor set as

$$\mathcal{C} = \lim_{n\to\infty} h_1 \star h_2 \star \cdots \star h_n \star g_n\,. \quad (42)$$

In the Fourier domain we have

$$\begin{aligned}
\mathcal{F}\{\mathcal{C}\} &= \lim_{n\to\infty} \mathcal{F}\{g_n\} \prod_{k=1}^{n} \mathcal{F}\{h_k\} \\
&= \lim_{n\to\infty} \exp(-j3^{-n}\Omega/2)\mathrm{sinc}(3^{-n}\Omega/2) \\
&\quad \times \prod_{k=1}^{n} \exp(-j3^{-k}\Omega)\cos(3^{-k}\Omega) \\
|\mathcal{F}\{\mathcal{C}\}|^2 &= \lim_{n\to\infty} \mathrm{sinc}^2(3^{-n}\Omega/2) \prod_{k=1}^{n} \cos^2(3^{-k}\Omega) \\
&= \prod_{n=1}^{\infty} \cos^2(3^{-k}\Omega)\,.
\end{aligned} \quad (43)$$

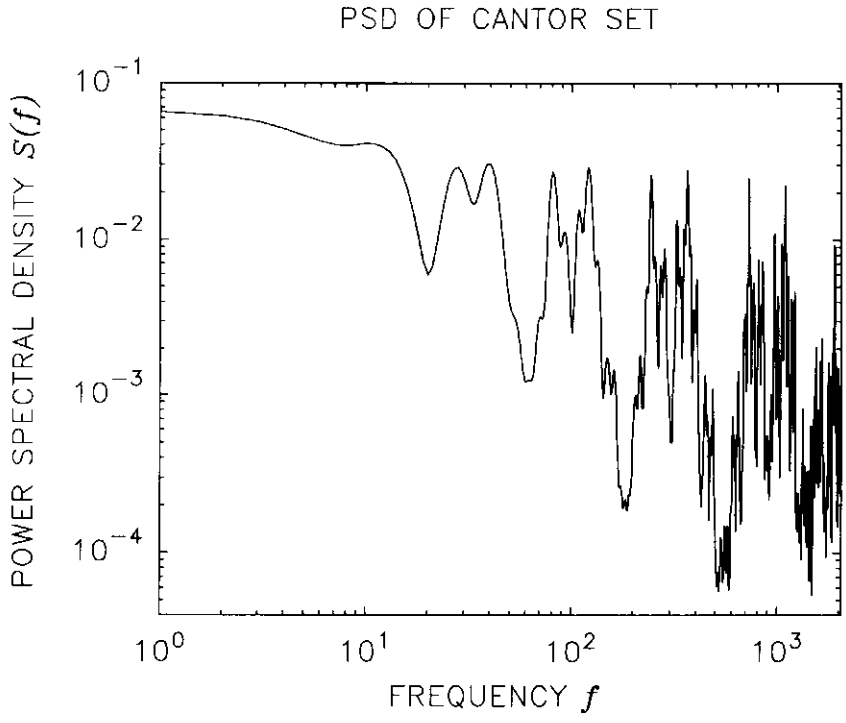

Fig. 6 Doubly logarithmic plot of the power spectral density of the sixteenth-generation approximation to the Cantor set, formed by starting with the unit interval and removing the middle third segments. Although fractal structure is evident in the form of the graph, scaling regions, which would appear as straight lines on this plot, cannot be seen.

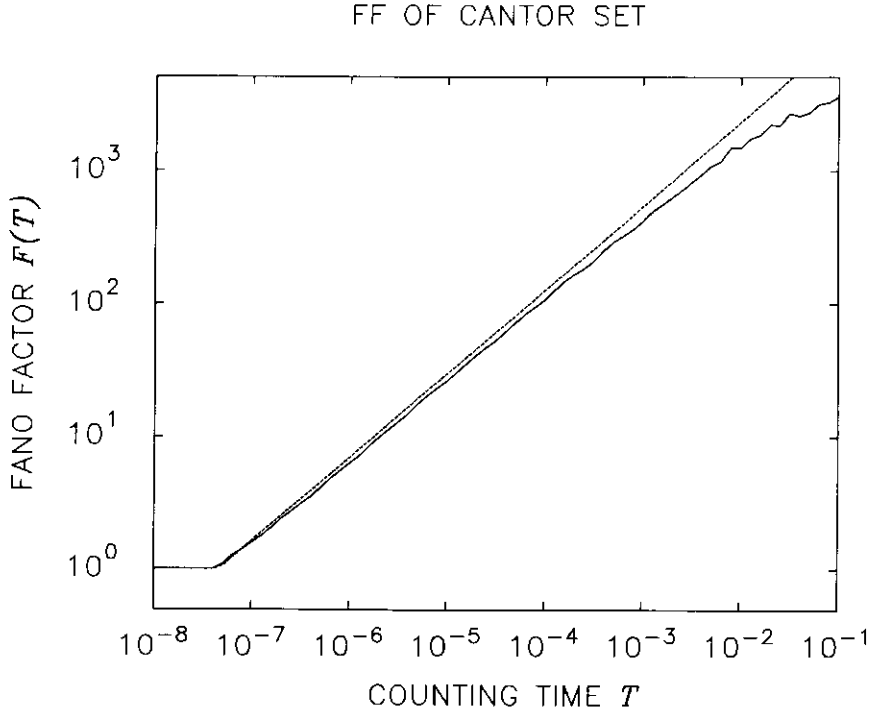

Fig. 7 Doubly logarithmic plot of the Fano factor of the sixteenth-generation approximation to the Cantor set, formed by starting with the unit interval and removing the middle third segments (solid curve). Scaling is evident over many decades, and the asymptotic form $F(T) = (T/T_0)^D$ (dotted line) fits the FF curve well with the dimension $D = \log(2)/\log(3)$, and T_0 set to the smallest interval in the approximation.

This result has also been obtained by somewhat different methods.[59]

Figure 6 shows the PSD of $x_{16}(t)$, which exhibits no discernible scaling region despite $x_{16}(t)$ being self-similar over more than seven decades. In contrast, the FF of the Cantor set does indeed show scaling, and with the correct scaling exponent. Consider the mth stage in the construction of the Cantor set, which consists of 2^m intervals or points. Now divide the original unit interval into 3^n equal segments, with $1 \ll n < m$, each of length $T = 3^{-n}$; of these a proportion $p = 2^n/3^n$ will contain $M = 2^m/2^n$ points each of the Cantor set approximation. By the binomial theorem, the resulting Fano factor will be

$$\begin{aligned} F(T) &= M(1-p) = (2^m/2^n)(1-2^n/3^n) \\ &= 2^m(T^D - T)\,, \end{aligned} \tag{44}$$

where $D = \log(2)/\log(3)$ is the fractal dimension of the Cantor set. Since n is large, then $T \ll T^D$, so that

$$F(T) \approx MT^D = (T/T_0)^D\,, \tag{45}$$

which is indeed scaling with the correct scaling exponent. A plot of the FF of $x_{16}(t)$, displayed in Fig. 7, shows an extensive scaling region. Thus the FF highlights some forms of fractal behavior which the PSD fails to show.

5. ESTIMATION OF FRACTAL DIMENSION: SIMULATION

The analytical results presented in Sec. 4 provide an indication of the expected performance of the PSD- and FF-based dimension estimators. However, for statistics such as the variance of the estimate derived from the FF, and for a picture of the overall behavior of these estimators, we turn to simulations of various mathematical models of FSPPs. We simulated superpositions of standard fractal renewal (point) processes, sums of alternating fractal renewal processes driving a Poisson point process (the FBNDP), and fractal Gaussian noise driving a Poisson point process (the FGNDP), as well as a homogeneous Poisson point process which serves as a check on our theory. We also simulated sums of alternating fractal renewal processes driving an integrate-and-reset process (see Sec. 3.5). In contrast to the sum driving a Poisson point process as illustrated in Fig. 3, the sum was integrated until the result attained a value of unity. A point was then generated in the output point process, the integrator was reset to zero, and the process was repeated over the length of the simulation. This process lacks the additional randomness of the driven Poisson process; since this forms only a relatively small component of the total fluctuation in the process, however, the integrate-and-reset process yielded nearly identical results to those of the FBNDP. We therefore do not discuss it further except to note that it exhibited a slightly more positive bias in the PSD-based dimension estimate for all values of the fractal dimension D compared to the FBNDP.

All simulations were designed with an average time between events of unity and a total length of 10^6, so that the expected number of events was also 10^6, and with fractal intercept times T_0 near 10. For each type of FSPP, 100 simulations were performed for each of three values of the fractal dimension: $D = 0.2$, 0.5, and 0.8.

The following procedures were employed to compute the dimension estimates. For the PSD-based estimate, the absolute event times (not the interevent times) were quantized into 2^{16} bins, which therefore formed a rate estimate of the process with counting window $10^6/2^{16} \approx 15.259$ time units. Then a discrete Fourier transform was performed, followed by replacing the Fourier components with their square magnitudes. Finally, a least-squares fit was obtained on the logarithm of these quantities vs. the logarithm of the frequencies for a selected range of frequencies. As expected, for all three FSPPs simulated, increasing the range of frequencies included in the least-squares fit reduced the variance of the fractal-dimension estimate at the expense of the bias; a range of $10^{-6} \le f = \omega/2\pi \le 10^{-3}$ provided the best overall tradeoff and was selected for all three FSPP simulations, although a lower limit of 10^{-5} yielded similar results. A total simulation length of 10^6 produces a frequency resolution of 10^{-6}, for a total of 10^3 values of the PSD. Equation (17) predicts a standard deviation of 0.027 for this case, while the negative bias due to a finite data length as predicted by Eq. (24) is presented in Table 2 for the values of D employed. The derivation of the bias requires the assumption $\omega L \gg 1$, but over the range of ω used this quantity falls as low as $\omega_L L = 2\pi$. This value appears to exceed unity by a sufficient quantity in practice however, since employing $\omega_L L = 20\pi$ resulted in a nearly identical bias calculated from the simulation.

Table 2 Expected fractal-dimension estimator bias due to finite data size. The PSD-based estimate exhibits a smaller magnitude of bias than that of the Fano factor.

	$D = 0.2$	$D = 0.5$	$D = 0.8$
PSD:	−0.001	−0.004	−0.024
FF:	−0.006	−0.020	−0.099

Note: **FF** Fano factor
PSD power spectral density

For the FF-based estimate, the variance and mean were calculated for a number of counting times, with the times determined by geometric series of ten counting times per decade. As with the PSD-based estimate, a larger range of counting times yielded increased bias and decreased variance. The best overall performance was achieved for a range of $10^0 \leq T \leq 10^5$. Equation (37) predicts a negative bias for the FF-based estimate, also presented in Table 2. The simulation results are displayed in Figs. 8–11 and Tables 3–5.

5.1 Superposition of Standard Fractal Renewal Point Processes

As indicated in Sec. 2.2, perhaps the simplest FSPP is a single standard fractal renewal point process, with an interevent-time survivor function following the abrupt-cutoff power-law form of Eq. (8). Employing an exponent $d \equiv 2 - D$, which falls in the range $1 < d < 2$ for fractal dimensions D between zero and unity, results in fractal behavior in the PSD, FF, and CR with the same fractal dimension D, but provides a more quickly decaying tail in the survivor function which renders the outer cutoff B unnecessary. Thus it may be eliminated and still result in an interevent time with finite mean. Eliminating the outer cutoff yields better power-law behavior in the PSD and FF, and simplifies simulation. A point process in one Euclidean dimension must have a fractal dimension between zero and unity, and employing a power-law exponent d

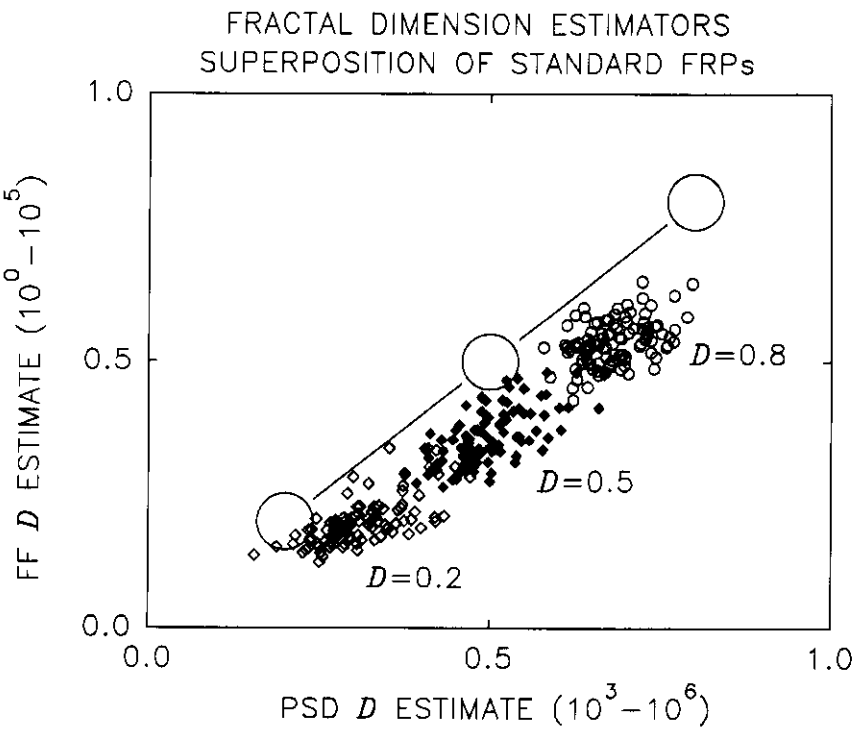

Fig. 8 Scatterplot comparing PSD- and FF-based estimates of the fractal dimension for the superposition of standard fractal renewal processes. Estimates of the fractal dimension computed from a least-squares fit of a doubly logarithmic plot of the power spectral density over the frequency range $10^{-6} \leq f = \omega/2\pi \leq 10^{-3}$ (or time range 10^3–10^6) form the abscissa, while the ordinate consists of estimates from a doubly logarithmic plot of the Fano factor over the time range $10^0 \leq T \leq 10^5$. The large circles mark the points corresponding to zero error in estimating the fractal dimension for the three values of D used to simulate the process: $D = 0.2$, 0.5, and 0.8. The line connecting the circles indicates where the two estimates would coincide.

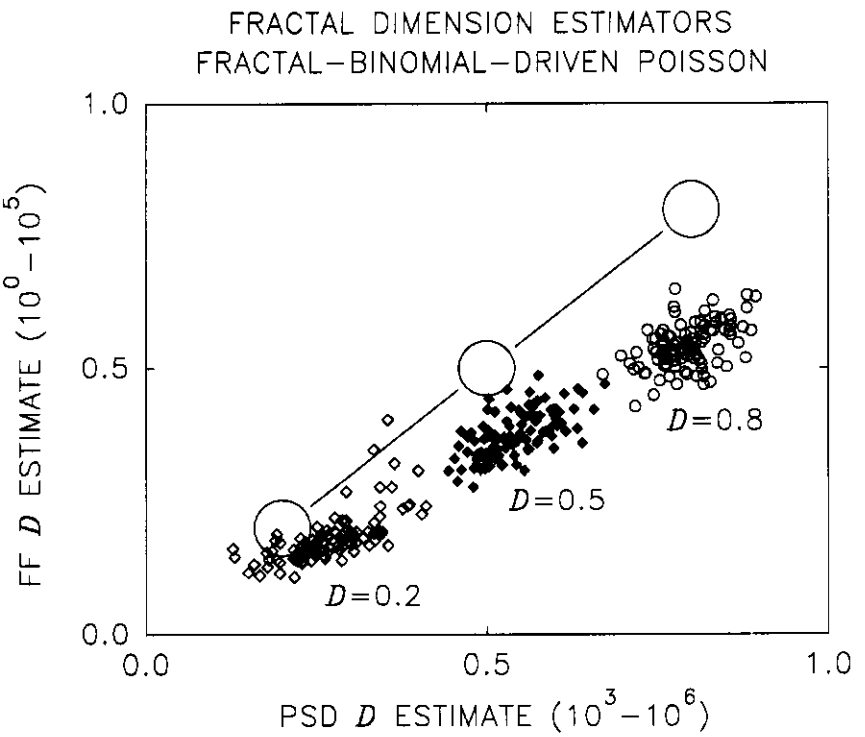

Fig. 9 Scatterplot comparing PSD- and FF-based estimates of the fractal dimension for the sum of alternating fractal renewal processes driving a Poisson point process (FBNDP). Estimates of the fractal dimension computed from a least-squares fit of a doubly logarithmic plot of the power spectral density over the frequency range $10^{-6} \leq f = \omega/2\pi \leq 10^{-3}$ (or time range 10^3–10^6) form the abscissa, while the ordinate consists of estimates from a doubly logarithmic plot of the Fano factor over the time range $10^0 \leq T \leq 10^5$.

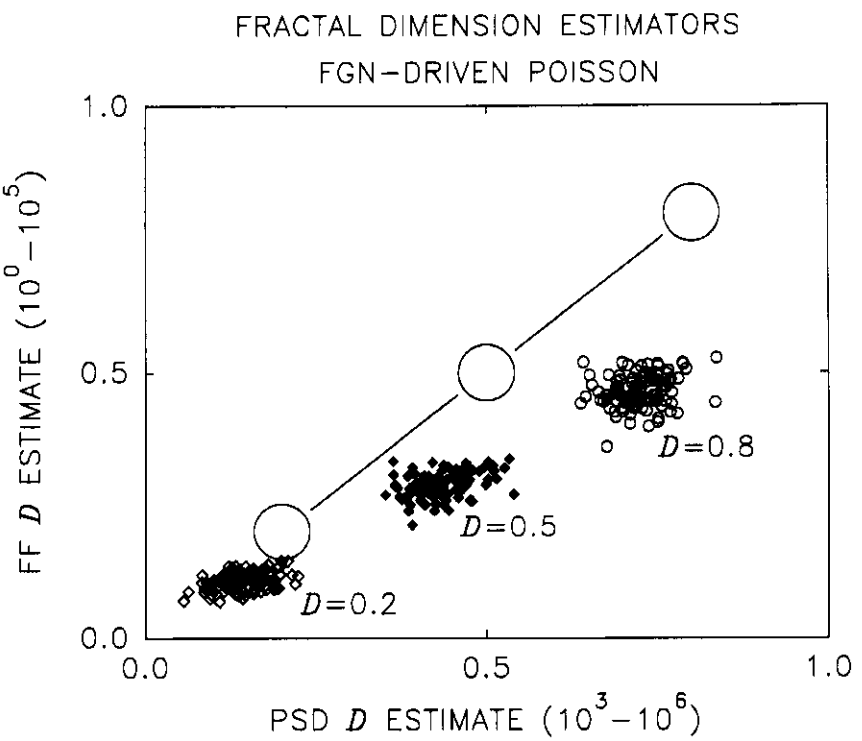

Fig. 10 Scatterplot comparing PSD- and FF-based estimates of the fractal dimension for fractal Gaussian noise driving a Poisson point process. Estimates of the fractal dimension computed from a least-squares fit of a doubly logarithmic plot of the power spectral density over the frequency range $10^{-6} \leq f = \omega/2\pi \leq 10^{-3}$ (or time range 10^3–10^6) form the abscissa, while the ordinate consists of estimates from a doubly logarithmic plot of the Fano factor over the time range $10^0 \leq T \leq 10^5$.

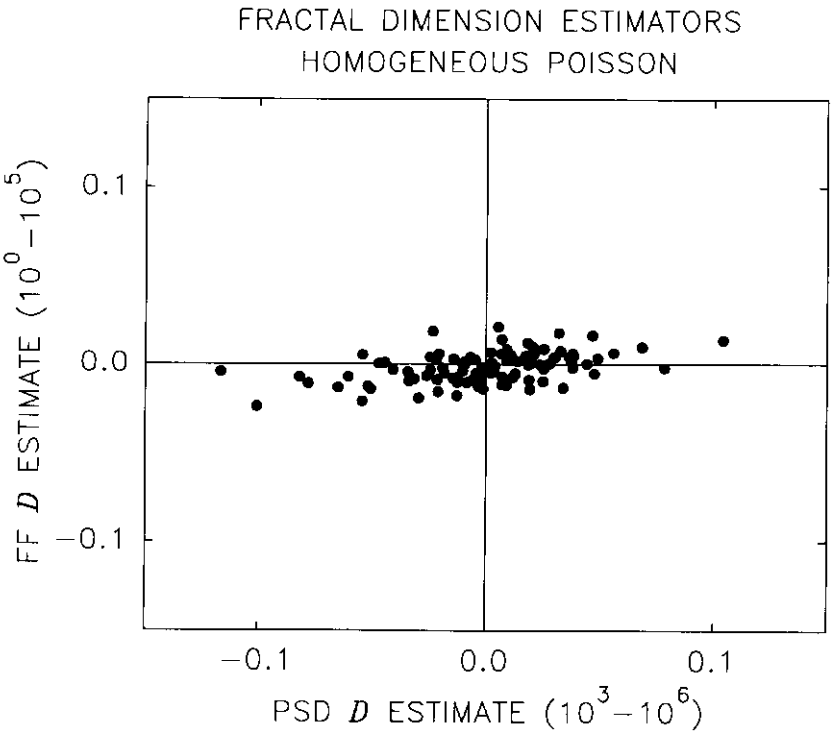

Fig. 11 Scatterplot comparing PSD- and FF-based estimates of the fractal dimension for a homogeneous Poisson point process. Estimates of the fractal dimension computed from a least-squares fit of a doubly logarithmic plot of the power spectral density over the frequency range $10^{-6} \leq f = \omega/2\pi \leq 10^{-3}$ (or time range 10^3–10^6) form the abscissa, while the ordinate consists of estimates from a doubly logarithmic plot of the Fano factor over the time range $10^0 \leq T \leq 10^5$.

Table 3 Superposition of ten standard fractal renewal point processes: Statistics of fractal-dimension estimators for 100 trials. For this FSPP, the PSD-based estimate exhibits bias towards the median value of $D = 0.5$. The correlation coefficient was computed between the PSD- and FF-based estimates of the fractal dimension.

		$D = 0.2$	$D = 0.5$	$D = 0.8$
PSD:	Bias	+0.114	−0.001	−0.119
	Bias (Theory)	+0.044	−0.002	−0.099
	Standard Deviation	0.063	0.057	0.048
FF:	Bias	+0.002	−0.142	−0.262
	Standard Deviation	0.047	0.052	0.043
Correlation Coefficient		0.708	0.633	0.422

Note: **FF** Fano factor
PSD power spectral density

in the range $1 < d < 2$ indeed leads to an FSPP with dimension $D = 2 - d$.[5] The probability density function then assumes the form

$$p(t) = \begin{cases} 0 & \text{for } t \leq A\,, \\ dA^d t^{-(d+1)} & \text{for } t > A\,. \end{cases} \tag{46}$$

However, the resulting Fano factor $F(T)$ has a dip near $T = T_0$, caused by the abrupt cutoff in the interevent time density; furthermore, the power spectral density exhibits excessive oscillations for the same reason. Improvement results from employing

Table 4 Sum of ten alternating fractal renewal point processes driving a Poisson point process: Statistics of fractal-dimension estimators for 100 trials

		$D = 0.2$	$D = 0.5$	$D = 0.8$
PSD:	Bias	+0.070	+0.042	−0.006
	Bias (Theory)	+0.027	−0.009	−0.039
	Standard Deviation	0.062	0.051	0.045
FF:	Bias	−0.018	−0.124	−0.258
	Standard Deviation	0.048	0.044	0.043
Correlation Coefficient		0.692	0.625	0.477

Note: **FF** Fano factor
PSD power spectral density

Table 5 Fractal Gaussian noise driving a Poisson point process: Statistics of fractal-dimension estimators for 100 trials

		$D = 0.2$	$D = 0.5$	$D = 0.8$
PSD:	Bias	−0.056	−0.064	−0.076
	Standard Deviation	0.036	0.041	0.038
FF:	Bias	−0.093	−0.211	−0.335
	Standard Deviation	0.017	0.024	0.032
Correlation Coefficient		0.412	0.341	0.201

Note: **FF** Fano factor
PSD power spectral density

a smoother interevent time density

$$p(t) = \begin{cases} dA^{-1}e^{-dt/A} & \text{for } t \le A, \\ de^{-d}A^d t^{-(d+1)} & \text{for } t > A, \end{cases} \tag{47}$$

which is continuous for all interevent times t.

Other fractal probability density functions were also considered, in an attempt to reduce the effects of the transition at $T = A$, and to extend the region of power-law behavior in the PSD and FF to still higher frequencies and shorter time scales. In particular, the three other analytic forms considered

$$p(t) = \frac{d(d-1)}{(2^d-2)\Gamma(2-d)} A^d (1-e^{-t/A})^2 t^{-(d+1)} \tag{48}$$

$$p(t) = \frac{D}{\Gamma(1-D)}[(A^{-1}+B^{-1})^D - B^{-D}]^{-1} \times (1-e^{-t/A})e^{t/B}t^{-(D+1)} \tag{49}$$

$$p(t) = \frac{1}{\Gamma(d)} A^d e^{-A/t} t^{-(d+1)} \tag{50}$$

resulted in no improvement over the form in Eq. (47). All would have added considerable complexity to the simulation, and were therefore rejected.

We simulated a superposition of ten identical independent standard FRPs, utilizing the smoother interevent-time probability density function in Eq. (47). With an average interevent time of unity for the superposed process, the fractal intercept time T_0 became $T_0 = 3.89$, 8.10, and 9.98 for $D = 0.2$, 0.5, and $D = 0.8$, respectively.

Figure 8 presents the values for the fractal dimension estimated from the 100 simulations of each value of D used, with the estimates derived from the PSD forming the horizontal axis and those from the FF the vertical axis. Table 3 presents the associated bias and standard deviation for both estimates. Also presented are the theoretically expected values of the bias for the PSD-based estimates, which arise from two sources. First, employing the interevent time probability density

function in Eq. (47), exact rather than limiting forms of the PSD were obtained numerically, and used in calculation of an empirical dimension. Second, including the estimated bias due to finite data length from Table 2 formed the expected values of the bias in Table 3. The last line displays the correlation coefficients between the two estimates. Figure 8 displays a bias in the PSD-based estimate towards the median value of 1/2, and in the FF-based estimate towards zero. These effects are presented quantitatively in Table 3, which further reveals an almost linear form for both estimators based on the three values of D simulated. The standard deviations of the estimators appear not to vary between estimators or across values of D.

5.2 Fractal-Binomial-Noise-Driven Poisson Point Process

Again employing the smoother interevent-time probability density function in Eq. (47), we simulated a collection of ten identical independent alternating FRPs and utilized their sum as a stochastic rate function for a Poisson point process (see Sec. 2.3). All of the alternating renewal processes were symmetric, so that the number in either state at any given instant follows the binomial distribution with individual probabilities $p = 1/2$. This is the fractal-binomial-noise-driven-Poisson point process or FBNDP. For ten such processes, the sum already follows the Gaussian amplitude distribution fairly closely, with a skewness of zero and a kurtosis of 2.8; corresponding values for an exactly Gaussian process are zero and 3.0. Thus the sum indeed approximates FGN, and the resulting point process approximates the FGNDP.

Since the rate function may be multiplied by an arbitrary constant without changing the nature of the process, an extra free parameter exists. We chose to eliminate this parameter by setting the cutoff time A equal to the Fano factor intercept time T_0, to within the resolution of the number of alternating fractal renewal processes (which must be an integer). With an average interevent time of unity for the superposed process, the fractal intercept time T_0 became 13, 20, and 15 for $D = 0.2$, 0.5, and $D = 0.8$, respectively.

Figure 9 presents the values from the fractal dimension estimated for the 100 simulations of each value of D used, with the estimates derived from the PSD forming the horizontal axis and those from the FF the vertical axis. Table 4 presents the associated bias and standard deviation for both estimates. Also presented are expected values of the bias for the PSD-based estimates, computed using the same method as in Table 3 but with formulas for alternating FRPs. The last line again displays the correlation coefficients between the two estimates. This process exhibits different behavior than the standard FRP in one respect; the PSD-based estimate appears biased toward higher values of D in Fig. 9, and Table 4 supports this with a near-zero bias for $D = 0.8$. The bias in the FF-based estimate and all standard deviations appear similar to the results for the standard FRP, however.

5.3 Fractal-Gaussian-Noise-Driven Poisson Point Process

We also simulated FGN directly and employed it as a stochastic rate function for a Poisson point process. Exact simulation of a FGNDP by obtaining first passage times for the associated biased fractal Brownian motion appears impractical; rather, a discrete-time approximation of FGN typically serves as the rate function. We employed the relatively simple method of forming a sequence $\{X(k)\}$ with an amplitude decreasing as $k^{-D/2}$ and independent random phases uniformly distributed in the interval $[0, 2\pi)$, and then taking the inverse discrete Fourier transform. The resulting approximation to FGN, $x(n)$, is periodic with period equal to the length of the sequence, however, and does not well represent a segment of a longer, aperiodic process.

This problem was ameliorated by retaining only half of the resulting sequence.[60,61] Specifically, we used an original sequence of length $2^{16} = 65\,536$ points and kept the first $2^{15} = 32\,768$, for a time step between FGN points of $10^6/2^{15} \approx 30.518$ time units. Besides the necessity of approximating the ideal, continuous-time FGN by a discrete-time sequence, the FGN rate function must be prevented from assuming negative values, since a negative rate function has no meaning for a doubly stochastic Poisson point process. In practical simulations, negative values are obviated by choosing parameters such that the probability of negative values occurring is small. In contrast, the nonlinear method of simply setting negative numbers to zero introduces energy at higher frequencies and unacceptably changes the fractal nature of the process. We therefore chose to set $T_0 = 25$ for all simulations of the FGNDP; for this value of the Fano factor

intercept time, the mean rate exceeds the standard deviation by more than a factor of five for all three values of D used, so that the probability of any specific point assuming a negative value becomes $(1/2)\mathrm{erfc}[5/\sqrt{2}] \approx 2.86652 \times 10^{-7}$. With 100 simulations of length 32 768 points for each of 3 values of D, the probability of at least one negative value among independent points would increase to ≈ 0.94. However, the 32 768 points in each simulation derive from a fractal process (FGN); they therefore exhibit a high correlation, and in fact the actual probability is much smaller. Indeed, no negative numbers occurred in any of the simulations, although lower values of T_0, or finer detail (larger number of points) in the discrete-time approximation to the FGN, led to a mean value only 4.5 times the standard deviation, and the consequent appearance of negative values in the FGN.

Since only three parameters completely describe FGN (the mean, the fractal dimension, and the intercept time or frequency), no free parameters exist, as with the standard FRP, but in distinction to the alternating-FRP-driven Poisson point process.

Figure 10 presents the values for the fractal dimension estimated for the 100 simulations of each value of D used, with the estimates derived from the PSD forming the horizontal axis and those from the FF the vertical axis. Table 5 presents the associated bias and standard deviation for both estimates, and the correlation coefficients between the two estimates. This process differs from both FRP-based processes. The standard deviation of the estimates decreases, especially for the FF-based estimate: Table 5 presents smaller values, and Fig. 10 exhibits tighter bunching. These values approach those of the HPP. Further, the bias in all cases is consistently more negative than for the FRP-based processes.

5.4 Homogeneous Poisson Point Process

A nonfractal stochastic point process must have a dimension equal to zero. The homogeneous Poisson point processes (HPP) has a PSD equal to S_0 for all frequencies, a Fano factor $F(T) = 1$ for all counting times, and a CR of G_0 independent of the delay; thus all three fractal dimensions derived analytically from these statistics are indeed identically zero. We therefore simulated a HPP as a test of our theory and estimation procedures.

Figure 11 presents the values for the fractal dimension estimated from 100 simulations of the HPP, with the estimates derived from the PSD forming the horizontal axis and those from the FF the vertical axis. The PSD-based estimate returned a mean of -0.001 with standard deviation 0.035, while for the Fano factor the corresponding values were -0.002 and 0.009. The correlation coefficient between the two estimators was 0.470. As expected, the means for both the PSD- and FF-based estimates are nearly zero, the value for a nonfractal process. Equations (24) and (37) approach values of 0 and -0.003 for the PSD- and FF-based estimator biases, respectively, which closely concur with the simulated HPP data. The standard deviation of the PSD-based estimate also agrees well with the value of 0.027 predicted by Eq. (17). Thus simulation of the HPP verifies the performance of the PSD- and FF-based estimators, showing that a nonfractal process indeed yields estimated dimensions of zero.

5.5 Discussion of Estimation and Simulation Methods

The simulation results presented in Figs. 8–10 and Tables 3–5 lead to a number of conclusions. Perhaps the most important is that even with an average of 10^6 events per simulation, dimension estimation is accurate only to within ± 0.1. Simulations with an average of 10^5 events (not shown) yielded dramatically worse performance, with standard deviations roughly three times larger and with increased bias, as expected. The relative paucity of information available in an experimentally observed point process evidently renders dimension estimation difficult in comparison with continuous-time processes.

The PSD-based estimate proved superior to that obtained from the FF in general. Although the standard deviation from the PSD exceeded that from the FF in all cases, the difference was small, and was dwarfed by the superior bias behavior of the PSD-based estimate, except for $D = 0.2$ in the standard and alternating FRPs. Both estimates exhibited a bias which grew increasingly more negative as D increased in all cases, but much more so for the FF; this probably stems from the FF-based estimate having a maximum possible value of unity, as shown in Eq. (30). That the FF-based estimate yielded a more negative bias accords with the observations of other researchers,[58] as does the tendency of the PSD-based estimate to lie between the

intended value D and 1/2, especially for the standard FRP.[54] Correlation coefficients between the two estimates ranged from +0.2 to +0.7, neither statistically close to unity nor to zero, the latter indicating that the two estimates suffer from similar sources of error; this holds for the HPP as well. That the correlation coefficients do not approach unity shows that the two estimates display somewhat different information, as expected, since the PSD-based estimate uses the time range 10^3–10^6 while that of the FF employs 10^0–10^5. Therefore the FF-based estimate, while often inferior to that of the PSD, presents information not readily available to the PSD, as was shown for the Cantor set in Sec. 4.3. Thus for stochastic signals, as well as deterministic sets, the FF plays an important role in determining the fractal dimension.

One method for reducing the bias of the dimension estimators consists of constructing an extensive table of simulation results, and compensating for the bias accordingly. A particular result from one or both of the estimators would then lead to a distribution of possible values of D which could have generated the process under study. However, the results for the three processes considered here show that different processes, although fractal with the same fractal dimension D over almost the same range of times, yield different values for the bias and standard deviation. Thus such an empirically generated distribution would only prove useful for a particular FSPP, and presupposes some *a priori* knowledge of the process. As with the question of cutoffs considered in Sec. 4, to which this is related, such an approach does not follow in the spirit of estimating a completely unknown signal and we therefore do not attempt to compensate for the bias in this manner.

Another candidate method for improving the performance of the dimension estimators lies in removing the asymptotic value of the measure at high frequencies or short counting times before estimating the dimension. For the PSD, the variance in the PSD estimate precludes subtracting any fixed value; the result would often be negative, and a power-law fit becomes meaningless. The FF exhibits much less variance, especially at the shorter counting times at which the asymptotic value of unity is to be subtracted. However, this becomes problematic for many practical point processes, which have structure at time scales shorter than the fractal onset time T_0; the simple standard FRP with an abrupt cutoff in the interevent-time probability density function was rejected in Sec. 5.1 in favor of one with a smoother cutoff for precisely this reason. Once again, this requires knowledge of the particular details of the process and fails for the general case, so we do not attempt to compensate for short-time effects in the FF. As an example, consider the behavior of the FF of a single standard FRP with interevent time probability density function $p(t)$ given by Eq. (47). Within a time period much smaller than the cutoff time A, the probability of having two events almost vanishes, so that we may approximate the coincidence rate $G(\tau)$ by a constant value of $S_0 p(0)$. Substituting this value into Eq. (31) yields $F(T) \approx 1 + p(0)T$, so that after subtracting unity from the Fano factor, the remainder yields a slope of unity on a doubly-logarithmic plot. Indeed, all three FSPP simulation methods employed exhibited such linear behavior in $F(T) - 1$ for $1 \leq T \leq 3$.

For auditory-nerve recordings in particular, however, power-law behavior in $F(T) - 1$ does indeed exist well below $T = T_0$, and a sophisticated algorithm exists for eliminating confounding non-fractal effects from the data.[62] A smoothed hazard function of the data is estimated, sometimes depending on the previous interval as well, and a FF constructed from it. Then this FF is subtracted from the FF obtained from the data. Although an analytical proof of the validity of this approach does not exist, the method nonetheless works well in practice.

Finally, we note that in addition to the bias and standard deviations of the fractal-dimension estimators attaining large values, they do not closely accord with the predictions obtained in Sec. 4. Values of the standard deviation for the PSD-based estimate range from 0.036 to 0.063, with a mean of 0.049, while the predicted value remains constant at 0.027. The bias for the two FRPs differs from predictions by as much as 0.070. The question naturally arises: what generates these differences? The answer appears to lie in the fractal fluctuations of the point processes themselves, and to be an integral part of fractal-dimension estimation. Fractal behavior in a point process manifests itself in terms of clusters of points on all (or a large set of) time scales; for an FSPP, which is random, the relative abundance of clusters will vary. Thus a particular section of an FSPP with a fractal dimension D may be best described by a different number, while an adjacent segment would result in yet another value of the estimated dimension. This gives an indication of another source of variance in the

estimates. Since dimension estimation is highly nonlinear, the average of a sequence of estimates from sub-sections of a larger section will not in general equal the estimate obtained by examining the larger section as a whole; this systematic difference leads to bias. Longer segments of data, such as 10^7 or even 10^8 points, would presumably exhibit less bias and variance in both estimators, approaching the limit of $\hat{D} = D$ for an infinite-length process. Therefore, without knowing more details about the FSPP under study, for a data segment of 10^6 points the fractal dimension may only be estimated to within ± 0.1. This inherent fluctuation in FSPP highlights the impracticality of detecting a change in the fractal dimension over time in any but the longest data sets. Even in a stationary process, estimators will yield apparently different values for adjacent segments which derive from a single process with a constant fractal dimension, so a putative change in the estimated dimension cannot be attributed with certainty to either a real changing dimension or to fractal fluctuations. Also for this reason, the concept of multifractals has only limited usefulness in general fractal stochastic point processes, since the information required to determine the multifractal spectrum does not appear to exist.

6. CONCLUSION

We have investigated the properties of fractal stochastic point processes (FSPPs). We developed several mathematical formulations of FSPPs, and showed that over a broad range of conditions superpositions of large numbers of arbitrary independent identical FSPPs must converge to the fractal-Gaussian-noise-driven Poisson point process. We presented in detail the analytical properties of the PSD- and FF-based fractal-dimension estimators for FSPPs. We considered the practical aspects involved in simulating FSPPs, and simulated four FSPPs with values of the specified fractal dimension of $D = 0.2$, 0.5 and 0.8. As revealed in both the theoretical properties of the estimators and in the simulated results, we discussed factors which confound the estimation of the fractal dimension: the finite length of the simulation, the structure or type of FSPP employed, and the fluctuations inherent in any FSPP. We conclude that segments of FSPPs with as many as 10^6 points yield estimates of their fractal dimension accurate to only within ± 0.1.

ACKNOWLEDGMENTS

It is a pleasure to thank Bong Kyun Ryu for his close reading of the manuscript and helpful suggestions. This work was supported by the Joint Services Electronics Program through the Columbia Radiation Laboratory and by the Office of Naval Research under Grant No. N00014-92-J-1251.

REFERENCES

1. B. B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, New York, 1983).
2. S. B. Lowen and M. C. Teich, "Power-law shot noise," *IEEE Trans. Info. Theory* **36**, 1302–1318 (1990).
3. S. B. Lowen and M. C. Teich, "Doubly stochastic Poisson point process driven by fractal shot noise," *Phys. Rev.* **A43**, 4192–4215 (1991).
4. M. C. Teich, "Fractal neuronal firing patterns," in *Single Neuron Computation*, eds. T. McKenna, J. Davis, and S. Zornetzer (Academic, Boston, 1992), pp. 589–625.
5. S. B. Lowen and M. C. Teich, "Fractal renewal processes generate $1/f$ noise," *Phys. Rev.* **E47**, 992–1001 (1993).
6. S. B. Lowen, Fractal Stochastic Processes, Ph.D. Thesis (Columbia University, New York City, 1992).
7. W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. (McGraw Hill, New York, 1976), p. 197.
8. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. (McGraw-Hill, New York, 1991).
9. D. R. Cox and P. A. W. Lewis, *The Statistical Analysis of Series of Events* (Methuen, London, 1966), pp. 73–75.
10. H. G. E. Hentschel and I. Procaccia, "The infinite number of generalized dimensions of fractals and strange attractors," *Physica* **D8**, 435–444 (1983).
11. P. Grassberger, "Generalized dimensions of strange attractors," *Phys. Lett.* **A97**, 227–230 (1983).
12. J. Theiler, "Estimating fractal dimension," *J. Opt. Soc. Am.* **A7**, 1055–1073 (1990).
13. F. Grüneis and H.-J. Baiter, "More detailed explication of a number fluctuation model generating $1/f$ pattern," *Physica* **A136**, 432–452 (1986).
14. F. A. Haight, *Handbook of the Poisson Distribution* (Wiley, New York, 1967).
15. D. R. Cox, "Some statistical methods connected with series of events," *J. Roy. Stat. Soc.* **B17**, 129–164 (1955).
16. S. B. Lowen and M. C. Teich, "Generalised $1/f$ shot noise," *Electron. Lett.* **25**, 1072–1074 (1989).
17. S. B. Lowen and M. C. Teich, "Fractal shot noise," *Phys. Rev. Lett.* **63**, 1755–1759 (1989).
18. B. E. A. Saleh and M. C. Teich, "Multiplied-Poisson noise in pulse, particle, and photon detection," *Proc. IEEE* **70**, 229–245 (1982).

19. S. B. Lowen and M. C. Teich, "Fractal renewal processes," *IEEE Trans. Info. Theory* **39**, 1669–1671 (1993).
20. S. B. Lowen and M. C. Teich, "Fractal renewal processes as a model of charge transport in amorphous semiconductors," *Phys. Rev.* **B46**, 1816–1819 (1992).
21. J. M. Berger and B. B. Mandelbrot, "A new model for the clustering of errors on telephone circuits," *IBM J. Res. Dev.* **7**, 224–236 (1963).
22. B. B. Mandelbrot, "Self-similar error clusters in communication systems and the concept of conditional stationarity," *IEEE Trans. Comm. Tech.* **13**, 71–90 (1965).
23. F. Grüneis and T. Musha, "Clustering Poisson process and $1/f$ noise," *Jpn. J. Appl. Phys.* **25**, 1504–1509 (1986).
24. D. R. Cox and V. Isham, *Point Processes* (Chapman and Hall, London, 1980).
25. B. B. Mandelbrot, "A fast fractional Gaussian noise generator," *Water Resources Res.* **7**, 543–553 (1971).
26. T. Lundahl, W. J. Ohley, S. M. Kay, and R. Siffert, "Fractional Brownian motion: a maximum likelihood estimator and its application to image texture," *IEEE Trans. Med. Imag.* **5**, 152–161 (1986).
27. P. Flandrin, "Wavelet analysis and synthesis of fractional Brownian motion," *IEEE Trans. Info. Theory* **38**, 910–917 (1992).
28. M. A. Stoksik, R. G. Lane and D. T. Nguyen, "Accurate synthesis of fractional Brownian motion using wavelets," *Electron. Lett.* **30**, 383–384 (1994).
29. K. Matthes, J. Kerstan and J. Mecke, *Infinitely Divisible Point Processes* (Wiley, Chichester, UK, 1978).
30. G. Pfister and H. Scher, "Dispersive (non-Gaussian) transient transport in disordered solids," *Adv. Phys.* **27**, 747–798 (1978).
31. J. Orenstein, M. A. Kastner and V. Vaninov, "Transient photoconductivity and photo-induced optical absorption in amorphous semiconductors," *Phil. Mag.* **B46**, 23–62 (1982).
32. M. A. Kastner, "The peculiar motion of electrons in amorphous semiconductors," in *Physical Properties of Amorphous Materials*, eds. D. Alser, B. B. Schwartz and M. C. Steele (Plenum, New York, 1985), pp. 381–396.
33. T. Tiedje and A. Rose, "A physical interpretation of dispersive transport in disordered semiconductors," *Solid State Commun.* **37**, 49–52 (1980).
34. V. K. Bhatnagar and K. L. Bhatia, "Frequency dependent electrical transport in bismuth-modified amorphous germanium sulfide semiconductors," *J. Non-cryst. Sol.* **119**, 214–231 (1990).
35. M. F. Shlesinger, "Fractal time and $1/f$ noise in complex systems," *Ann. New York Acad. Sci.* **504**, 214–228 (1987).
36. W. Tomaszewicz, "Multiple-trapping carrier transport due to modulated photogeneration," *Phil. Mag. Lett.* **61**, 237–243 (1990).
37. W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the self-similar nature of Ethernet traffic," in *Proc. ACM SIGCOMM 1993* (1993), pp. 183–193.
38. A. Erramilli and W. Willinger, "Fractal properties in packet traffic measurements," in *Proc. St. Petersburg Reg. ITC Sem.* (St. Petersburg, Russia, 1993), pp. 144–158.
39. B. K. Ryu and H. E. Meadows, "Performance analysis and traffic behavior of Xphone videoconferencing application on an Ethernet," in *Proc. 3rd Int. Conf. Comp. Commun. Netw.*, ed. W. Liu (1994), pp. 321–326.
40. B. K. Ryu and S. B. Lowen, "Modeling self-similar traffic with the fractal-shot-noise-driven Poisson process," Cent. for Telecomm. Res., Tech. Rep. 392-94-39 (Columbia University, New York, 1994).
41. B. Sakmann and E. Neher, *Single-Channel Recording* (Plenum, New York, 1983).
42. L. J. DeFelice and A. Isaac, "Chaotic states in a random world: relationship between the nonlinear differential equations of excitability and the stochastic properties of ion channels," *J. Stat. Phys.* **70**, 339–354 (1993).
43. L. S. Liebovitch and T. I. Tóth, "Using fractals to understand the opening and closing of ion channels," *Ann. Biomed. Eng.* **18**, 177–194 (1990).
44. M. C. Teich, "Fractal character of the auditory neural spike train," *IEEE Trans. Biomed. Eng.* **36**, 150–160 (1989).
45. L. S. Liebovitch and J. Fischbarg, "Membrane pores: a computer simulation of interacting pores analyzed by $g_1(\tau)$ and $g_2(\tau)$ correlation functions," *J. Theor. Biol.* **119**, 287–297 (1986).
46. M. C. Teich, R. G. Turcott and S. B. Lowen, "The fractal doubly stochastic Poisson point process as a model for the cochlear neural spike train," in *The Mechanics and Biophysics of Hearing (Lecture Notes in Biomathematics, Vol. 87)*, eds. P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero and C. R. Steele (Springer-Verlag, New York, 1990), pp. 354–361.
47. M. C. Teich and S. B. Lowen, "Fractal patterns in auditory nerve-spike trains," *IEEE Eng. Med. Biol. Mag.* **13**, 197–202 (1994).
48. A. R. Kumar and D. H. Johnson, "Analyzing and modeling fractal intensity point processes," *J. Acoust. Soc. Am.* **93**, 3365–3373 (1993).
49. M. C. Teich, D. H. Johnson, A. R. Kumar, and R. G. Turcott, "Rate fluctuations and fractional power-law noise recorded from cells in the lower auditory pathway of the cat," *Hear. Res.* **46**, 41–52 (1990).
50. S. B. Lowen and M. C. Teich, "Fractal auditory-nerve firing patterns may derive from fractal switching in sensory hair-cell ion channels," in *Noise in*

Physical Systems and 1/f fluctuations (AIP Conference Proceedings **285**), eds. P. H. Handel and A. L. Chung, (American Institute of Physics, New York, 1993), pp. 781–784.
51. R. G. Turcott and M. C. Teich, "Long-duration correlation and attractor topology of the heartbeat rate differ for healthy patients and those with heart failure," *Proc. SPIE* **2036** (Chaos in Biology and Medicine), 22–39 (1993).
52. R. G. Turcott and M. C. Teich, "Long-duration correlation properties of the ECG recorded from normal subjects and patients with heart failure," *Ann Biomed. Eng.*, submitted.
53. M. Meesman, F. Grüneis, P. Flachenecker and K. Kniffki, "A new method for analysis of heart rate variability: counting statistics of $1/f$ fluctuations," *Biol. Cybern.* **68**, 299–306 (1993).
54. O. I. Yordanov and N. I. Nickolaev, "Self-affinity of time series with finite domain power-law power spectrum," *Phys. Rev.* **E49**, 2517–2520 (1994).
55. S. B. Lowen and M. C. Teich, "Auditory-nerve action potentials from a nonrenewal point process over short as well as long time scales," *J. Acoust. Soc. Am.* **92**, 803–806 (1992).
56. A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, New Jersey, 1975).
57. S. B. Lowen and M. C. Teich, "Estimating the Dimension of a Fractal Point Process," *Proc. SPIE* **2036** (Chaos in Biology and Medicine), 64–76 (1993).
58. F. Grüneis, M. Nakao, M. Yamamoto and T. Musha, "$1/f$ fluctuations in stochastic point processes: New tools for characterizing neuronal outputs," in *Noise in Physical Systems and* $1/f$ *fluctuations*, eds. T. Musha, S. Sato and M. Yamamoto, (IOS Press, Tokyo, 1992), pp. 583–586.
59. C. P. Dettman, N. E. Frankel and T. Taucher, "Structure factor of Cantor sets," *Phys. Rev.* **E49**, 3171–3178 (1994).
60. R. F. Voss, "Fractals in nature: from characterization to simulation," in *The Science of Fractal Images*, eds. H.-O. Peitgen and D. Saupe (Springer-Verlag, New York, 1988), pp. 21–70.
61. D. Saupe, "Algorithms for random fractals," in *The Science of Fractal Images*, eds. H.-O. Peitgen and D. Saupe (Springer-Verlag, New York, 1988), pp. 71–136.
62. O. E. Kelly, Analysis of Long-Range Dependence in Auditory-Nerve Fiber Recordings, Masters Thesis (Rice University, Houston, Texas, 1994).

APPENDIX: MINIMUM-SQUARED-ERROR SLOPE ESTIMATION OF A FUNCTION

Consider a function $f(x)$ for which a linear approximation $y = mx + b$ is desired over the interval $A < x < B$. The total squared error is given by

$$\chi^2 = \int_A^B [f(x) - (mx + b)]^2 \, dx \,. \tag{51}$$

Partial derivatives with respect to both m and b will assume a value of zero where the quantity χ^2 assumes a minimum. We focus on the slope, m. Defining

$$I_n \equiv \int_A^B x^n \, dx \,, \tag{52}$$

the equation

$$I_0 \partial\chi^2/\partial m - I_1 \partial\chi^2/\partial b = 0 \tag{53}$$

yields

$$m = \frac{I_0 \int_A^B x f(x) dx - I_1 \int_A^B f(x) dx}{I_0 I_2 - I_1^2} \,. \tag{54}$$

To find the exponent of a power-law function, we use logarithms

$$\begin{aligned} I_n &= \int_A^B \ln^n(x) dx/x \\ &= (n+1)^{-1} \Big[\ln^{n+1}(B) - \ln^{n+1}(A) \Big] \,, \end{aligned} \tag{55}$$

so that for $B \gg A$ we have

$$I_0 I_2 - I_1^2 \approx \frac{1}{12} \Big[\ln^4(A) + \ln^4(B) \Big] \,, \tag{56}$$

and

$$\begin{aligned} m &= \hat{D} \\ &\approx \frac{12[\ln(B) - \ln(A)] \int_A^B \ln[f(x)] \ln(x) dx/x - 6[\ln^2(B) - \ln^2(A)] \int_A^B \ln[f(x)] dx/x}{\ln^4(B) + \ln^4(A)} \,. \end{aligned} \tag{57}$$

For the PSD-based dimension estimate, the function has the form

$$f(x) = cx^{-D}\left[1 - bx^{-a}\sin(x)\right] \tag{58}$$

where bx^{-a} is small, so that

$$\begin{aligned}\ln[f(x)] &= \ln(c) + D\ln(x) + \ln[1 - bx^{-a}\sin(x)] \\ &\approx \ln(c) + D\ln(x) - bx^{-a}\sin(x)\,.\end{aligned} \tag{59}$$

The two integrals in Eq. (57) then become

$$\begin{aligned}\int_A^B x^{-a}\sin(x)dx/x &\approx \int_A^\infty x^{-(a+1)}\sin(x)dx \\ &\approx A^{-(a+1)}\cos(A)\end{aligned} \tag{60}$$

$$\begin{aligned}&\int_A^B x^{-a}\sin(x)\ln(x)dx/x \\ &\quad\approx \ln(A)\int_A^\infty x^{-(a+1)}\sin(x)dx \\ &\quad\approx A^{-(a+1)}\ln(A)\cos(A)\,.\end{aligned} \tag{61}$$

Since typically $A = \omega_L L$ is closer to unity than $B = \omega_U L \gg 1$, we may disregard the terms in $\ln(A)$, yielding

$$m = \hat{D} \approx D - 6bA^{-(a+1)}\cos(A)\ln^{-2}(B)\,. \tag{62}$$

For the FF-based dimension estimate, however, the function has a simpler form

$$f(x) = cx^D(1 - bx^a) \tag{63}$$

where bx^a is small, so that

$$\begin{aligned}\ln[f(x)] &= \ln(c) + D\ln(x) + \ln(1 - bx^a) \\ &\approx \ln(c) + D\ln(x) - bx^a\,.\end{aligned} \tag{64}$$

The two integrals in Eq. (57) become

$$\int_A^B x^a\,dx/x \approx \int_0^B x^{a-1}\,dx = a^{-1}B^a \tag{65}$$

$$\begin{aligned}\int_A^B x^a\ln(x)dx/x &\approx \int_0^B x^{a-1}\ln(x)dx \\ &= a^{-1}B^a\left[\ln(B) - a^{-1}\right].\end{aligned} \tag{66}$$

For the FF we have the reverse case as for the PSD, in that typically $B = T_U/L$ is closer to unity than $A = T_L/L \ll 1$; this permits us to disregard the terms in $\ln(B)$, so that

$$m = \hat{D} \approx D - 6a^{-1}bB^a\ln^{-2}(A)\,. \tag{67}$$