Fractal-Based Point Processes

2005

Steven Bradley Lowen
Harvard Medical School
McLean Hospital

Malvin Carl Teich
Boston University
Columbia University

WILEY

12

Analysis and Estimation

In the 1950s, **Harold Edwin Hurst (1880–1978)**, a British physicist who spent a great deal of his life in Egypt, developed rescaled range analysis, a statistical measure that revealed long-term dependence and the absence of a characteristic scale in the outflows of the River Nile.

Working with James A. Barnes in the 1960s, the American physicist **David W. Allan (born 1936)** introduced a measure of variability based on differences between successive numbers of events; computation of the Allan variance makes use of Haar wavelets.

Given a segment of a fractal-based point process, it is often desirable to identify the point-process model that predicts the data as closely as possible; this can help to elucidate the mechanisms that underlie the data. In particular, we often wish to estimate the fractal exponent that characterizes the data.

Because of the sparseness of point-process data, however, the identification of a point process is a difficult enterprise. Under many circumstances it cannot be achieved even in principle. Nevertheless, a few situations exist in which we can identify the underlying point process by using a number of statistics in concert. More often, as we will demonstrate in Sec. 12.1 by example, only partial identification can be achieved by analyzing the point process.

Since the estimation problem is far more amenable to solution than is the identification problem, we devote the lion's share of this chapter to investigating procedures for estimating fractal-related parameters associated with an arbitrary fractal-based point process. As discussed in Sec. 12.2.1, we take a nonparametric approach to estimation, making no assumptions about the nature of the point process. We devote particular attention to obtaining estimates for the fractal exponent $\widehat{\alpha}$; the bias, standard deviation, and root-mean-square error of the estimators are of principal interest. We discuss heart rate variability analysis as an example in which careful estimation of the fractal exponent is important.

We obtain our results by making use of simulations of a fractal-Gaussian-process-driven Poisson process, which we introduced in Sec. 6.3.3, using a typical set of parameters. We make use of this process because of its widespread applicability, as discussed in Sec. 12.2.3. The simulations enable us to compare the performance of a large collection of estimators, as detailed in Sec. 12.3. Within the bounds of our study, as discussed in Sec. 12.4, the normalized Haar-wavelet variance $\widehat{A}(T)$ and the

rate spectrum $\widehat{S}_\lambda(f,T)$ emerge as measures of choice for estimating fractal exponents associated with unidentified fractal-based point processes.

For the particular set of parameters examined, we achieve optimal fractal-exponent estimation using the normalized Haar-wavelet variance $\widehat{A}(T)-1$ with five geometrically spaced counting times per decade, weighting of these counting times by $1/\sqrt{T}$, and oversampling of the point process by a factor of two. Optimal estimation using the rate-based spectrum is achieved by with $\widehat{S}_\lambda(f,T)-\widehat{\mathrm{E}}[\mu]$ and calls for the use of uniformly spaced frequency intervals. The performance of the normalized Haar-wavelet variance slightly exceeds that of the rate-based spectrum, albeit at increased computational cost.

Although we expect the estimation results established in this chapter to prove useful in most circumstances, we caution the reader that the *a priori* information available within the realm of all possible experimental scenarios spans far too wide a range for one approach to yield the best results in all cases.

12.1 IDENTIFICATION OF FRACTAL-BASED POINT PROCESSES

We have discussed the identification problem from a simple perspective in Secs. 5.5.4 and 11.5.3, and in a number of Problems sprinkled throughout the text.[1] We will again revisit the point-process identification issue in Sec. 13.6, in connection with computer network traffic.

The identification of the underlying point-process associated with an observed point process is possible only in special cases. More often, the analysis of a point process leads to the identification of only some of its features. Such partial identification is highlighted by the following examples:

- If the estimated rate spectrum (Sec. 3.4.5) of the process, $\widehat{S}_\lambda(f,T)$, strongly indicates the presence fractal behavior, whereas the estimated interval spectrum (Sec. 3.3.3) of the process, $\widehat{S}_\tau(\mathrm{f})$, does not, then the point process under study may well belong to the family of fractal renewal processes, as discussed in Sec. 5.5.4.

  We can confirm this presumption by shuffling the intervals and then recomputing the spectra, which provides a nonparametric test of this hypothesis (see Sec. 11.5). A renewal process, whether fractal or not, is invariant to such shuffling since its interevent intervals are independent and identically distributed. Surrogate data analysis permits us to achieve a good measure of discrimination between fractal and fractal-rate point processes, as discussed in Sec. 11.5.3. It also generally permits the separation of the fractal and nonfractal features of a point process.

[1] See, in particular, Probs. 5.2, 5.3, 5.4, 5.5, 6.8, 7.7, 7.8, 7.9, 7.10, 10.6, 10.7, 10.8, 10.9, 11.3, and 11.12.

The estimated generalized dimension $\widehat{D}_q$ (Sec. 3.5.4) frequently offers a means for definitively discriminating between fractal-rate and fractal point processes (Prob. 5.5).

- In the domain of doubly stochastic Poisson processes (Sec. 4.3), a fractal-binomial-noise-driven Poisson process $dN(t)$ (Sec. 8.4) comprising a small number M of component alternating fractal renewal processes $X(t)$ exhibits a particular interevent interval probability density $p_\tau(t)$. When the underlying fractal binomial noise $X_\Sigma(t)$ (Sec. 8.3.1) assumes a value of zero, no events contribute to $dN(t)$. Otherwise, $dN(t)$ comprises a locally homogeneous Poisson process, with relatively few intervals spanning changes in the rate; this yields an interval density that consists of sums of exponentials.

  The overall interval density $p_\tau(t)$ therefore approximates exponential decay for short times and power-law decay for longer times. This behavior is most evident when the mean rate of $dN(t)$ greatly exceeds the mean switching rates of the constituent alternating renewal processes $X(t)$, and it requires that at least the off times for $X(t)$ have a power-law distribution.

- Similarly, the interevent-interval density of a fractal-shot-noise-driven Poisson process (Sec. 10.3) can sometimes reveal the statistics of the times between the primary events. When the rate of the primary Poisson process is sufficiently small in comparison with the durations of the impulse response functions, significant periods of time elapse between the termination of one impulse response function and the arrival of the next primary event. During this period, no intervals can occur. Therefore, $p_\tau(t)$ essentially follows the form of the primary Poisson process, as provided in Eq. (10.22).

We also note some of the manifold difficulties associated with point-process identification:

- In general, no single statistic is sufficient to identify or characterize a fractal-based point process.

- Even the partial identification of a fractal-based point process generally requires a large quantity of data. Consider the fractal-shot-noise-driven Poisson process discussed above as an example. The interval density $p_\tau(t)$ for $\beta = \frac{1}{2}$ (dotted curve in Fig. 10.6) differs from the homogeneous-Poisson-process density [Eq. (4.3)] only for the largest 10^{-10} of the intervals.

- A general difficulty associated with identifying fractal-rate point processes stems from the fact that the process of generating point events from a rate process destroys information about fluctuations in the rate that occur over time scales shorter than the local interevent interval. Hence, the sampling eradicates detail that is an intrinsic part of the rate process (see Sec. 5.5.4).

  As an example, one can choose parameters for which it is quite easy to distinguish segments of: (1) a fractal Gaussian process (Sec. 6.3.3), (2) fractal binomial noise (Sec. 8.3.1), and (3) fractal shot noise (Chapter 9). However, when

used as rates for a Poisson process, these three diverse rate functions can yield essentially indistinguishable fractal-rate point processes (but see Prob. 12.1).

As another example, given the superposition of a fractal doubly stochastic Poisson process and a homogeneous Poisson process, it is nearly impossible to distinguish the individual component processes.

- The identification of fractal-based point processes is confounded by the large variety of forms that they take, many of which have quite similar statistics. One can easily construct a collection of distinct processes that do not belong to any of the families that we have considered. For example, we can apply the block-shuffling operation (Sec. 11.5.1) to a fractal-shot-noise-driven Poisson process.

The identification problem proves a bit less difficult for the family of integrate-and-reset point processes (Sec. 4.4), by virtue of their deterministic kernel. Nevertheless, just as with the Poisson kernel, all information resident in the rate over time scales significantly shorter than the local interevent time is not carried forward to the ensuing point process.

12.2 FRACTAL PARAMETER ESTIMATION

We turn now to parameter estimation for fractal-based point processes. This is a far easier task than identification.

Well-established techniques are available for estimating various conventional measures of a point process, such as the mean and variance of the interevent intervals, the mean rate of the process, and more complex measures such as the spectrum (see, for example, Cox & Lewis, 1966). We expressly consider the *estimation of fractal parameters* in point processes, a topic that has received scant attention.

12.2.1 Nonparametric estimation

We cast the estimation problem as follows: Given a segment of a fractal-rate point process, we seek an estimate $\widehat{\alpha}$ of the true fractal exponent α of the entire process from which the segment was extracted. We often seek an estimate of the onset time or frequency as well. Several effects contribute to estimation error for finite-length segments, regardless of the methods used (Lowen & Teich, 1995; Thurner et al., 1997).

The fractal exponent provides a measure of the relative strengths of fluctuations over various time or frequency scales; a particular fractal exponent leads to a particular distribution of power over these scales. Variance in the estimated values stems from the inherent randomness of the strengths of fluctuations in finite data sets. A collection of finite realizations of a fractal-based point process with the same parameters will exhibit fractal fluctuations of varying strengths, thereby leading to a distribution of estimated fractal exponents and onset times or frequencies.

Whereas fluctuations lead to variance in estimators, cutoffs lead to bias. Cutoffs arise from limitations in the measurement process: noise and finite precision lead to a minimum practical time scale, whereas the limited duration of any data set defines a maximum time scale (see Sec. 2.3.1). Cutoffs can also derive from the data itself, such as when behavior with different characteristics (nonfractal or even fractal with a different exponent) occurs over adjacent time-scale ranges.

Although algorithms exist that can accurately compensate for the effects of variance and bias, they presuppose detailed *a priori* knowledge of the process, which violates the principle of estimating an unknown signal. We therefore do not attempt to compensate for these effects in this manner, except for very simple cases justified by previous knowledge of the system that generates the data. Indeed, we do not attempt to specifically estimate parameters for a fractal point process; rather we employ the formalism of a fractal-rate point process, which is encountered far more frequently.

In particular, we do not consider maximum-likelihood techniques because they require an exact model, despite the fact that they yield the best possible performance in a certain sense. By definition, applying one model to data generated by another model no longer yields the maximum-likelihood estimate of a parameter; instead it returns an estimate of unknown utility. Furthermore, even when knowledge of a model does permit the use of maximum-likelihood methods, the loss of accuracy entailed in employing a nonparametric estimation technique is small (Veitch & Abry, 1999). We therefore eschew maximum-likelihood and related techniques in favor of a nonparametric approach.

A closely related issue pertains to selecting the range of times or frequencies over which to calculate a fractal measure. Given a monofractal model, a χ^2-approach provides a powerful tool for automatically selecting the appropriate power-law region (Abry et al., 2000). Indeed, selection of the appropriate scaling range proves both important and nontrivial in the general case (see Abry et al., 2000, 2003, for a discussion of this issue). However, since this selection also depends on the decision criterion (χ^2 limit vs. range of dependent variable), we provide estimates of fractal exponents over several different ranges of time or frequency.

As an aside, such *a priori* knowledge also plays an important role in estimating fractal exponents for discrete-time processes. Some estimators have superb characteristics when applied to data generated from a restricted class of models but perform poorly in general, whereas robust estimators yield good (but not superb) results in many cases (Taqqu, Teverovsky & Willinger, 1995). Often a speed/accuracy tradeoff exists as well. Again, detailed knowledge of the underlying model permits the use of an estimation technique tailored to the data at hand.

12.2.2 Example: Fractal exponent of the human heartbeat

The human heartbeat provides an example of the usefulness of estimating the fractal exponent of a point process. Fluctuations in the sequence of heartbeat interevent intervals, over time scales ranging from minutes to hours, can help assess the presence of, and likelihood of acquiring, cardiovascular disease (Malik et al., 1996). This noninvasive approach has come to be called **heart rate variability** (HRV) analysis

(Hon & Lee, 1965), whether attention is directed to fluctuations of the actual heart rate or of the interevent intervals.

A whole host of heart-rate-variability measures have been developed and examined over the years. Although the vast majority of these are nonfractal (scale dependent) in nature, the fractal exponents $\widehat{\alpha}$ associated with a number of measures have become a part of the armamentarium used for the analysis of the heartbeat point process. These measures include the interval-based spectrum $\widehat{S}_\tau(\mathsf{f})$ (Kobayashi & Musha, 1982; Turcott & Teich, 1996), the normalized interval-based Haar-wavelet variance $\widehat{A}_\tau(k)$ (Thurner et al., 1998; Ashkenazy et al., 1998), the rate spectrum $\widehat{S}_\lambda(f,T)$ (Turcott & Teich, 1996), and the normalized Haar-wavelet count variance $\widehat{A}(T)$ (Turcott & Teich, 1993, 1996).

Heart-rate-variability measures that can discriminate patients with congestive heart failure from normal subjects have received particular attention. Teich et al. (2001) recently carried out such a study using 16 measures, of which five were scaling exponents: $\widehat{\alpha}_{A\tau}$, $\widehat{\alpha}_{S\tau}$, $\widehat{\alpha}_A$, $\widehat{\alpha}_Y$, and $\widehat{\alpha}_U$. The results of this investigation revealed that the interval-based Haar-wavelet variance at a scale near 32 heartbeat intervals, $\widehat{A}_\tau(32)$, along with its interval-based spectral counterpart near 1/32 cycles/interval, $\widehat{S}_\tau(1/32)$, are the most reliable of the measures, even for electrocardiogram records just minutes long. However, some evidence suggests that the scaling exponents $\widehat{\alpha}_{A\tau}$ and $\widehat{\alpha}_Y$ outperform scale-dependent measures for predicting *mortality* following myocardial infarction (Ashkenazy et al., 2001). Hence, heart-rate-variability analysis provides an example in which careful estimation of the fractal exponent is an important task.

12.2.3 Simulation and fractal-exponent estimation

To illustrate how the fractal exponent of a point process may be estimated using nonparametric techniques, we turn to a specific, but important, example: the **fractal-Gaussian-process-driven Poisson process** introduced in Sec. 6.3.3. This point process has widespread applicability and therefore provides a good testbed for our analysis.

We generate the fractal Gaussian driving process via a simple spectral method (Peitgen & Saupe, 1988) that enables us to use the fast Fourier transform with an array of size $M = 2^{17} = 131\,072$. After the inverse Fourier transform, we discard half of the array to reduce periodicity effects, leaving $M/2 = 2^{16} = 65\,536$ elements. We use a simulation duration $L_0 = 1.1 \times 10^4$. Each element of the resulting fractal Gaussian process therefore corresponds to a duration of $2L_0/M \doteq 0.167847$, and serves as the rate for a (locally) homogeneous Poisson process for that same duration. Thus, every $2L_0/M$ time units, the rate changes to the value specified by the next element in the array. While this method yields a good approximation to a fractal Gaussian process over long time scales, the piecewise-constant construction essentially eliminates fluctuations over time scales significantly shorter than $2L_0/M$.

As discussed in Secs. 6.1 and 6.2, Gaussian processes in general, and fractal Gaussian processes in particular, can assume negative values. These are not permitted for

point-process rate functions and must be eliminated. Eschewing nonlinear methods for dealing with this issue, we instead choose a mean rate that yields positive elements in all simulations. For the example at hand, we chose the parameters as follows: mean rate $\mathrm{E}[\mu] = 100$, duration $L_0 = 1.1 \times 10^4$, fractal exponent $\alpha = 0.8$, and onset frequency $f_S = 0.2$. The resulting rate has a coefficient of variation $C_\mu \doteq 0.176181$. For $M/2 = 65\,536$ and 100 simulations, the complementary error function provides that all rates lie above zero with a probability greater than 0.95, and indeed this was the case for our simulations.[2] For some of the interval-based measures we examine, it proves simpler to employ data sets for which the number of intervals is an integral power of two. The simulations have an expected number of events equal to $1\,100\,000$; of the 100 simulations generated, the one with the least number of events has $1\,067\,365$. We therefore retain the first $2^{20} = 1\,048\,576$ of the intervals in each of the 100 simulations. A further rationale for the choice of these particular values will emerge as the chapter unfolds. These truncated simulations have total durations L that vary; by construction, none can exceed the value $L_0 = 11\,000$ employed in the original simulations. We measure $\widehat{\mathrm{E}}[L] = 10\,511.2$ (yielding $\widehat{\mathrm{E}}[\tau] = \widehat{\mathrm{E}}[L]/N = 10\,511.2/1\,048\,576 \doteq 0.010024$) and $\widehat{\sigma}_L = 187.0$. Finally, within simulations the variability in the rate slightly favors small and large intervals at the expense of the mean, yielding an average interval coefficient of variation $\widehat{\mathrm{E}}[C_\tau] = 1.0327$. For the homogeneous Poisson process, in contrast, such variability does not occur, and we have $C_\tau = 1$ exactly (see Sec. 4.1).

We begin by examining the estimation of the fractal exponent via the normalized Haar-wavelet variance $\widehat{A}(T)$. As discussed in Sec. 3.4.3, Allan and Barnes introduced the unnormalized version of this measure for discrete-time processes in 1966. The presentation in this section is intended to serve as an example that provides a general format for carrying out the analysis. The performance of other measures set forth in Chapter 3 will follow in subsequent sections of this chapter. We will carry out a more thorough study of fractal-exponent estimation via $\widehat{A}(T)$ in Sec. 12.3.8.

To first order, the estimated normalized Haar-wavelet variance should increase as a power-law function of the counting time, which, according to Eq. (5.2), obeys

$$\widehat{A}(T) \approx (T/T_A)^\alpha. \tag{12.1}$$

A representation that is more suitable in many cases follows the relation

$$\widehat{A}(T) \approx 1 + (T/T_A)^\alpha, \tag{12.2}$$

as provided in Eq. (5.6). We performed the simulations detailed above, and estimated the normalized Haar-wavelet variance $\widehat{A}(T)$ at a sequence of counting times T for each of the 100 runs separately. As with prior displays of the normalized Haar-wavelet variance, we chose these to increase geometrically by factors of $10^{0.1}$, thereby providing 10 counting times per decade of the overall counting-time range considered.

[2] The strong correlations among the rates makes this probability still closer to unity.

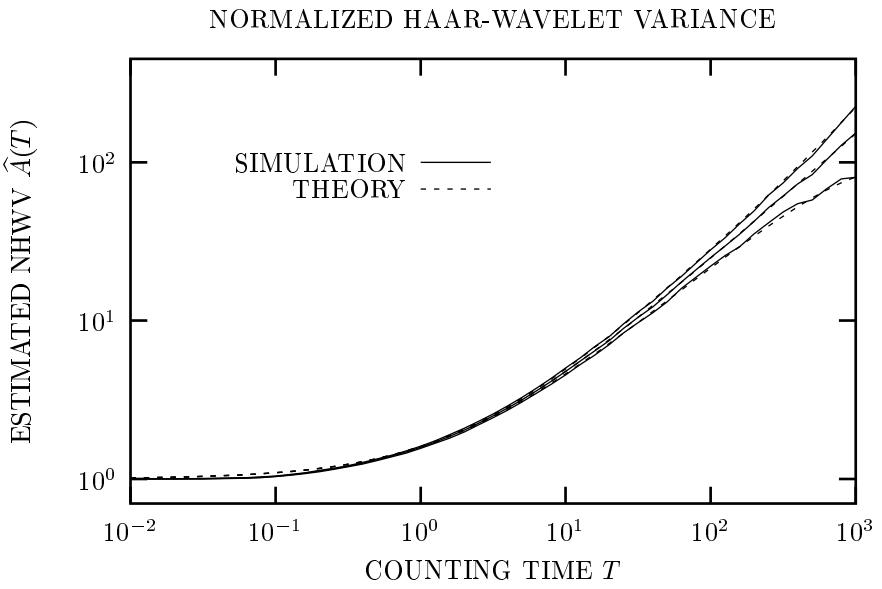

Fig. 12.1 Estimated normalized Haar-wavelet variance $\widehat{A}(T)$ vs. counting time T for a fractal-Gaussian-process-driven Poisson process. Using different random seeds, we simulated 100 runs of this process and estimated the normalized Haar-wavelet variance for each. We present the mean value (center solid curve) and mean $\pm$ one standard deviation (upper and lower solid curves) for these simulations. The corresponding three theoretical (dashed) curves derive from Eqs. (5.44), (5.45), (12.21), and (12.25). Note the increase in the uncertainty as the counting time T increases, and the slight dip in the simulation outcome near $T = 0.1$.

Figure 12.1 presents the corresponding simulation results. The solid curves correspond to values calculated from these individual $\widehat{A}(T)$ estimates: mean (center solid curve) and mean $\pm$ one standard deviation (upper and lower solid curves).

The dashed curves in Fig. 12.1 display theoretical results. Equations (5.44) and (5.45) yield $T_A \doteq 1.8939$; the dashed curve at the center corresponds to the representation provided in Eq. (12.2). Results established in Sec. 12.3.8, our more thorough study of estimation via the normalized Haar-wavelet variance, provide the upper and lower dashed curves, which represent theoretical values for the mean $\pm$ one standard deviation [see Eq. (12.25)].

As with all finite data sets, increasing the counting time T yields fewer members of the sequence of counts, $\{Z_k(T)\}$, which, in turn, leads to increased variability in the estimated normalized Haar-wavelet variance, $\widehat{A}(T)$. This results in a wider separation between the mean $\pm$ one standard deviation curves in Fig. 12.1 for larger values of the counting time T. For this reason, employing values of $\widehat{A}(T)$ near $T = L$ has dubious value; we typically set an upper limit $T \leq T_{\max} \equiv L/10$ for all count-based statistics.

Time Range		$\widehat{\alpha}_A$ from $\widehat{A}(T)$			$\widehat{\alpha}_{A-1}$ from $\widehat{A}(T)-1$		
$T_{\min}$	$T_{\max}$	Bias	SD	RMSE	Bias	SD	RMSE
10^0	10^1	-0.321	0.019	0.321	0.007	0.028	0.029
10^0	10^2	-0.193	0.020	0.194	0.001	0.023	0.023
10^0	10^3	-0.128	0.039	0.134	-0.009	0.039	0.040
10^1	10^2	-0.082	0.052	0.097	0.000	0.056	0.056
10^1	10^3	-0.061	0.077	0.098	-0.018	0.078	0.080
10^2	10^3	-0.060	0.197	0.206	-0.046	0.199	0.204

Table 12.1 Performance of fractal-exponent estimates $\widehat{\alpha}$ for 100 different simulations of a fractal-Gaussian-process-driven Poisson process, obtained via estimates of the normalized Haar-wavelet variance $\widehat{A}(T)$. We chose the counting times to increase geometrically by factors of $10^{0.1}$, thereby providing 10 counting times per decade of the overall counting-time range used to obtain the estimate; the first two columns lists this range. The next three columns represent the bias, standard deviation (SD), and resulting root-mean-square error (RMSE) for calculations carried out on the logarithm of the estimated normalized Haar-wavelet variance $\widehat{A}(T)$; the final three columns display the same calculations carried out on $\widehat{A}(T)-1$. Best results (least root-mean-square error) obtain by using the counting-time range $10^1 \le T \le 10^2$ for $\widehat{A}(T)$, and by using $10^0 \le T \le 10^2$ for $\widehat{A}(T)-1$.

Apart from statistical deviations about the mean, the mean value itself also deviates from ideal fractal behavior, as a result of the method for constructing the fractal Gaussian process used in the simulation. As indicated above, we employ a discrete-time approximation; this leads to a generated process that remains fixed for periods of $2L/M$, or approximately 0.160 time units, bearing in mind that L differs among the simulations. For counting times T significantly less than $2L/M$, the rate rarely changes between counts $Z_k(T)$ in adjacent counting windows. This leads to results that differ little from those of a homogeneous Poisson process, and provides an explanation for why the simulation (solid curve) lies closer to unity than the simple theoretical result (dashed curve) for small counting times T (this slight difference appears in many other measures presented in this chapter as well). Hence, for this process with small counting times, $A(T) \approx 1$ provides a somewhat better approximation than does Eq. (12.2). We can obtain further improvement by making use of an empirical fit that more precisely accommodates the piecewise-constant construction of the fractal Gaussian process, namely

$$A(T) = 1 + \frac{(T/T_A)^{\alpha}}{1 + (T/T_1)^{-2}}\,, \tag{12.3}$$

with $T_1 \doteq 0.124$. Equation (12.3) follows the simulation results almost perfectly (not shown).

The deviations from ideal scaling behavior displayed in graphical form in Fig. 12.1 directly carry over to estimates $\widehat{\alpha}$ of the fractal exponent. We examined the statistics

of the fractal exponents estimated from individual simulations, rather than from the aggregate results shown in Fig. 12.1. Specifically, for each simulation we calculated the logarithms of the estimated normalized Haar-wavelet variance $\widehat{A}(T)$ and of the counting time T, obtained the least-squares best fit of a straight line to this curve over a specified range of counting times T, and called the resulting slope $\widehat{\alpha}$. (Fitting the logarithms of the values to a linear function turns out to be superior to fitting the original values to a power-law function; see Sec. 12.3.8.)

We then determined the statistics of $\widehat{\alpha}$ over the 100 simulations, namely its sample bias

$$\widehat{\mathrm{E}}[\widehat{\alpha} - \alpha],$$

standard deviation (SD)

$$\widehat{\mathrm{Var}}^{1/2}[\widehat{\alpha}\,],$$

and root-mean-square error (RMSE)

$$\widehat{\mathrm{E}}^{1/2}[(\widehat{\alpha} - \alpha)^2],$$

which is equal to the square root of the sum of the squares of the standard deviation and the bias. We followed this procedure for various ranges of the counting time, spanning 10^0 through 10^3 in decade steps. Nonfractal properties of the data set, such as dead time, often heavily influence behavior for counting times substantially smaller than T_A; for this reason, we do not consider $T_{\min} < 1$. Finally, we repeated the entire procedure for the logarithm of $[\widehat{A}(T) - 1]$, rather than of $\widehat{A}(T)$. Table 12.1 presents the results of this procedure.

Examination of the third column in Table 12.1 reveals the strong negative bias that emerges by using $\widehat{A}(T)$ directly. The form of Eq. (12.2) implies that the slope on a doubly logarithmic plot will lie significantly below α, because of the presence of the unity term; this occurs for all ranges of the counting time, but is most apparent for the shorter ones. This observation accords with the curved appearance of Fig. 12.1. Were $\widehat{A}(T)$ to follow a pure power-law form, as in Eq. (12.1), Fig. 12.1 would behave as a straight line. The standard deviation of $\widehat{\alpha}$ (fourth column) increases with counting time, in concert with the increase in the standard deviation of $\widehat{A}(T)$. Taken together, the bias and standard deviation of $\widehat{\alpha}$ result in a root-mean-square error that never lies below 0.09, largely as a result of the bias, which, in turn, arises from the constant unity term.

Subtracting a constant term of unity yields better estimates for the fractal exponent. The magnitude of the bias decreases markedly, although the bias still proves significant for the largest and smallest counting times. For the smallest of these, the departure of $\widehat{A}(T)$ from the simple form of Eq. (12.2), resulting from the discrete-time approximation employed for the fractal Gaussian process, gives rise to the slight dip in the simulation outcome (relative to theory) in Fig. 12.1 near $T = 0.1$. This, in turn, leads to an estimated slope *larger* than α, and therefore an estimated fractal exponent $\widehat{\alpha} > \alpha$.

Turning to the largest counting times, all bias calculations including $T = 10^3$ exhibit negative values. In all three cases the magnitude of the bias exceeds the

standard error (standard deviation divided by the square root of the number of trials) by a factor of two or three, and the bias values appear approximately to follow a Gaussian distribution, making this difference highly significant. This likely arises from periodicity effects in the simulated fractal Gaussian process array; the circular nature of the original array of size M reduces the variance below the value that would obtain for a true (nonperiodic) fractal Gaussian process. While eliminating half of the array reduces this effect, it nevertheless remains statistically significant, particularly for longer times.[3] This deviation also appears at the right-most edge of Fig. 12.1, where the simulated mean $\widehat{A}(T)$ curve lies slightly below its theoretical value.

Intermediate counting times yield excellent bias values, which are quite small for $10^0 \leq T \leq 10^2$; this range of counting times yields the best overall performance as well, with a root-mean-square error of only 0.023. Finally, we note that the bias values, while significantly different from zero, do not significantly affect the root-mean-square error values; they cause an increase of less than four percent of the value that would obtain from the standard deviation alone.

Using the functional form provided in Eq. (12.3) to model the results in Fig. 12.1 would provide still better estimates $\widehat{\alpha}$. However, as indicated in Sec. 12.2.1, this violates the spirit of analyzing an unknown process. Indeed, subtracting unity from the estimated normalized Haar-wavelet variance is truly acceptable only if we know that the process approximately follows the form of Eq. (12.2). In fact, replacing the Poisson kernel with an integrate-and-reset process yields an estimated normalized Haar-wavelet variance that more closely follows Eq. (12.1) for intermediate counting times. Subtracting unity from such values of $\widehat{A}(T)$ would then lead to a *positive* bias in the estimate $\widehat{\alpha}$, and would likely even lead to negative values for $\widehat{A}(T) - 1$ over some range of counting times. Certainly, progressively more complex models for $\widehat{A}(T)$ yield progressively better estimates $\widehat{\alpha}$ when extensive *a priori* knowledge of the process exists [see Bardet, Lang, Moulines & Soulier (2000) and Bardet et al. (2003) for a sophisticated treatment]. However, these models become progressively less tenable in the absence of such knowledge. In this chapter we make use of both Eqs. (12.1) and (12.2) in estimating the fractal exponent.

We conclude this section by noting that other simulation methods can provide results that follow Eq. (12.2) even more closely. In later sections of this chapter, where we evaluate the performance of various estimation approaches, it turns out that the slight difference between simulated and ideal behavior does not affect the performance measures significantly more (or less) than it does for the normalized Haar-wavelet variance examined here. These differences give rise to estimated fractal exponents with expected values that differ from the nominal value of 0.8. This leads to a small apparent bias, but as shown in the right-most three columns of Table 12.1,

[3] A further systematic source of error derives from the fact that the expectation of the logarithm differs from the logarithm of the expectation. This difference, which arises in all exponent estimates based on doubly logarithmic plots, can prove important when statistics become sparse, such as at the longest counting times for $\widehat{A}(T)$. In this particular case, explicit forms are available for bias correction (Veitch & Abry, 1999). However, these forms are valid only for Gaussian-distributed counts, and the calculated values overestimate the bias by a factor of three for our simulations.

we may generally neglect the effects of the bias, and therefore the effects of non-ideal fractal behavior, for the purposes of evaluating various point-process measures.

Furthermore, the results developed in this section illustrate some of the issues involved in analyzing non-ideal point processes. The simulation approach presented here is therefore instructive since most real data sets depart from ideal fractal behavior to some extent.

We now proceed to use the same simulated data to evaluate the performance of various other measures.

12.3 PERFORMANCE OF VARIOUS MEASURES

12.3.1 Limitations of measures not based on counts

Of the collection of measures presented in Chapter 3, some prove more useful than others for estimating the parameters of fractal behavior such as fractal exponents and onset times or frequencies. As suggested in Sec. 5.5.1, we can cast aside interval-based measures from consideration when analyzing a general fractal-based point process. These measures generally prove less useful than count-based and point-process-based measures because they reliably reveal fractal behavior only in special cases.

To establish this, we examine the results of applying interval-based measures to two specific fractal-based point processes, and demonstrate that all such measures fail with one process or the other.

First we consider a fractal renewal process (Chapter 7). As a member of the renewal point-process family, the interevent intervals are independent so that all of the measures set forth in Secs. 3.3.2–3.3.6 return simple, nonfractal forms. These measures thus *cannot* distinguish between a fractal renewal process and a nonfractal one such as the homogeneous Poisson process. Figure 5.4b) illustrates the indistinguishability of the rescaled range statistic for these two processes. The interevent interval density (or distribution) set forth in Sec. 3.3.1 forms the sole exception. Since it readily reveals fractal behavior in fractal renewal processes it *can* distinguish the two forms.

As the second example, we consider a doubly stochastic Poisson process driven by a fractal Gaussian process (Sec. 6.3.3) whose rate has a low coefficient of variation, $C_\mu \ll 1$. In this case, all of the measures set forth in Secs. 3.3.2–3.3.6 *can* serve to discriminate between this process and a homogeneous Poisson process. However, the interevent-interval density and distribution *cannot* distinguish between them; although the interval ordering differs, quite similar interval densities characterize the two processes [see Eqs. (4.33) and (4.3)]. Figure 5.5a) illustrates the indistinguishability of the interval densities for these two processes.

We conclude that the statistics in Secs. 3.3.2–3.3.6 fail for the former process whereas those in Sec. 3.3.1 fail for the latter, so that no interevent-interval statistic proves useful in all cases. Even in cases where dependencies among intervals do exist, the measures in Secs. 3.3.2–3.3.6 can suffer from problems of interpretation since these measures describe the point process in terms of a time axis that is warped

with respect to real time, as shown in Fig. 3.1e). In particular, the interval-based autocorrelation and spectrum reliably describe fractal behavior only when the coefficient of variation of the intervals is small (DeBoer et al., 1984; Turcott & Teich, 1996).

Nevertheless, some interval-based measures can be useful for analyzing point processes whose fractal characteristics derive only from the relative ordering of the interevent intervals, and not from their distribution. This holds for the general doubly stochastic Poisson process, results for which we report throughout this chapter. We therefore include four often-used interval-based measures within our purview: rescaled range analysis (Sec. 3.3.5), detrended fluctuation analysis (Sec. 3.3.6), the interval-based wavelet variance (Sec. 3.3.4), and the interval-based spectrum (periodogram) (Sec. 3.3.3). We continue to bear in mind, however, that these measures do not always yield reliable results for general fractal-based point processes.

The point-process-based measures presented in Sec. 3.5 also prove impractical in the general case. Generalized dimensions for fractal-rate point processes assume integer (nonfractal) values, as discussed in Sec. 3.5.4, whereas the remaining measures set forth in Sec. 3.5 all have count-based analogs (some with superior statistics) that prove far easier to calculate. For example, estimates of the coincidence rate $G(t)$ based on a finite-length data set containing N points comprise $N(N-1)/2$ Dirac delta functions representing the delay times that happened to occur in that particular data set, and another N delta functions at zero delay time. Increasing the number of points in the data set yields a greater number of delta functions, but does not cause the coincidence-rate estimate to converge to a smooth form. One solution is to bin the events. Rather than using the point process $dN(t)$ to generate the coincidence rate $G(t)$, this procedure effectively makes use of the counting process $\{Z_k(T)\}$ to construct an autocorrelation $R_Z(k,T)$, which yields smooth results for finite data sets. Indeed, Eq. (3.54) illustrates how the autocorrelation derives from a smoothed version of the coincidence rate (see Prob. 12.4).

We are left with four count-based measures whose merit we wish to assess for the estimation problem at hand: the normalized variance, normalized Haar-wavelet variance, autocorrelation, and rate spectrum. In addition we study the performance of the four interval-based measures indicated immediately above. We proceed to examine these in turn.

12.3.2 Normalized variance

As mentioned in Sec. 3.4.2, the normalized variance suffers from bias and thus proves problematical as a measure. To explicitly demonstrate this, we begin with the estimate of the count-sequence variance,

$$\begin{aligned}\widehat{\mathrm{Var}}[Z(T)] &= (M-1)^{-1}\sum_{n=0}^{M-1}\left\{Z_n(T)-\widehat{\mathrm{E}}[Z(T)]\right\}^2 \\ &= (M-1)^{-1}\sum_{n=0}^{M-1}\left\{Z_n(T)-M^{-1}\sum_{m=0}^{M-1}Z_m(T)\right\}^2\end{aligned}$$

$$
\begin{aligned}
&= (M-1)^{-1} \sum_{n=0}^{M-1} Z_n^2(T) \\
&\qquad - M^{-1}\,(M-1)^{-1} \sum_{m=0}^{M-1} \sum_{n=0}^{M-1} Z_m(T)\, Z_n(T) \\
&= M^{-1} \sum_{n=0}^{M-1} Z_n^2(T) \\
&\qquad - M^{-1}\,(M-1)^{-1} \sum_{m=0}^{M-1} \sum_{n\neq m}^{M-1} Z_m(T)\, Z_n(T), \qquad (12.4)
\end{aligned}
$$

where $M = \mathrm{int}(L/T)$ represents the number of counts and $\mathrm{int}(x)$ is the largest integer not exceeding x.

The count-sequence variance has an expected value given by

$$
\begin{aligned}
&\mathrm{E}\big\{\widehat{\mathrm{Var}}[Z(T)]\big\} \\
&\quad = M^{-1} \sum_{n=0}^{M-1} \mathrm{E}\big[Z_n^2(T)\big] \\
&\qquad - M^{-1}\,(M-1)^{-1} \sum_{m=0}^{M-1} \sum_{n\neq m}^{M-1} \mathrm{E}[Z_n(T)\, Z_m(T)] \qquad (12.5) \\
&\quad = \mathrm{E}[Z^2(T)] - 2 \sum_{k=1}^{M} \frac{(M-k)\, R_Z(k,T)}{M(M-1)}.
\end{aligned}
$$

To evaluate this expression, we make use of Eqs. (5.13) (for $k = 0$), (5.14), and (5.45), which yield

$$
\begin{aligned}
&\mathrm{E}\big\{\widehat{\mathrm{Var}}[Z(T)]\big\} \\
&\quad = \mathrm{E}^2[Z(T)] + \mathrm{E}[Z(T)] + \mathrm{E}[Z(T)]\,(T/T_F)^{\alpha} \\
&\qquad - \sum_{k=1}^{M} \frac{2(M-k)}{M(M-1)} \Big\{ \mathrm{E}^2[Z(T)] + \mathrm{E}[Z(T)] \\
&\qquad\quad + \tfrac{1}{2}\mathrm{E}[Z(T)]\,(T/T_F)^{\alpha} \left[(k+1)^{\alpha+1} + (k-1)^{\alpha+1} - 2k^{\alpha+1}\right] \Big\} \\
&\quad = \mathrm{E}[Z(T)]\,(T/T_F)^{\alpha} \\
&\qquad \times \left\{ 1 - \sum_{k=1}^{M} \frac{(M-k)}{M(M-1)} \left[(k+1)^{\alpha+1} + (k-1)^{\alpha+1} - 2k^{\alpha+1}\right] \right\} \\
&\quad \approx \mathrm{E}[Z(T)]\,(T/T_F)^{\alpha} \\
&\qquad \times \left\{ 1 - \int_{x=1}^{M} \frac{(M-x)}{M(M-1)} \left[(x+1)^{\alpha+1} + (x-1)^{\alpha+1} - 2x^{\alpha+1}\right] dx \right\}
\end{aligned}
$$

$$\begin{aligned} &\approx \mathrm{E}[Z(T)]\,(T/T_F)^{\alpha} \\ &\quad \times \left\{ 1 - \int_{x=0}^{M} \frac{(M-x)}{M^2} \left[\alpha(\alpha+1)\, x^{\alpha-1} \right] dx \right\} \\ &= \mathrm{E}[Z(T)]\,(T/T_F)^{\alpha} \left[1 - M^{\alpha-1} \right] \\ &= \mathrm{E}[Z(T)]\,(T/T_F)^{\alpha} \left[1 - (T/L)^{1-\alpha} \right]. \end{aligned} \tag{12.6}$$

The bias in Eq. (12.6) carries over to the estimate of the normalized variance $\widehat{F}(T)$ as a similar bias. The result is a spuriously low fractal-exponent estimate $\widehat{\alpha}$. This appears in the simulation results presented in Fig. 12.2 and Table 12.2, which illustrate the substantial departure of this statistic from precise T^{α} behavior.

Moreover, as indicated in Sec. 5.2.3, the normalized variance cannot rise faster than T^1. As a result of the inherent bias in its estimate, and its inability to reveal fractal exponents greater than unity, we do not recommend use of the normalized variance $\widehat{F}(T)$. The normalized Haar-wavelet variance $\widehat{A}(T)$ is a far superior statistic, as comparison of Fig. 12.2 with 12.1, and Table 12.2 with 12.1, affirms.

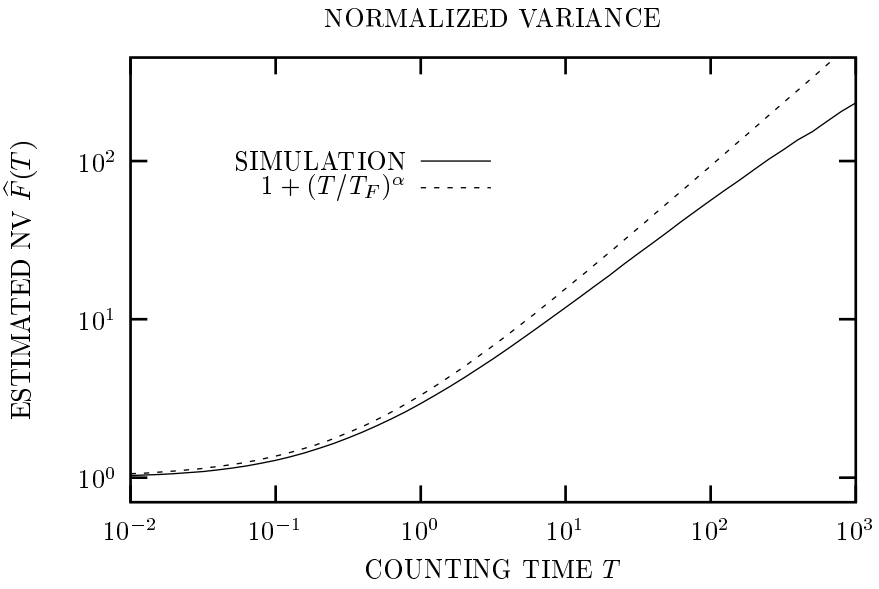

Fig. 12.2 Estimated normalized variance $\widehat{F}(T)$ vs. counting time T, based on the same simulations as those used to generate Fig. 12.1 (a fractal-Gaussian-process-driven Poisson process). We present the mean value (solid curve) along with the simple theoretical form provided in Eq. (5.44b) (dashed). The simulation results deviate from this ideal fractal behavior, especially for long counting times T, illustrating the substantial bias inherent in the normalized variance statistic.

Time Range		$\widehat{\alpha}_F$ from $\widehat{F}(T)$			$\widehat{\alpha}_{F-1}$ from $\widehat{F}(T)-1$		
$T_{\min}$	$T_{\max}$	Bias	SD	RMSE	Bias	SD	RMSE
10^0	10^1	-0.191	0.022	0.192	-0.051	0.018	0.054
10^0	10^2	-0.151	0.028	0.153	-0.073	0.026	0.077
10^0	10^3	-0.159	0.050	0.167	-0.112	0.048	0.122
10^1	10^2	-0.126	0.039	0.132	-0.096	0.037	0.103
10^1	10^3	-0.167	0.074	0.183	-0.151	0.074	0.168
10^2	10^3	-0.221	0.141	0.262	-0.216	0.141	0.258

Table 12.2 Performance of fractal-exponent estimates $\widehat{\alpha}$ obtained via estimates of the normalized variance $\widehat{F}(T)$. See Table 12.1 for details. Best results (least root-mean-square error) obtain by using $10^1 \le T \le 10^2$ for $\widehat{F}(T)$, and by using $10^0 \le T \le 10^1$ for $\widehat{F}(T) - 1$. We find that $\widehat{\alpha} < \alpha$ in all cases; the deviation increases with increasing values of T. Bias is problematic for this estimator. For all ranges of counting time, the normalized Haar-wavelet variance $\widehat{A}(T)$ returns significantly better results.

12.3.3 Count autocorrelation

The autocorrelation $R_Z(k,T)$ of the counting sequence $\{Z_k(T)\}$, as a function of the delay k, forms a windowed version of the coincidence rate $G(t)$, as demonstrated by Eq. (3.54). Like the coincidence rate, the limit for large delays dominates the fractal portion for all but the smallest delays. Rearranging Eq. (5.44d) yields

$$R_Z(k,T)/\mathrm{E}[Z(T)] = \mathrm{E}[Z(T)] + (T/T_R)^{\alpha}\, k^{\alpha-1}. \tag{12.7}$$

Since $\alpha < 1$ (a condition that is required for this measure to demonstrate any fractal behavior), the last term decreases with increasing k, making estimates of its form difficult for all but small values of k.

We therefore use instead the autocovariance (autocorrelation minus the square of the mean), which eliminates the dominant constant term. Unlike other measures considered in this chapter, the count-based autocorrelation (or autocovariance) takes two arguments; analyses as a function of delay number k require a counting time T as a parameter. Choosing a long counting time leads to loss of information at shorter time scales. In contrast, small counting times lead to excessive noisiness in the resulting computed autocovariance and, as mentioned earlier, nonfractal behavior often dominates for these small times. We choose $T = 1$ as a compromise for this simulation, which proves close to optimal.

Figure 12.3 and Table 12.3 present the simulation results for the normalized autocovariance, that is, the autocorrelation minus the square of the mean, normalized by the variance, as a function of the delay number k:

$$R_2(k) \equiv \frac{R_Z(k,T) - \mathrm{E}^2[Z(T)]}{\mathrm{Var}[Z(T)]}\,. \tag{12.8}$$

Unlike the presentations in Tables 12.1 and 12.2, we present only one set of columns here; we must subtract the square-mean number of counts or essentially constant behavior results. This estimator performs poorly in comparison with the normalized Haar-wavelet variance $\widehat{A}(T)$; it is inferior even to the normalized variance $\widehat{F}(T)$.

Since the autocovariance has two arguments, one can consider this quantity as a function of the counting time T with a fixed delay number k. Examination of Eqs. (5.44d) and (12.7) indeed reveals a T^{α} dependency. However, rearrangement of these equations yields a diminished (and therefore relatively more noisy) version of the normalized variance:

$$\begin{aligned} R_Z(k,T)/\mathrm{E}[Z(T)] - \mathrm{E}[Z(T)] &= (T/T_R)^{\alpha}\, k^{\alpha-1} \\ &= [F(T)-1]\,(T_F/T_R)^{\alpha}\, k^{\alpha-1} \\ &= \tfrac{1}{2}\alpha(\alpha+1)\, k^{\alpha-1}\,[F(T)-1]. \end{aligned} \quad (12.9)$$

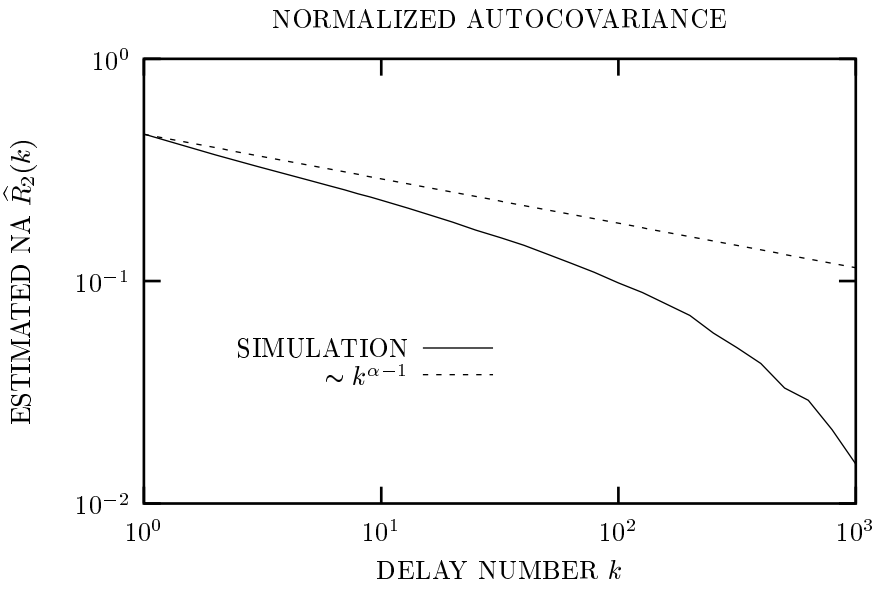

Fig. 12.3 Plot of $\widehat{R}_2(k) = \{\widehat{R}_Z(k,T) - \widehat{\mathrm{E}}^2[Z(T)]\}\big/\widehat{\mathrm{Var}}[Z(T)]$ vs. number of intervals k, the estimated normalized autocovariance, based on the same simulations as those used to generate Figs. 12.1 and 12.2 (a fractal-Gaussian-process-driven Poisson process). We obtained values of the autocovariance and variance for each simulation run, and then averaged the ratio of these two quantities (solid curve). This normalization provides results in the spirit of Figs. 12.1 and 12.2. We chose $T = 1$ for this simulation. A simple theoretical form proportional to $k^{\alpha-1}$, following Eq. (5.44d), also appears (dashed curve). The large disagreement between the simulation results and the theoretical behavior, especially at large delays k, illustrates the bias inherent in the count-based autocovariance.

Delay Range		$\widehat{\alpha}_{R2}$ from $\{\widehat{R}_Z(k,T) - \widehat{\mathrm{E}}^2[Z(T)]\}/\widehat{\mathrm{Var}}[Z(T)]$		
$k_{\min}$	$k_{\max}$	Bias	SD	RMSE
10^0	10^1	-0.097	0.034	0.103
10^0	10^2	-0.139	0.058	0.150
10^0	10^3	-0.259	0.106	0.280
10^1	10^2	-0.185	0.100	0.211
10^1	10^3	-0.360	0.163	0.395
10^2	10^3	-0.611	0.417	0.739

Table 12.3 Performance of fractal-exponent estimates $\widehat{\alpha}$ obtained via estimates of the count-based autocovariance, $\widehat{R}_Z(k,T) - \widehat{\mathrm{E}}^2[Z(T)]$, for different ranges of the delay k. See Table 12.1 for details. Best results (least root-mean-square error) obtain for the shortest delay ranges; however, the bias is excessive for all delays and the standard deviation is large for all but the shortest delays. Results returned by the normalized variance $\widehat{F}(T)$ are better, and those returned by the normalized Haar-wavelet variance $\widehat{A}(T)$ are uniformly superior. This estimator does not appear to be useful.

Equation (12.9) evidently suffers from all of the deficiencies of the normalized variance $F(T)$; moreover, it is afflicted by a multiplicative factor $k^{\alpha-1}$, always less than unity, that decreases the useful term while leaving the constant, nonvarying terms unchanged.

We conclude that the autocorrelation $\widehat{R}_Z(k,T)$ is deficient as a statistic when considered as a function either of the delay k or the counting time T. We do not recommend its use.

12.3.4 Rescaled range

The rescaled range $U(k)$ of the interevent-interval sequence $\{\tau_k\}$, as a function of the delay k, yields information about dependencies among the intervals. Although it fails to reveal fractal behavior in the fractal renewal process (see Sec. 12.3.1), and presents information on a warped time axis, it enjoys substantial use as a measure of fractal activity and is therefore worth examining. Hurst introduced this measure in 1951, as discussed in Sec. 3.3.5. The rescaled range varies as $\sqrt{k}$ for independent intervals and, more generally as k^H, where $H = (1+\alpha)/2$, for a fractal process with fractal exponent α (see Sec. 6.3.1).

To facilitate the comparison of this measure with others, we examine the statistics of

$$U_2(k) \equiv U^2(k)/k\,. \tag{12.10}$$

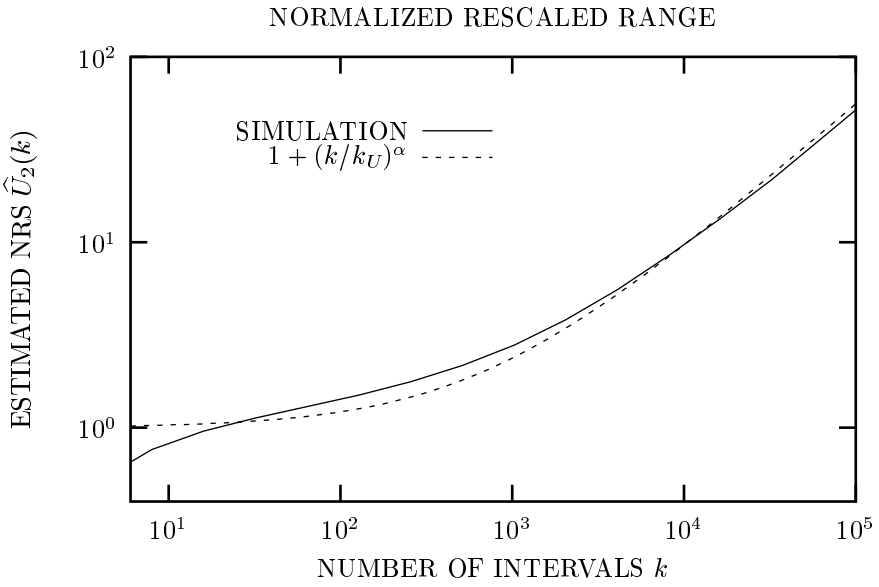

Fig. 12.4 Normalized estimated rescaled range $\widehat{U}_2(k) \equiv \widehat{U}^2(k)/k$, as a function of the number of intervals k, based on the same simulations as those used to generate Figs. 12.1–12.3 (a fractal-Gaussian-process-driven Poisson process). Results begin at $k = 4$, and increase by a factor of 2 per displayed point thereafter. We obtained values of the R/S statistic for each simulation run, normalized them so that they came close to unity for independent intervals [converted to $\widehat{U}_2(k)$], and then averaged all 100 runs together (solid curve). A theoretical form akin to those shown in Figs. 12.1–12.3 also appears (dashed curve), employing $k_U = 671.3$, which minimizes the sum of squares of the differences in the logarithms. A simpler form, proportional to k^α, provides an even worse fit to the data (not shown). The deviation of the simulation results from these curves, even for large interval numbers k, illustrates the large bias inherent in the rescaled range statistics.

For a fractal-based point process, $U_2(k)$ varies as k^α, much like the normalized Haar-wavelet variance $A(T)$ which varies as T^α. In the context of the simulations employed in this chapter, for small values of k the underlying fractal rate changes little so that the intervals are relatively independent in these local neighborhoods. We therefore expect that $U(k) \sim \sqrt{k}$ so that $U_2(k) \sim k^0$ for small k. On the other hand, large values of k should encompass significant fractal fluctuations, so that $U(k) \sim k^H$ and $U_2(k) \sim k^\alpha$.

We expect the value of k where these two behaviors meet (the onset value) to be of the order of the fractal onset times for the normalized variance and normalized Haar-wavelet variance (about 1 time unit), multiplied by the average rate (100 per time unit); the onset should therefore occur at about $k = 100$. Figure 12.4 and Table 12.4 present simulation results for the normalized estimated rescaled range $\widehat{U}_2(k)$. The averaged results presented in Fig. 12.4 follow the theoretical power-law form k^α

Time Range		$\widehat{\alpha}_{U2}$ from $\widehat{U}^2(k)/k$			$\widehat{\alpha}_{U2-1}$ from $\widehat{U}^2(k)/k-1$		
$k_{\min}$	$k_{\max}$	Bias	SD	RMSE	Bias	SD	RMSE
10^2	10^3	-0.537	0.008	0.537	-0.194	0.016	0.195
10^2	10^4	-0.386	0.010	0.386	-0.154	0.012	0.155
10^2	10^5	-0.282	0.015	0.282	-0.116	0.015	0.117
10^3	10^4	-0.265	0.022	0.266	-0.112	0.025	0.114
10^3	10^5	-0.175	0.026	0.177	-0.077	0.027	0.082
10^4	10^5	-0.069	0.101	0.122	-0.033	0.105	0.110

Table 12.4 Performance of fractal-exponent estimates $\widehat{\alpha}$ obtained via normalized estimates of the rescaled range (R/S), $\widehat{U}_2(k) \equiv \widehat{U}^2(k)/k$, as a function of the number of intervals k. See Table 12.1 for details. Best results (least root-mean-square error) obtain by using $10^4 \le k \le 10^5$ for $\widehat{U}_2(k)$, and by using $10^3 \le k \le 10^5$ for $\widehat{U}_2(k) - 1$. Results obtained by using $\widehat{A}(T)$, the normalized Haar-wavelet variance, are uniformly superior to those based on the normalized rescaled range. This estimator does not appear to be useful.

only approximately, seriously diverging from it for smaller values of k; the function $1 + (k/k_U)^\alpha$ with $k_U \approx 671.3$ provides a better fit to $\widehat{U}_2(k)$ but still differs considerably from it. Employing $\widehat{U}_2(k)$ instead of $\widehat{U}(k)$ permits us to subtract the short-delay asymptote (which is approximately unity), but doing so improves the performance only marginally inasmuch as the asymptote bears only a slight likeness to the simulation. Overall, the normalized Haar-wavelet variance $\widehat{A}(T)$ proves substantially superior to the rescaled range for estimating the fractal exponent $\widehat{\alpha}$.

12.3.5 Detrended fluctuation analysis

The detrended fluctuation $Y(k)$ of the interevent-interval sequence $\{\tau_k\}$, as a function of the delay k, yields information about dependencies among the intervals related to those provided by the rescaled range $U(k)$. Like the rescaled range, it fails to reveal fractal behavior in the fractal renewal process (see Sec. 12.3.1) but we nevertheless investigate its performance. Taqqu & Teverovsky (1998) have pointed out that this measure exhibits significant bias and variance, except for the special case of Gaussian-distributed sequences.

The detrended fluctuation resembles the rescaled range in that $Y(k)$ typically varies as k^H, where $H = (1 + \alpha)/2$ for a process with fractal exponent α. Using Eq. (3.25) to proceed along lines similar to those followed in Sec. 12.3.4, we define a normalized version of this measure:

$$Y_2(k_2) \equiv \frac{15\, Y^2(k+2)}{(k+2)\,\mathrm{Var}[\tau]}\,. \tag{12.11}$$

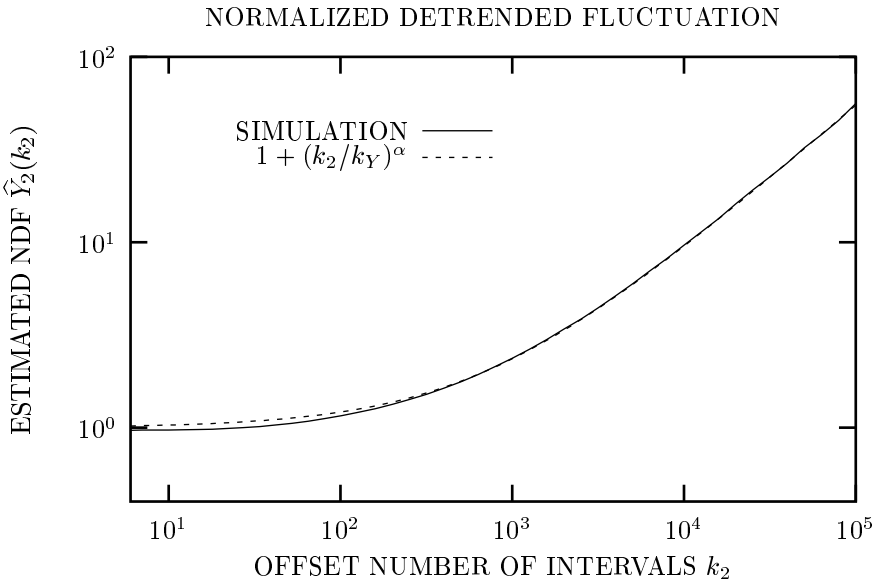

Fig. 12.5 Normalized estimated detrended fluctuation shown in Eq. (12.11), as a function of the offset number of intervals k_2, based on the same simulations as those used to generate Figs. 12.1–12.4 (a fractal-Gaussian-process-driven Poisson process). We define $k_2 \equiv k + 2$, the offset delay number. Results begin with $k_2 = 5$ ($k = 3$), and increase by factors of $10^{0.1}$ thereafter (excluding multiple copies of identical values of k_2). We obtained values of the detrended fluctuation statistic for each simulation run, normalized them to yield values equal to unity for independent intervals as specified in Eq. (12.11), and then averaged all 100 runs together (solid curve). The simple form provided in Eq. (12.12) fits the simulated data very well (dashed curve).

We present the simulation results for the normalized estimated detrended fluctuation $\widehat{Y}_2(k)$ in Fig. 12.5 and Table 12.5. Overall, we expect this measure to vary as

$$Y_2(k_2) \approx 1 + (k_2/k_Y)^\alpha. \tag{12.12}$$

Equation (12.12) does indeed fit the simulations presented in Fig. 12.5 very well, with only one free parameter, k_Y, chosen to minimize the sum of squares of the differences in the logarithms. This takes a value of $k_Y = 684.6$, which is within an order of magnitude of our original estimate in Sec. 12.3.4.[4]

[4] Although Eq. (12.11) reduces to $Y_2(k_2) = 1 - 4/k_2^2$, given Eq. (3.25), for our simulations it nevertheless yields results that closely approach those of Eq. (12.12). We therefore employ it rather than a simpler form without the $+2$ offset. The origin of the offset may lie in the warping involved in translating the process from time in seconds to dimensionless interval number, which, in turn, may affect the representations of the relevant correlations. A number of *ad hoc* methods have been devised to address this issue (see, for example, Buldyrev et al., 1995), but they do not provide as good a fit to our simulations.

Time Range		$\widehat{\alpha}_{Y2}$ from $\widehat{Y}_2(k_2)$			$\widehat{\alpha}_{Y2-1}$ from $\widehat{Y}_2(k_2)-1$		
$k_{2\min}$	$k_{2\max}$	Bias	SD	RMSE	Bias	SD	RMSE
10^2	10^3	-0.490	0.011	0.490	0.123	0.025	0.125
10^2	10^4	-0.336	0.013	0.337	0.046	0.017	0.049
10^2	10^5	-0.225	0.019	0.226	0.020	0.020	0.028
10^3	10^4	-0.189	0.027	0.191	-0.001	0.031	0.031
10^3	10^5	-0.111	0.034	0.116	-0.007	0.035	0.036
10^4	10^5	-0.052	0.087	0.101	-0.014	0.089	0.091

Table 12.5 Performance of fractal-exponent estimates $\widehat{\alpha}$ obtained using normalized estimates of the detrended fluctuation as defined in Eq. (12.11), as a function of offset delay number $k_2 \equiv k+2$. See Table 12.1 for details. Best results (least root-mean-square error) obtain by using $10^4 \le k_2 \le 10^5$ for $\widehat{Y}_2(k_2)$, and by using $10^2 \le k_2 \le 10^5$ for $\widehat{Y}_2(k_2)-1$. This measure yields errors nearly as small as those returned by the normalized Haar-wavelet variance $\widehat{A}(T)$; it is substantially superior to the rescaled range statistic examined in Fig. 12.4 and Table 12.4.

Employing $\widehat{Y}_2(k_2)$ instead of $\widehat{Y}(k)$ again permits the short-delay asymptote of about unity to be subtracted, and in this case the excellent fit of Eq. (12.12) to the simulated data imparts substantial improvement with this procedure. Detrended fluctuation analysis thus yields results only slightly inferior to those provided by the normalized Haar-wavelet variance $\widehat{A}(T)$ for the point process under study.

12.3.6 Interval wavelet variance

The wavelet variance $\mathrm{Var}[W_{\psi,\tau}(k,l)]$ of the interevent-interval sequence $\{\tau_k\}$ as a function of the delay k, as set forth in Sec. 3.3.4, provides another window on the second-order properties of this sequence. Although different in construction from the rescaled range $U(k)$ and detrended fluctuation $Y(k)$, it nevertheless has much in common with these two measures. As before, we consider the merits of this measure, bearing in mind the same cautions regarding its inability to discriminate fractal renewal processes from nonfractal point processes (see Sec. 12.3.1). We choose the Haar wavelet to simplify processing and to minimize the variance of the resulting estimate (see Sec. 5.2).

Unlike rescaled range and detrended fluctuation analysis, the wavelet variance $\mathrm{Var}[W_{\psi,\tau}(k,l)]$ typically varies as k^{α} rather than as k^H for a process with fractal exponent α. Hence, we need only divide by the variance of interevent intervals to render this quantity dimensionless, as provided in Eq. (3.21):

$$A_\tau(k) \equiv \frac{\mathrm{Var}[W_{\psi,\tau}(k,l)]}{\mathrm{Var}[\tau]}. \tag{12.13}$$

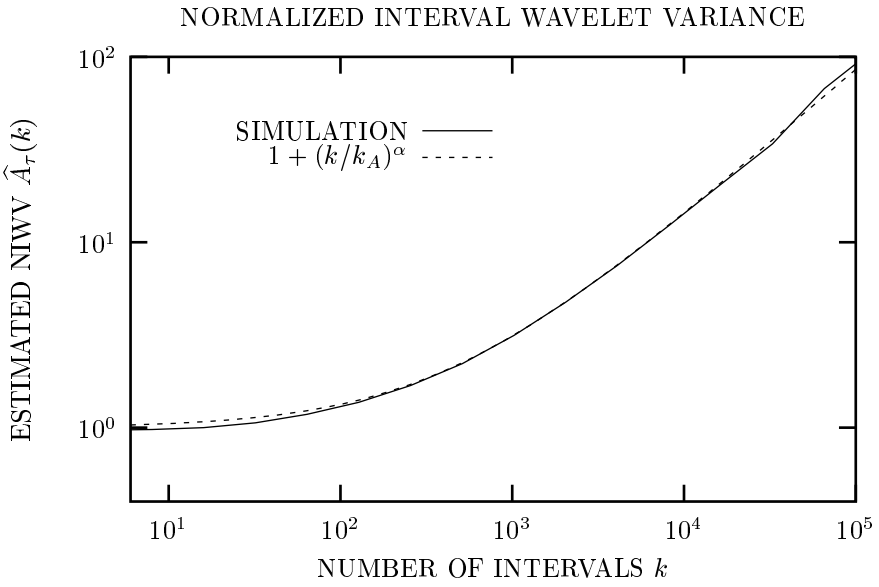

Fig. 12.6 Normalized estimated interval-based wavelet variance $\widehat{A}_\tau(k)$ shown in Eq. (12.13), as a function of wavelet scale (number of intervals) k, based on the same simulations as those used to generate Figs. 12.1–12.5 (a fractal-Gaussian-process-driven Poisson process). Results begin with $k = 2$, and increase by factors of 2 thereafter. We obtained values of the wavelet variance statistic for each simulation run, normalized them to yield values equal to unity for independent intervals as specified in Eq. (12.13), and then averaged all 100 runs together (solid curve). The simple theoretical form shown in Eq. (12.14) fits the simulated data very well (dashed curve). The increased difference between the two curves for the largest numbers of intervals derives from the stochastic fluctuations in the simulations themselves. Changing only the random seeds used in the simulations, but leaving all other parameters intact, yields an averaged $\widehat{A}_\tau(k)$ that exhibits different fluctuations, albeit of similar magnitude (not shown).

We again assume that the warping required to transform time in seconds to dimensionless interval number only marginally affects the results. As a parallel to Eq. (12.12), we expect the normalized interval-based wavelet variance to follow the form

$$A_\tau(k) \approx 1 + (k/k_A)^\alpha. \tag{12.14}$$

Equation (12.14) indeed provides an excellent fit to the simulated data, as shown in Fig. 12.6, by choosing the one free parameter to be $k_A \doteq 389.2$. This value minimizes the sum of squares of the differences in the logarithms. Again, the parameter k_A lies within an order of magnitude of our original estimate in Sec. 12.3.4. As with the rescaled range and detrended fluctuation, using $\widehat{A}_\tau(k)$, rather than $\widehat{\mathrm{Var}}[W_{\psi,\tau}(k,l)]$, permits us to subtract the short-delay asymptote near unity. As with $Y_2(k_2)$, Eq. (12.14) fits the simulated data extremely well and offers a considerable decrease in the estimation error; the results are only slightly inferior to

Time Range		$\widehat{\alpha}_{A\tau}$ from $\widehat{A}_\tau(k)$			$\widehat{\alpha}_{A\tau-1}$ from $\widehat{A}_\tau(k)-1$		
$k_{\min}$	$k_{\max}$	Bias	SD	RMSE	Bias	SD	RMSE
10^2	10^3	-0.454	0.024	0.454	0.057	0.055	0.079
10^2	10^4	-0.269	0.020	0.270	0.016	0.025	0.030
10^2	10^5	-0.179	0.029	0.182	0.008	0.029	0.030
10^3	10^4	-0.147	0.059	0.158	-0.005	0.067	0.067
10^3	10^5	-0.083	0.054	0.099	-0.001	0.056	0.056
10^4	10^5	0.029	0.298	0.299	0.054	0.306	0.311

Table 12.6 Performance of fractal-exponent estimates $\widehat{\alpha}$ obtained via normalized estimates of the interval-based wavelet variance as defined in Eq. (12.13), as a function of wavelet scale k. See Table 12.1 for details. Best results (least root-mean-square error) obtain by using $10^3 \leq k \leq 10^5$ for $\widehat{A}_\tau(k)$, and by using $10^2 \leq k \leq 10^4$ or 10^5 for $\widehat{A}_\tau(k)-1$. This measure yields errors comparable to those generated by normalized detrended fluctuation analysis, and nearly as small as those returned by the normalized Haar-wavelet variance.

those provided by the normalized Haar-wavelet variance $\widehat{A}(T)$ (compare Tables 12.6 and 12.1). For the Poisson-based process at hand, multiplying the abscissa by half the average interevent time[5] yields a plot that is nearly coincident with that of the normalized Haar-wavelet variance $\widehat{A}(T)$. In particular, we obtain a more precise value for the fractal onset number:

$$k_A = \frac{T_A}{\mathrm{E}[\tau]/2} \doteq 1.8939/0.005 = 378.78, \tag{12.15}$$

which is in excellent accord with the rough estimate of $k_A \doteq 389.2$ obtained empirically.

Agreement of this kind between $\widehat{A}(T)$ and $\widehat{A}_\tau(k)$ does not emerge in general, however. Consider, for example, a point process whose interval standard deviation lies well below the interval mean. For small numbers of intervals k, the interval-based wavelet variance $A_\tau(k)$ can assume arbitrarily small values, as a result of the small relative standard deviation of the intervals. However, $A(T)$ must approach an asymptote of unity for small values of T, so it can diverge significantly from $A_\tau(2T/\mathrm{E}[\tau])$. The human-heartbeat point process provides a case in point; examination of Figs. 5.2 and 5.8 reveals precisely this difference. The underlying reason for this disparity lies in the integrate-and-reset kernel associated with useful models for the heartbeat process. Poisson kernels, in contrast, yield plots of $\widehat{A}(T)$ and $\widehat{A}_\tau(k)$ that nearly coincide, as shown in the simulations presented here.

[5] The factor of a half arises from the definition of the count-based normalized Haar-wavelet variance $A(T)$, for which T encompasses half the duration of the Haar wavelet [see Eq. (3.40)].

12.3.7 Interval spectrum

The interval-based spectral estimate (periodogram) $\widehat{S}_\tau(\mathsf{f})$ forms the last interval-based measure we consider. As with other such measures, we note its most serious limitation, its inability to discriminate fractal renewal processes from nonfractal point processes (see Sec. 12.3.1).

Like the wavelet variance $\mathrm{Var}[W_{\psi,\tau}(k,l)]$, but unlike the rescaled range and detrended fluctuation, the interval-based periodogram typically varies as $1/\mathsf{f}^\alpha$ rather than as $1/\mathsf{f}^H$ for a process with fractal exponent α. Normalizing by $\mathrm{Var}[\tau]$, the estimated high-frequency asymptote as discussed in Sec. 3.3.3, yields a dimensionless form that achieves a value of unity for large interval frequencies f. We again assume that the warping required to transform time in seconds to dimensionless interval number only marginally affects the results. In another parallel to Eq. (12.12), we expect

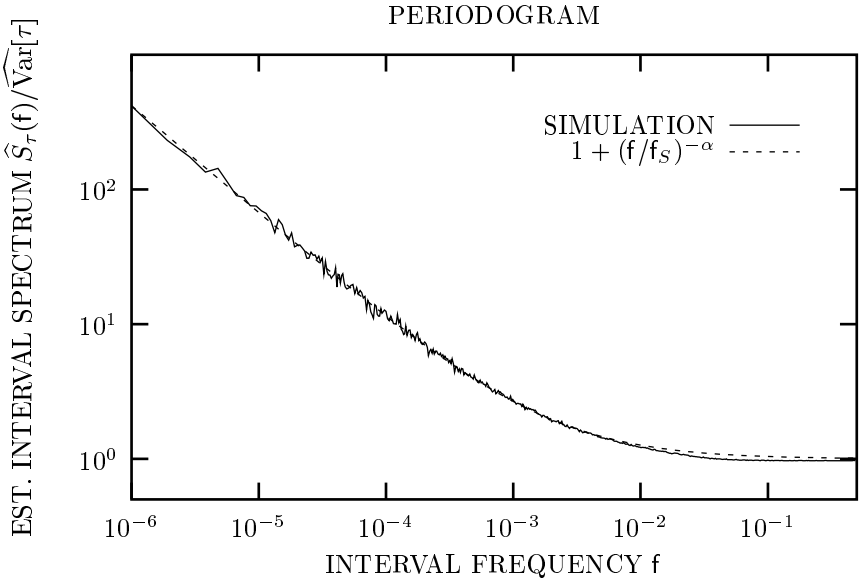

Fig. 12.7 Normalized estimated interval-based spectrum (periodogram) $S_\tau(\mathsf{f})/\widehat{\mathrm{Var}}[\tau]$ vs. interval frequency f, based on the same simulations as those used to generate Figs. 12.1–12.6 (a fractal-Gaussian-process-driven Poisson process). For each simulation, we calculated the (fast) Fourier transform of the data, using all 2^{20} intervals in a single transform without windowing. Next, we took the square of the magnitude, and divided by the sample variance of the intervals of that simulation. We then collected values into blocks whose frequencies differed by less than a factor of 1.02, and plotted a single point comprising the mean of the frequencies and the mean of the spectral estimate. Finally, we averaged all 100 curves together. This appears as the solid curve. Equation (12.16) generates a simple theoretical curve (dashed) which fits the simulated data very well.

Freq. Range		$\widehat{\alpha}_{S\tau}$ from $\widehat{S}_\tau(\mathsf{f})$			$\widehat{\alpha}_{S\tau-\mathrm{Var}[\tau]}$ from $\widehat{S}_\tau(\mathsf{f}) - \widehat{\mathrm{Var}}[\tau]$		
$\mathsf{f}_{\min}$	$\mathsf{f}_{\max}$	Bias	SD	RMSE	Bias	SD	RMSE
10^0	10^1	0.008	0.606	0.606	0.095	0.843	0.848
10^0	10^2	-0.032	0.137	0.141	-0.174	0.179	0.250
10^0	10^3	-0.168	0.037	0.172	-0.070	0.052	0.087
10^0	10^4	-0.436	0.015	0.436	0.309	0.377	0.487
10^1	10^2	-0.054	0.227	0.234	-0.230	0.320	0.394
10^1	10^3	-0.189	0.040	0.193	-0.059	0.057	0.082
10^1	10^4	-0.446	0.015	0.447	0.324	0.389	0.506
10^2	10^3	-0.256	0.067	0.264	-0.010	0.089	0.090
10^2	10^4	-0.483	0.016	0.483	0.384	0.448	0.589
10^3	10^4	-0.562	0.023	0.563	0.644	0.746	0.985

Table 12.7 Performance of fractal-exponent estimates $\widehat{\alpha}$ for 100 different simulations of a fractal-Gaussian-process-driven Poisson process obtained from estimates of the interval spectrum (periodogram) $\widehat{S}_\tau(\mathsf{f})$. The simulations were the same as those employed to produce Table 12.1. We calculated the Fourier transform, and took the square magnitude. The first two columns specify the range of frequencies employed for calculating $\widehat{S}_\tau(\mathsf{f})$. We express frequency in terms of its product with the number of retained intervals N; thus, the first row corresponds to the frequency range $1/N \leq \mathsf{f} \leq 10/N$. The next three columns represent the bias, standard deviation (SD), and resulting root-mean-square error (RMSE) for calculations employing the interval spectrum $\widehat{S}_\tau(\mathsf{f})$. We next sought to remove the effects of the high-frequency asymptote $\widehat{\mathrm{Var}}[\tau]$, and re-calculate the fractal exponent. Since the interval spectrum fluctuates about this asymptote, many of the resulting differences would lie below zero, making calculation of the logarithm meaningless. We therefore employed averaging and thresholding. Specifically, we averaged the value of $\widehat{S}_\tau(\mathsf{f})$ at a frequency $\mathsf{f} = n/N$ with the $n/16$ values following it in frequency, and then subtracted the estimated asymptote $\widehat{\mathrm{Var}}[\tau]$. Finally, to avoid the occasional nonpositive number that would still result, we replaced all numbers less than $\varepsilon\widehat{\mathrm{Var}}[\tau]$ with $\varepsilon\widehat{\mathrm{Var}}[\tau]$, where $\varepsilon = 2^{-12} = 0.000244140625$. We estimated the fractal exponent employing these modified values; these results appear in columns 6–8. Subtracting the high-frequency asymptote does indeed effectively remove the bias, returning reduced root-mean-square errors. Best results (least root-mean-square error) obtain by using $10^0 \leq N\mathsf{f} \leq 10^2$ for $\widehat{S}_\tau(\mathsf{f})$, and by using $10^1 \leq N\mathsf{f} \leq 10^3$ for $\widehat{S}_\tau(\mathsf{f}) - \widehat{\mathrm{Var}}[\tau]$.

the normalized interval-based periodogram to follow the form

$$S_\tau(\mathsf{f}) \approx 1 + (\mathsf{f}/\mathsf{f}_S)^{-\alpha}. \tag{12.16}$$

Figure 12.7 shows that Eq. (12.16) indeed fits the simulated data well, with $\mathsf{f}_S = 0.001893$ yielding the least mean-square difference in the logarithms. As with the

other interval-based measures, using $\widehat{S}_\tau(\mathsf{f})/\widehat{\mathrm{Var}}[\tau]$ instead of $\widehat{S}_\tau(\mathsf{f})$ makes it possible to subtract unity from the resulting statistic. Since Eq. (12.16) fits the data well, this operation reduces the bias in the subsequent estimates of α. The root-mean-square estimation error indeed improves significantly, although it does not achieve nearly the results of the normalized Haar-wavelet variance $\widehat{A}(T)$ (compare Tables 12.7 and 12.1).

The link that we demonstrated between $A(T)$ and $A_\tau(k)$ also holds between the interval-based and rate-based periodograms, and for the same reason: the simulations employ a Poisson kernel. Again, this connection suggests a more accurate estimate of a fractal onset value, the interval frequency in this case

$$\mathsf{f}_S = \mathrm{E}[\tau] f_S \doteq 0.010024 \times 0.2 = 0.0020048, \tag{12.17}$$

which agrees well with the value 0.001893 estimated above. We reiterate that the close link between interval-based and rate-based measures fails to hold in the general case.

12.3.8 Normalized Haar-wavelet variance

We now consider again, in greater detail, the use of the normalized Haar-wavelet variance $\widehat{A}(T)$ for obtaining the fractal exponent $\widehat{\alpha}$. Our preliminary study of the properties of this statistic appeared in Sec. 12.2.3.

In contrast to the normalized variance $\widehat{F}(T)$ and the autocorrelation $\widehat{R}_Z(k,T)$, considered in Secs. 12.3.2 and 12.3.3, respectively, the normalized Haar-wavelet variance $\widehat{A}(T)$ has the distinct merit that it is free of bias, as shown in Sec. 12.2.3.[6] This statistic offers another important advantage in addition, and that is its wavelet origin. Values computed at scales that differ by factors of two are nearly independent, as are values computed for a given counting time but in nonoverlapping counting windows (Tewfik & Kim, 1992).

One consequence of this wavelet property is that a plot of $\widehat{A}(T)$ appears noisier than that of $\widehat{F}(T)$ for a given set of data. Figure 12.8 illustrates this, where we graphically present $\widehat{A}(T)$ and $\widehat{F}(T)$ from an individual simulation run. We used averages of multiple runs such as these to generate Figs. 12.1 and 12.2. Figure 12.8 demonstrates that, despite its noisier appearance, $\widehat{A}(T)$ proves superior to $\widehat{F}(T)$ for quantifying fractal behavior. In particular, it has the distinct merits that its deviations from ideal fractal behavior have zero mean and are nearly independent. Tables 12.1 and 12.2, and Figs. 12.1 and 12.2, demonstrate this superiority for our canonical point-process model.

Although normalized variances employing other wavelet bases have been developed (Teich et al., 1996; Heneghan et al., 1996), fractal-based point processes rarely exhibit fractal exponents exceeding two (see Sec. 5.2.2). This, together with the reduced effective scaling range of more complex wavelets, usually renders them less

[6] As discussed in Sec. 12.2.3, the putative residual bias evident in the sixth column of Table 12.1 is apparently the result of small deviations from ideal fractal behavior that arise in the simulation itself, and that $\widehat{A}(T) - 1$ faithfully reports.

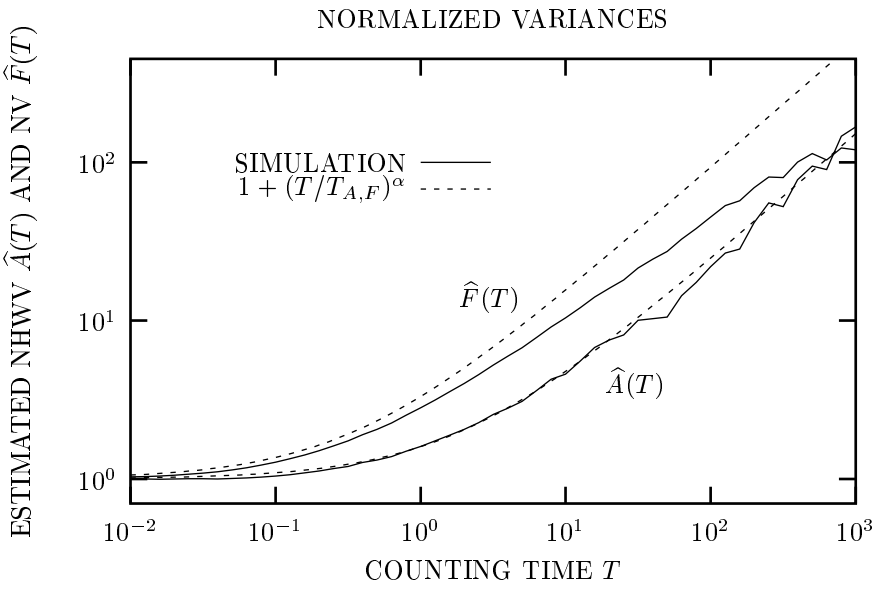

Fig. 12.8 Estimated normalized Haar-wavelet variance $\widehat{A}(T)$ (lower solid curve) and estimated normalized variance $\widehat{F}(T)$ (upper solid curve) vs. counting time T. Averaging these curves with 99 others that are similar yields the results displayed in Figs. 12.1 and 12.2. We provide mean theoretical results from Eqs. (5.44c) and (5.44b), respectively, for comparison (dashed curves). The estimate of the normalized Haar-wavelet variance $\widehat{A}(T)$ exhibits greater fluctuations than that of the normalized variance $\widehat{F}(T)$, yet it exhibits significantly superior performance.

useful than the Haar wavelet, as discussed in Sec. 5.2.5. Such wavelets do prove insensitive to linear or higher-order polynomial trends, unlike the Haar. However, the trends observed in experimental point processes rarely follow exact polynomial forms, so that this insensitivity does not provide a significant advantage. Finally, although there is near independence between wavelet transforms at scales that differ by a factor of two or more, and at different times within the same scale, a small residual correlation does remain. Wavelets other than the Haar do serve to reduce this correlation (Tewfik & Kim, 1992), but the concomitant reduced effective scaling range again typically outweighs any advantage gained by the slight decrease in the within-scale correlation. In the following, we therefore restrict our attention to the Haar version of the count-based wavelet variance, $A(T)$.

We now examine three sampling issues that prove important in obtaining optimal estimates of the fractal exponent $\widehat{\alpha}$ from $\widehat{A}(T)$: (1) **counting-time increments**; (2) **counting-time weighting**; and (3) **oversampling**. We examine these in turn.

Counting-time increments. We first examine the values of T that offer the most accurate estimates of the fractal exponent, when analyzing a fixed-length data set using

the normalized Haar-wavelet variance $\widehat{A}(T)$. Wavelet theory provides a solution. For a discrete-time sequence, a wavelet transform sampled at scales that increase by powers of two retains all of the information in the data; an inverse transform of such a dyadic wavelet transform returns exactly the original sequence. Two caveats apply in the current circumstances, however. First, computing the variance at a given scale collapses all of the transform results at that scale into a single number, thereby eliminating much of the original information. Second, our interest lies in point-process data rather than in discrete-time sequences; however, the difference between the two proves unimportant at longer counting times where fractal behavior becomes significant.

Despite these caveats, optimal sampling turns out not to differ greatly from dyadic sampling. We investigated this issue by recalculating the results shown in Table 12.1 using various counting-time increments. In particular, we reexamine the root-mean-square error of $\widehat{\alpha}$ obtained from $\widehat{A}(T)-1$, which made use of factor of $10^{0.1}$ increments of the counting time, as reported in the right-most column of Table 12.1. Results appear in the top half of Table 12.8 for four counting-time increments: factors of $10^{0.3}$, $10^{0.2}$, $10^{0.1}$, and $10^{0.05}$. Comparing the entries in these four columns reveals that close to optimal results obtain with a spacing of $10^{0.2} \doteq 1.584893$ while nearly dyadic spacing ($10^{0.3} \doteq 1.995262 \approx 2$) yields noticeably inferior performance. A spacing of slightly less than a factor of two appears to yield the best tradeoff between estimation accuracy and computational load.

Counting-time weighting. The optimal weighting for different counting times follows from the accuracy of the normalized Haar-wavelet variance $\widehat{A}(T)$ at a given value of T, which we now derive. From the ideal value provided in Eq. (3.40), we obtain its estimate

$$\widehat{A}(T) = \frac{\widehat{\mathrm{E}}\{[Z_k(T) - Z_{k+1}(T)]^2\}}{2\widehat{\mathrm{E}}[Z_k(T)]} = \frac{\widehat{\mathrm{E}}[C^2_{\mathrm{Haar},N}(2T,k)]}{2\widehat{\mathrm{E}}[\lambda_k(T)]}\,. \tag{12.18}$$

The statistics of $\widehat{A}(T)$ prove difficult to analyze since the calculation involves the division of two random quantities. However, the numerator exhibits far more relative variation than the denominator, since estimates of second-order quantities such as autocorrelations and autocovariances generally fluctuate more than estimates of the mean, a first-order statistic. Moreover, for a given data set, the proportional error in estimating the denominator remains constant for all counting times, since this error derives directly from the same estimate of the mean rate regardless of counting time. After taking the logarithm, this error becomes a constant offset, irrelevant for the estimation of α.

We therefore focus on the numerator alone, since replacing the denominator with its expected value yields the same result:

$$\begin{aligned} \widehat{A}(T) &= \frac{\widehat{\mathrm{E}}[C^2_{\mathrm{Haar},N}(2T,k)]}{2\widehat{\mathrm{E}}[\lambda_k(T)]} \\ &\approx \frac{\widehat{\mathrm{E}}[C^2_{\mathrm{Haar},N}(2T,k)]}{2\mathrm{E}[\lambda_k(T)]} \end{aligned}$$

Time Range			RMSE of $\widehat{\alpha}_{A-1}$ from $\widehat{A}(T)-1$			
$T_{\min}$	$T_{\max}$	Weighting	$10^{0.3}$	$10^{0.2}$	$10^{0.1}$	$10^{0.05}$
10^0	10^1	Equal	0.033	0.031	0.029	0.028
10^0	10^2		0.025	0.022	0.022	0.023
10^0	10^3		0.052	0.045	0.040	0.038
10^1	10^2		0.080	0.056	0.056	0.056
10^1	10^3		0.116	0.088	0.080	0.077
10^2	10^3		0.258	0.223	0.203	0.196
10^0	10^1	$T^{-1/2}$	0.033	0.030	0.029	0.028
10^0	10^2		0.022	0.019	0.019	0.019
10^0	10^3		0.023	0.021	0.019	0.019
10^1	10^2		0.080	0.053	0.053	0.054
10^1	10^3		0.082	0.059	0.055	0.053
10^2	10^3		0.232	0.201	0.186	0.182

Table 12.8 Root-mean-square error of $\widehat{\alpha}$ based on $\widehat{A}(T)-1$, using the same simulated data and counting-time ranges used in Table 12.1. This table provides results for different counting-time increments within those time ranges (left-most two columns). The counting times increase geometrically by factors of $10^{0.3}$, $10^{0.2}$, $10^{0.1}$ (used in Table 12.1), and $10^{0.05}$, thereby providing, respectively, $3\frac{1}{3}$, 5, 10, and 20 counting times per decade of the time range. The top half of the table represents equal weighting, whereas the bottom half represents weighting inversely proportional to the square root of the counting time. Comparison of the entries in the four RMSE columns shows that estimator accuracy generally improves with decreasing counting-time increments (that is, as the number of counting times per decade increases). Exceptions to this rule occur in a few cases, such as in the second row, where the fluctuations in the individual simulations result in estimates of the normalized Haar-wavelet variance that happen, on the whole, to lie further from ideal behavior for increments of $10^{0.05}$ than $10^{0.2}$. A different simulation, with a different set of 100 random seeds, would yield slightly different results. Counting times incremented by factors of $10^{0.2}$ provide nearly optimal results at half the computational burden of increments by factors of $10^{0.1}$, and at a quarter the burden of increments by factors of $10^{0.05}$. Comparing the two weighting schemes reveals that accurately weighted results equal or exceed equally weighted results, as expected. Finally, comparing rows shows that a counting-time range of $10^0 \le T \le 10^2$ proves best for equal weighting, and that increasing the range improves the accurately weighted results with few exceptions.

$$= \quad \widehat{\mathrm{E}}\big[C^2_{\mathrm{Haar},N}(2T,k)\big] \,/ 2\mathrm{E}[\mu]T. \tag{12.19}$$

Beginning with the bias, Eq. (3.39) yields

$$\widehat{\mathrm{E}}\big[C^2_{\mathrm{Haar},N}(2T,k)\big] \quad = \quad [2T(M-1)]^{-1} \sum_{k=0}^{M-2} [Z_k(T) - Z_{k+1}(T)]^2$$

$$\begin{aligned}
\mathrm{E}\Big\{\widehat{\mathrm{E}}\big[C^2_{\mathrm{Haar},N}(2T,k)\big]\Big\} &= [2T(M-1)]^{-1}\sum_{k=0}^{M-2}\mathrm{E}\Big\{[Z_k(T)-Z_{k+1}(T)]^2\Big\} \\
&= (2T)^{-1}\,\mathrm{E}\Big\{[Z_0(T)-Z_1(T)]^2\Big\} \qquad (12.20)\\
&= \mathrm{E}\big[C^2_{\mathrm{Haar},N}(T,k)\big]\,, \qquad (12.21)
\end{aligned}$$

where $M \equiv \mathrm{int}(L/T)$, L represents the duration of the data segment, and $\mathrm{int}(\cdot)$ returns the largest integer not greater than its argument. Equation (12.20) follows from the stationarity of the point process.

Estimates of the Haar-wavelet variance therefore have zero bias, and all error in its estimate, $\widehat{\mathrm{E}}\big[C^2_{\mathrm{Haar},N}(2T,k)\big]$, must derive from its variance. To obtain the variance, we make use of two further simplifications. First, we assume that wavelet-transform results are exactly independent although, as we mentioned at the beginning of this section, theoretical results show that some correlation does remain in the transformed values, albeit a small amount (Tewfik & Kim, 1992). Second, we consider large counting times T, in particular $T \gg \mathrm{E}[\tau]$, so that the counting statistics become nearly Gaussian. With these simplifications, we obtain

$$\begin{aligned}
&[4T^2(M-1)^2]\,\mathrm{Var}\Big\{\widehat{\mathrm{E}}\big[C^2_{\mathrm{Haar},N}(T,k)\big]\Big\} \\
&\quad= \sum_{k=0}^{M-2}\sum_{l=0}^{M-2}\Big[\mathrm{E}\Big\{[Z_k(T)-Z_{k+1}(T)]^2\,[Z_l(T)-Z_{l+1}(T)]^2\Big\} \\
&\qquad\qquad - \mathrm{E}\Big\{[Z_k(T)-Z_{k+1}(T)]^2\Big\}\,\mathrm{E}\Big\{[Z_l(T)-Z_{l+1}(T)]^2\Big\}\Big] \\
&[4T^2(M-1)]\,\mathrm{Var}\Big\{\widehat{\mathrm{E}}\big[C^2_{\mathrm{Haar},N}(T,k)\big]\Big\} \\
&\quad\approx \Big[\mathrm{E}\Big\{[Z_0(T)-Z_1(T)]^4\Big\} - \mathrm{E}^2\Big\{[Z_0(T)-Z_1(T)]^2\Big\}\Big] \qquad (12.22)\\
&\quad\approx \Big[3\mathrm{E}^2\Big\{[Z_0(T)-Z_1(T)]^2\Big\} - \mathrm{E}^2\Big\{[Z_0(T)-Z_1(T)]^2\Big\}\Big] \qquad (12.23)\\
&\quad= 2\mathrm{E}^2\Big\{[Z_0(T)-Z_1(T)]^2\Big\} \\
&\quad= 2T^2\,\mathrm{E}^2\big[C^2_{\mathrm{Haar},N}(2T,k)\big]\,, \qquad (12.24)
\end{aligned}$$

where Eqs. (12.22) and (12.23) derive from the whiteness and Gaussian assumptions, respectively.

Combining Eqs. (12.21) and (12.24) then gives rise to the following simple result for the square of the coefficient of variation of the Haar-wavelet variance estimate: $1/[2(M-1)]$. Incorporating the assumption that the rate estimate exhibits much smaller fluctuations than does the Haar-wavelet variance [Eq. (12.19)] leads to a squared coefficient of variation for the estimated normalized Haar-wavelet variance

that takes the form

$$\frac{\mathrm{Var}\left[\widehat{A}(T)\right]}{\mathrm{E}^2\left[\widehat{A}(T)\right]} \approx \frac{T}{2(L-T)}\,, \tag{12.25}$$

where again $M \equiv \mathrm{int}(L/T)$. We use this result to generate the dashed curves in Fig. 12.1. Taking the logarithm of $\widehat{A}(T)$ yields

$$\begin{aligned}
\ln\left[\widehat{A}(T)\right] &= \ln\left\{A(T)\left[1+\frac{\widehat{A}(T)-A(T)}{A(T)}\right]\right\} \\
&= \ln[A(T)] + \ln\left[1+\frac{\widehat{A}(T)-A(T)}{A(T)}\right] \\
&\approx \ln[A(T)] + \frac{\widehat{A}(T)-A(T)}{A(T)} \qquad (12.26) \\
\mathrm{Var}\left\{\ln\left[\widehat{A}(T)\right]\right\} &\approx \mathrm{Var}\left[\frac{\widehat{A}(T)-A(T)}{A(T)}\right] \\
&= \mathrm{Var}\left[\frac{\widehat{A}(T)}{A(T)}\right] \\
&\approx \frac{T}{2(L-T)}. \qquad (12.27)
\end{aligned}$$

Equation (12.26) makes use of the fact that $\ln(1+x) \approx x$ for $x \ll 1$, and remains valid for $T/L \ll 1$; Eq. (12.27) results from Eq. (12.25).

Equation (12.27) provides a simple but fundamental result: the variance of the logarithm of the normalized Haar-wavelet variance estimate increases linearly with counting time T. This makes good sense; since each independent wavelet-transform value has a duration proportional to T, the number of independent samples in a fixed length of data varies inversely with T. Thus, the variance of the estimate, which is inversely proportional to the number of samples, varies in direct proportion to T.

For best accuracy in estimating α, then, weighting functions should vary inversely with the square root of T. We incorporated this weighting function in the entries reported in the lower half of Table 12.8. The results are indeed superior to those obtained using equal weighting, in all cases. Furthermore, for this weighting increasing the range of times used in estimating α almost always yields an improved estimate, as expected with a proper weighting function. Employing this $T^{-1/2}$ weighting function, and counting times incremented by factors of $10^{0.2}$, we obtain the best performance by using the largest possible range of counting times ($10^0 \le T \le 10^2$); the root-mean-square error is 0.019.

We note that the foregoing analysis fails without the use of logarithms. Rearranging Eq. (12.25) yields the variance of the estimate itself, as a function of counting time T:

$$\mathrm{Var}\left[\widehat{A}(T)\right] \approx \mathrm{E}^2\left[\widehat{A}(T)\right]\frac{T}{2(L-T)}$$

$$\begin{aligned} &\approx \left[(T/T_A)^{\alpha}\right]^2 T/2L \\ &= \tfrac{1}{2} T_A^{-2\alpha} L^{-1} T^{1+2\alpha} \\ &\sim T^{1+2\alpha}. \end{aligned} \tag{12.28}$$

Thus, fitting a power-law function directly to values of $\widehat{A}(T)$ results in a variance that increases as $T^{1+2\alpha}$. This implies that the optimal weighting function for estimating α itself depends on α, which violates the spirit of estimating an unknown signal. We thus employ logarithms and linear fits, rather than direct values of $\widehat{A}(T)$ and power-law fits.

Oversampling. We conclude this section by considering oversampled versions of the normalized Haar-wavelet variance estimate. Figure 3.6b) illustrates the method we have used to estimate the normalized Haar-wavelet variance to this point; the first counting duration begins at $t = 0$, with subsequent counting durations immediately following each other in turn.

Reiterating the construction of the normalized Haar-wavelet variance we obtain

$$\begin{aligned} \widehat{A}(T) \quad &= \quad (M-1)^{-1} \sum_{k=0}^{M-2} \left[Z_k(T) - Z_{k+1}(T)\right]^2 \Big/ \left[2TN(L)/L\right] \\ \left[2T(M-1)\,N(L)/L\right] \widehat{A}(T) \\ &= \quad \sum_{k=0}^{L/T-2} \left[2N(kT+T) - N(kT) - N(kT+2T)\right]^2, \end{aligned} \tag{12.29}$$

again with $M = \mathrm{int}(L/T)$. However, the time $t = 0$ bears no particular significance to the point process $dN(t)$; beginning the counting durations at $t = \frac{1}{2}T$ rather than at $t = 0$ is equally valid. In fact, an average of the two yields improved statistics:

$$\begin{aligned} &2\left[2T(M-1)\,N(L)/L\right] \widehat{A}_{2\times}(T) \\ &\quad = \quad \sum_{k=0}^{L/T-2} \left[2N(kT+\tfrac{3}{2}T) - N(kT+\tfrac{1}{2}T) - N(kT+\tfrac{5}{2}T)\right]^2 \\ &\quad + \quad \sum_{k=0}^{L/T-2} \left[2N(kT+T) - N(kT) - N(kT+2T)\right]^2, \end{aligned} \tag{12.30}$$

where the notation $\widehat{A}_{2\times}(T)$ indicates the average of the variances estimated over two different sets of counting durations. Continuing this process, one can average results over four sets of durations that differ in their starting times by $\frac{1}{4}T$, yielding $\widehat{A}_{4\times}(T)$:

$$\begin{aligned} &4\left[2T(M-1)\,N(L)/L\right] \widehat{A}_{4\times}(T) \\ &\quad = \quad \sum_{k=0}^{L/T-2} \left[2N(kT+\tfrac{7}{4}T) - N(kT+\tfrac{3}{4}T) - N(kT+\tfrac{11}{4}T)\right]^2 \end{aligned}$$

Time Range		RMSE of $\widehat{\alpha}_{A-1}$ from $\widehat{A}(T)-1$			
$T_{\min}$	$T_{\max}$	$10^{0.2}$, $1\times$	$10^{0.2}$, $2\times$	$10^{0.2}$, $4\times$	$10^{0.05}$, $1\times$
10^0	10^1	0.030	0.027	0.025	0.028
10^0	10^2	0.019	0.018	0.017	0.019
10^0	10^3	0.021	0.018	0.017	0.019
10^1	10^2	0.053	0.046	0.044	0.054
10^1	10^3	0.059	0.048	0.046	0.053
10^2	10^3	0.201	0.164	0.154	0.182

Table 12.9 Root-mean-square error of $\widehat{\alpha}$ based on $\widehat{A}(T)-1$, using the same simulated data and time ranges as employed for Table 12.1, but with different amounts of oversampling. Weighting varies as $T^{-1/2}$ in all cases. The results for counting-time increments of $10^{0.2}$ (labeled $1\times$) and $10^{0.05}$ (labeled $1\times$) coincide with those from Table 12.8. Other columns also derive from the same data, but make use of staggered counting durations. Computing the column labeled "$10^{0.2}$, $2\times$" uses a set of counting durations that begins at $t=0$, in conjunction with a second set that begins at $t=\frac{1}{2}T$; the average over all counting durations, from both sets, forms the estimate of the normalized Haar-wavelet variance. Similarly, the column labeled "$10^{0.2}$, $4\times$" makes use of four sets of counting times, beginning at $t=0$, $t=\frac{1}{4}T$, $t=\frac{1}{2}T$, and $t=\frac{3}{4}T$. For each row, double sampling decreases the error, while quadruple sampling brings little further improvement. Doubly sampled versions at a counting-time increment of $10^{0.2}$ consistently yield results superior to those obtained using singly sampled versions at a counting-time increment of $10^{0.05}$, although the latter imposes twice the computational load.

$$
\begin{aligned}
&+ \sum_{k=0}^{L/T-2} \left[2N(kT+\tfrac{3}{2}T) - N(kT+\tfrac{1}{2}T) - N(kT+\tfrac{5}{2}T)\right]^2 \\
&+ \sum_{k=0}^{L/T-2} \left[2N(kT+\tfrac{5}{4}T) - N(kT+\tfrac{1}{4}T) - N(kT+\tfrac{9}{4}T)\right]^2 \\
&+ \sum_{k=0}^{L/T-2} \left[2N(kT+T) - N(kT) - N(kT+2T)\right]^2 . \qquad (12.31)
\end{aligned}
$$

The results presented in Table 12.9 illustrate the improvement in fractal-exponent estimation accuracy obtained by using such oversampling. Having established in Table 12.8 that a weight of $T^{-1/2}$, a counting time increment of $10^{0.2}$, and subtracting unity from $\widehat{A}(T)$ optimizes the estimation process, we display results in Table 12.9 for this choice of parameters at $2\times$ and $4\times$ oversampling. We also reproduce from Table 12.8 results for increments of $10^{0.2}$ and $10^{0.05}$ without oversampling ($1\times$). For all six choices of counting-time ranges, $2\times$ oversampling turns out to provide a reduced root-mean-square error, whereas $4\times$ oversampling further improves performance only slightly. Both oversampling schemes yield performance superior to that obtained by using a counting-time increment of $10^{0.05}$ (right-most column of

Table 12.8), despite the fact that this latter increment has a computational load equal to that of an increment of $10^{0.2}$ at $4\times$ oversampling. Holding the ratio of oversampling constant, the performance largely resembles that portrayed in Table 12.8, thereby confirming the choice of $10^{0.2}$ for the counting-time increment, both for oversampling and for the original method. Similar results also obtain for estimates that are based on $\widehat{A}(T)$ rather than $\widehat{A}(T) - 1$ (not shown).

Overall, we conclude that optimal fractal-exponent estimation mandates the following, with the choices largely independent of each other: (1) a counting-time increment of $10^{0.2}$, (2) estimates based on $\widehat{A}(T) - 1$, (3) weighting in accordance with $T^{-1/2}$, and (4) counting-time ranges of 10^0–10^2 or 10^3.

12.3.9 Rate spectrum

The rate-based spectrum $S_\lambda(f, T)$ for a point process has an estimate known as the rate-based periodogram (see Sec. 3.4.5). Like the normalized Haar-wavelet variance examined in Secs. 12.2.3 and 12.3.8, this measure has no bias. Furthermore, errors about the true value are multiplicative, independent, and have the same statistics for all frequencies f (Oppenheim & Schafer, 1975, p. 547).[7] This multiplicative character suggests that the logarithm of the periodogram renders the errors additive with similar statistics. The independence and uniformity of the statistics with frequency suggest equal weighting.

What set of frequencies provides the best results? Although any frequency spacing will yield information about the spectrum of the underlying point process, linear spacing is far superior from a practical perspective: it permits the use of fast Fourier transform techniques, thereby vastly reducing the computational load. We make use of the rate spectrum rather than the point-process spectrum since fast Fourier transforms call for discrete-time data. As shown in Eq. (3.67), the two measures bear a close resemblance to each other for frequencies much lower than the inverse of the counting time. For the resemblance to hold, the Fourier transform must have at least $2fL$ elements, where f is the largest frequency of interest, and L is the duration of the data under study.

Since Fourier transforms operate on periodic sequences, the discontinuity between the first and last counting windows can introduce spurious components into the periodogram, thereby blurring fractal features. We examine this effect in detail in Sec. A.8.1. The effect of this discontinuity decays with frequency as f^{-2}, making good estimation of processes with $\alpha > 2$ difficult. Although values of α in excess of two rarely occur, as indicated in Sec. 5.2.2, the use of a simple Hanning window (Oppenheim & Schafer, 1975, p. 242) serves to increase the theoretical limit from $\alpha = 2$ to $\alpha = 6$. Figure 12.9 displays the rate spectrum estimate obtained with a Hanning window (dashed curve) and with a rectangular window (no windowing, solid curve). The simulated data were identical to those used to produce Figs. 12.1–12.7.

[7] While this reference presents a proof only for white Gaussian noise, these results hold to an excellent approximation for the situation at hand.

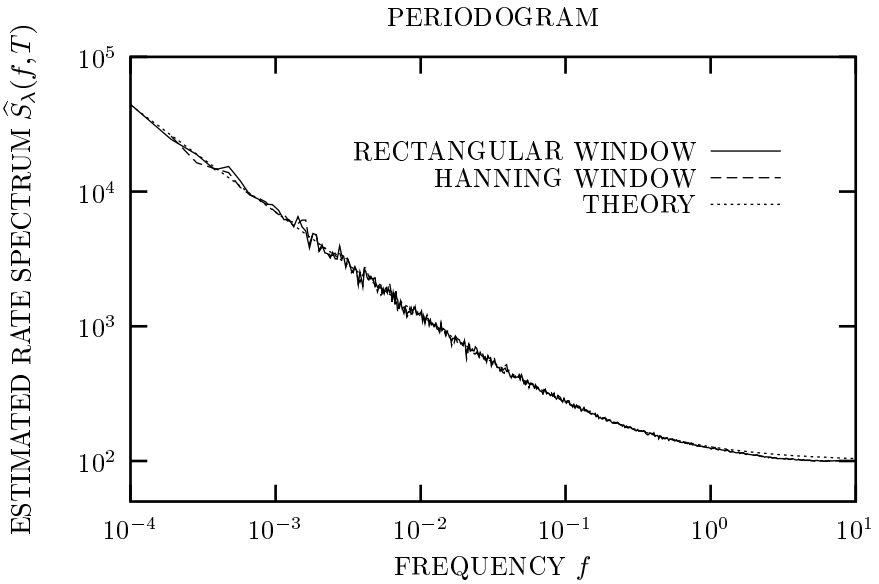

Fig. 12.9 Estimated rate spectrum (periodogram) $\widehat{S}_\lambda(f,T)$ vs. frequency f, based on the same simulations as those used to generate Figs. 12.1–12.7. For each simulation, we calculated its rate estimate using a counting time of $T = L/2^{18} \approx 0.0401$ time units (bearing in mind that L differs among the simulations). We then calculated the square magnitude of the fast Fourier transform of this rate estimate. Next, we collected values into blocks whose frequencies differed by less than a factor of 1.02, and plotted a single point comprising the mean of the frequencies and the mean of the spectral estimate. Finally, we averaged all 100 curves together, both ordinate and abscissa. We used two different windows: a Hanning (raised cosine) window, and a rectangular window equal to the duration L of the simulation (no windowing). The theoretical results directly follow Eq. (5.44a). The Hanning window yields a single large spurious value at $f = 1/L$ which we removed since it arises from the window itself. Otherwise, the results for both windows closely follow the theoretical results. The dip in the simulated curve near $f = 3$ derives from the discrete nature of the rate process used to generate the simulations (see Sec. 12.2.3).

We calculated the square magnitude of the Fourier transform of the rate estimate for each simulation, then averaged over frequencies within a factor of 1.02 of each other and, finally, averaged all 100 curves together. Except for a large peak arising from the cosine term in the Hanning window that we removed, the estimated spectra for both windows closely follow the theoretical form associated with Eq. (5.44a) (dotted curve). We conclude that windowing does not provide an advantage for these particular simulations, most likely because the fractal exponent $\alpha = 0.8$ lies below 2.

We next examined the fractal-exponent statistics based on the collection of individual simulations. In analogy with the method used for estimating the normalized

Haar-wavelet variance, we carried out a least-squares fit on the logarithm of the periodogram vs. the logarithm of frequency, over several ranges of frequencies, using equal weighting. For fits that include high frequencies ($10^3/L$ and $10^4/L$), the high-frequency asymptote $\lim_{f\to\infty} S_\lambda(f,T) = \mathrm{E}[\mu]$ causes the rate of decay to decrease with frequency, thereby leading to negative bias. Fits limited to low frequencies, on the other hand, suffer from excessive variance since fewer frequencies exist in these ranges. Overall, the best performance obtains for $1/L \le f \le 10^3/L$, but with a root-mean-square error of 0.103, it remains poor.

In an attempt to improve matters, we also performed fits to the periodogram after subtracting the estimate of the high-frequency asymptote, $\widehat{\mathrm{E}}[\mu]$. However, as a result of the variance inherent in the periodogram, this occasionally yielded nonpositive values. To rectify this, we employed averaging and thresholding. Specifically, we averaged the value of $\widehat{S}_\lambda(f,T)$ at a frequency $f = n/L$ with the $n/16$ values following it in frequency, and then subtracted the estimated asymptote $\widehat{\mathrm{E}}[\mu]$. Finally, we replaced all numbers less than $\varepsilon\widehat{\mathrm{E}}[\mu]$ with $\varepsilon\widehat{\mathrm{E}}[\mu]$, where $\varepsilon = 2^{-12} = 0.000244140625$, and performed a least-squares fit on the modified values. Indeed, removal of the high-frequency asymptote leads to a substantial improvement in the results. Larger frequency ranges generally yielded better accuracy, as was the case for the normalized Haar-wavelet variance with proper weighting. However, the smallest root-mean-square error, 0.059, occurred for the frequency range $10/L \le f \le 10^3/L$. The results of these simulations appear in Table 12.10. We note that our results are somewhat idiosyncratic, given the parameters used for averaging (16 adjacent frequency bins and $\varepsilon = 2^{-12}$); as with the interval spectrum, changing these values would likely lead to a different result.

While analytical results for the bias prove difficult to obtain in general, relatively simple results are available for the standard deviation, given a number of assumptions (Lowen & Teich, 1995). We proceed to derive this result. Starting with the sequence of counts $\{Z_k(T)\}$ obtained from a fractal-based point process, we take its Fourier transform. Each element of this transform comprises a sum of a relatively large number of terms (the counts), multiplied by trigonometric functions of magnitude not exceeding unity. This suggests the use of the central limit theorem, in which case the transform approaches a circularly symmetric complex Gaussian process. We assume that the Gaussian limit holds exactly. Converting to a spectrum estimate yields

$$\widehat{S}(n/L) = \mathrm{E}[\mu]\left[1 + (n/f_S L)^{-\alpha}\right]\exp(\epsilon_n)\,, \tag{12.32}$$

where n denotes the discrete index in the fast Fourier transform and ϵ_n represents some small error term near zero. For estimates of white noise, the errors for different indices n are independent and have a standard deviation equal to the mean (Oppenheim & Schafer, 1975); we assume that this holds for the fractal-rate process under study as well. We then have

$$\begin{aligned} \mathrm{E}[\exp(\epsilon_n)] &= 1 \\ \mathrm{E}\{[\exp(\epsilon_n) - 1]^2\} &= 1 \\ \mathrm{E}\{[\exp(\epsilon_n) - 1]\,[\exp(\epsilon_m) - 1]\} &= 0, \quad n \ne m. \end{aligned} \tag{12.33}$$

Freq. Range		$\widehat{\alpha}_S$ from $\widehat{S}_\lambda(f,T)$			$\widehat{\alpha}_{S-\mathrm{E}[\mu]}$ from $\widehat{S}_\lambda(f,T) - \widehat{\mathrm{E}}[\mu]$			Calc.
$f_{\min}$	$f_{\max}$	Bias	SD	RMSE	Bias	SD	RMSE	SD
10^0	10^1	-0.026	0.530	0.531	0.023	0.607	0.608	0.578
10^0	10^2	-0.003	0.130	0.130	-0.151	0.142	0.207	0.147
10^0	10^3	-0.096	0.036	0.103	-0.055	0.041	0.069	0.045
10^0	10^4	-0.322	0.014	0.322	0.076	0.028	0.081	0.014
10^1	10^2	0.007	0.196	0.196	-0.172	0.224	0.283	0.222
10^1	10^3	-0.108	0.040	0.115	-0.044	0.041	0.059	0.050
10^1	10^4	-0.330	0.015	0.331	0.081	0.029	0.086	0.014
10^2	10^3	-0.149	0.071	0.165	-0.016	0.064	0.066	0.077
10^2	10^4	-0.361	0.016	0.362	0.101	0.033	0.106	0.016
10^3	10^4	-0.439	0.022	0.439	0.172	0.060	0.182	0.025

Table 12.10 Performance of fractal-exponent estimates $\widehat{\alpha}$ for 100 different simulations of a fractal-Gaussian-process-driven Poisson process obtained using estimates of the rate spectrum (periodogram) $\widehat{S}_\lambda(f,T)$. The simulations were the same as those employed to produce Table 12.1. We counted the number of events falling into 2^{15} windows of equal length $T = L/2^{15} \approx 0.320$ (bearing in mind that L differs among the simulations), calculated the Fourier transform, and took the square magnitude. The first two columns specify the range of frequencies employed for calculating $\widehat{S}_\lambda(f,T)$. We express frequency in terms of its product with the duration of the simulation, L; thus, the first row corresponds to the frequency range $1/L \le f \le 10/L$. The next three columns represent the bias, standard deviation (SD), and resulting root-mean-square error (RMSE) for calculations employing the rate spectrum $\widehat{S}_\lambda(f,T)$. We next sought to remove the effects of the high-frequency asymptote $\widehat{\mathrm{E}}[\mu]$. Since the rate spectrum fluctuates about this asymptote, many of the resulting differences would lie below zero, making calculation of the logarithm meaningless. We therefore employed averaging and thresholding. Specifically, we averaged the value of $\widehat{S}_\lambda(f,T)$ at a frequency $f = n/L$ with the $n/16$ values following it in frequency, and then subtracted the estimated asymptote $\widehat{\mathrm{E}}[\mu]$. Finally, to ensure positive values, we replaced all numbers less than $\varepsilon\widehat{\mathrm{E}}[\mu]$ with $\varepsilon\widehat{\mathrm{E}}[\mu]$, where $\varepsilon = 2^{-12}$. We estimated the fractal exponent employing these modified values; the corresponding results appear in columns 6–8. Subtracting the high-frequency asymptote indeed reduces the bias significantly, returning reduced root-mean-square errors. Moreover, increasing the range generally improves the results, as we observed in the properly weighted outcomes for the normalized Haar-wavelet variance presented in the lower half of Table 12.8. The rightmost column presents the predicted theoretical values for the standard deviation; results for the corresponding simulations stand two columns to its left. The best results (minimum root-mean-square error) for $\widehat{S}_\lambda(f)$ obtain by using $1/L \le f \le 10^3/L$, whereas for $\widehat{S}_\lambda(f) - \widehat{\mathrm{E}}[\mu]$ the optimal range is $10/L \le f \le 10^3/L$. While the normalized Haar-wavelet variance $\widehat{A}(T)$ yields superior results, it does incur a larger computational burden.

Proceeding with the analysis, we estimate α from $\widehat{S}(n/L)$. Subtracting the constant term $\mathrm{E}[\mu]$ from Eq. (12.32), and taking logarithms of both independent and dependent variables, we define

$$\begin{aligned} x_n &\equiv \ln(n) \\ y_n &\equiv \ln\bigl\{\widehat{S}(n/L) - \mathrm{E}[\mu]\bigr\} \\ &= \ln\left(\mathrm{E}[\mu]\right) + \alpha\ln(f_S L) - \alpha\ln(n) + \epsilon_n. \end{aligned} \tag{12.34}$$

The estimate of α simply becomes the covariance between $\{x\}$ and $\{y\}$ divided by the variance of $\{x\}$ (Lowen & Teich, 1993a):

$$\begin{aligned} \widehat{\alpha}_S &= \frac{(M-1)^{-1}\displaystyle\sum_{n=1}^{M} x_n\, y_n - M^{-1}(M-1)^{-1}\left(\sum_{n=1}^{M} x_n\right)\left(\sum_{m=1}^{M} y_m\right)}{(M-1)^{-1}\displaystyle\sum_{n=1}^{M} x_n^2 - M^{-1}(M-1)^{-1}\left(\sum_{n=1}^{M} x_n\right)^2} \\ &= \alpha + \frac{\left(\displaystyle\sum_{n=1}^{M}\epsilon_n\right)\left(\sum_{m=1}^{M}\ln(m)\right) - M\displaystyle\sum_{n=1}^{M}\epsilon_n\ln(n)}{M\displaystyle\sum_{n=1}^{M}\ln^2(n) - \left(\sum_{n=1}^{M}\ln(k)\right)^2}, \end{aligned} \tag{12.35}$$

with $M = \mathrm{int}(L/T)$, the number of counts as defined in Eq. (12.4). Equation (12.35) has a mean of α by construction, as it should. After some algebra, an expression for the variance emerges:

$$\mathrm{Var}[\widehat{\alpha}_S] = \mathrm{Var}[\epsilon]\cdot\left[\sum_{n=1}^{M}\ln^2(n) - M^{-1}\left(\sum_{n=1}^{M}\ln(n)\right)^2\right]^{-1}. \tag{12.36}$$

The variance of ϵ depends on the distribution of the errors $\{\epsilon_n\}$ themselves, while the rest of Eq. (12.36), which is a deterministic function of M, approaches $1/M$ as $M \to \infty$. For the circularly symmetric Gaussian form assumed for the Fourier transform of $\{Z_k(T)\}$, the spectrum estimate $\widehat{S}(n/L)$ has an exponential distribution, which provides

$$\begin{aligned} \mathrm{Var}[\epsilon] &= \int_0^\infty \ln^2(t)\,\exp(-t)\,dt \\ &= \tfrac{1}{6}\pi^2 + C_{\mathrm{Euler}}^2 \\ &\doteq 1.978, \end{aligned} \tag{12.37}$$

where $C_{\mathrm{Euler}} \doteq 0.5772156649$ denotes Euler's constant.

The right-most column in Table 12.10 presents the expected theoretical values for the standard deviation, calculated from Eqs. (12.36) and (12.37). Here we substitute

$f_{\max}L$ and $f_{\min}L$ for the upper limit (M) and lower limit (unity), respectively, in the sums in Eq. (12.36), and replace M^{-1} with $f_{\max}L + 1 - f_{\min}L$. Results for the corresponding simulations for each frequency range stand two columns to the left. While none of the assumptions employed in the derivation provided above holds exactly, the calculated and simulated values nevertheless agree quite well, especially for more constrained frequency ranges ($f_{\max}L < 10^4$).

12.4 COMPARISON OF MEASURES

Based on the results provided in Sec. 12.3, we conclude that the normalized Haar-wavelet variance and the rate spectrum prove to be the most useful measures for estimating the exponent of an unidentified fractal-based point process. Furthermore, as shown in Sec. 5.2, these two measures are valid over the full range of fractal exponents normally encountered, $0 < \alpha < 2$.

Detrended fluctuation and the interval-based wavelet variance also perform well as estimators for the point process at hand, and the optimization techniques we used for the normalized Haar-wavelet variance would presumably provide comparable improvements for these two measures as well. Nevertheless, as with all interval-based approaches, they cannot reliably detect fractal behavior, let alone quantify it, so we do not recommend their general use.

The normalized Haar-wavelet variance returns the best performance, with a root-mean-square error of 0.018, in comparison with 0.059 for the rate spectrum, under optimal conditions for the estimation problem at hand (appropriate weighting, choice of counting time increment, oversampling, and subtraction of short time and high-frequency asymptotes). However, the former measure has significantly more computational burden than the latter. The ultimate choice of measure thus appears to depend on the relative costs of estimation error and processing time.

A number of other approaches exist for estimating fractal exponents (see, for example, Beran, 1992, 1994; Taqqu et al., 1995), but these expect real-valued, discrete-time sequences, generally with Gaussian statistics. Although counting turns a point process into a positive-integer-valued discrete-time sequence, the question then becomes the choice of counting time. Short counting times yield statistics far from a Gaussian form, whereas long counting times fail to capture the short-time-scale information inherent in a point process; this issue bears some similarity to the choice of range for a dependent variable, mentioned at the beginning of this chapter. In general, discrete-time methods do not appear to offer the performance and robustness desired in a fractal-exponent estimator for point processes. However, two notable exceptions exist. First, for extremely long point-process realizations, the very large numbers of events renders the choice of counting time far less critical, so that conversion of a data set to a discrete-time form becomes relatively straightforward. Second, for calculation of the rate-based spectral estimate, we employ an effective counting time L/M, with L again the duration of the realization and M the size of the (fast) Fourier transform used. In this case, the choice of counting time (or, equivalently, M) proves

relatively unimportant as long as the counting time does not exceed the Nyquist limit for frequencies of interest.

Other measures considered in this Chapter have made use of counting sequences as intermediate steps, but have avoided the limitations associated with choosing a fixed value of counting time. Little advantage appears to accrue from using traditional fractal methods based on discrete-time sequences, which confirms our conclusion that the normalized Haar-wavelet variance and the rate spectrum are the measures of choice.

We offer a final caveat in closing this chapter. We chose the simulation set analyzed because of its generic nature. The fractal-Gaussian-process-driven Poisson process characterizes a number of observed point processes; furthermore, superpositions of many fractal-based point processes converge to it. The value $\alpha = 0.8$ offers a significant range of variation of the various measures with time and frequency, yet it lies sufficiently below unity to preclude problems with measures that fail for $\alpha \geq 1$.[8] However, just as *a priori* information heavily influences the identification of a point process, as described in Sec. 12.1, so too does prior information affect the choice of an optimal measurement statistic.

While we expect that the results established here will prove useful in many circumstances, the *a priori* information available in all possible experiments spans far too great a range for one approach to yield optimal results in all cases. Different applications will surely involve different ranges of α; the estimation of other fractal-based parameters, such as fractal onset times and frequencies, are certainly of interest, and more subtle features of various measures will come to the fore. Rather than attempting to catalog such a vast parameter space, we have chosen instead to direct our presentation to investigating how well various statistical measures function for estimating the fractal exponent of an unidentified fractal-based point processes, and, most importantly, to setting forth a collection of techniques and mathematical relations that should find use in a broad variety of applications.

Problems

12.1 *Discriminating fractal-rate point processes via their interevent-interval densities* Suppose we have a realization of a fractal-rate point process, and we know that it derives from a Poisson process driven by either: (1) a fractal Gaussian process, (2) fractal binomial noise, or (3) fractal shot noise. Suppose further that processes (2) and (3) approach process (1) fairly closely, but not exactly, of course. The measures used in this chapter to quantify fractal behavior cannot distinguish among these three possibilities since all three processes generate power-law forms akin to those provided in Eq. (5.44) (see Sec. 12.1). Discuss how the interevent-interval probability density (or distribution) might assist us in determining which of the three rate processes is

[8] We previously examined the normalized Haar-wavelet variance and spectrum for this process, as well as for the fractal-Gaussian-process-driven integrate-and-reset process (with jitter), for $\alpha = 0.2$, 0.8, and 1.5 (Thurner et al., 1997).

responsible for generating the realization at hand. How does the situation differ if the three rate processes drive an integrate-and-reset kernel instead?

12.2 *Robustness/error tradeoff in estimation* In Sec. 12.2 we pointed out that a tradeoff exists between robustness and error. For nearly all estimation tasks, some estimators provide excellent results for a restricted class of data, whereas others yield useful results for a wider range of data at the cost of somewhat greater overall error. Discuss this issue in the context of estimating the fractal exponents for a fractal renewal process and for a fractal-shot-noise-driven Poisson process.

12.3 *Bias/variance tradeoff in estimation* In addition to the tradeoff between robustness and error discussed in Prob. 12.2, a tradeoff also exists between bias and variance in many estimators. Demonstrate this tradeoff explicitly for the estimation of $\widehat{\alpha}$ from the normalized Haar-wavelet variance $\widehat{A}(T)$ by computing the correlation coefficient between the absolute values in the third and fourth columns of Table 12.1, which represent the bias and standard deviation, respectively. Repeat this calculation for the estimation of $\widehat{\alpha}$ from $\widehat{A}(T) - 1$ by making use of the sixth and seventh columns of Table 12.1. What might account for the difference in the correlation coefficients obtained in the two cases?

12.4 *Coincidence-rate and spectrum estimation* Discuss the problems involved in attempting to construct an estimate of the coincidence rate $G(t)$, and show that addressing these problems leads to the count-based autocorrelation $R_Z(k, T)$ or a similar measure. Also show that the point-process spectrum $S_N(f)$ does not suffer from the same shortcoming. Why then do we employ the rate spectrum $S_\lambda(f, T)$ instead of $S_N(f)$?

12.5 *Bias in normalized-variance estimates* For a fractal-based point process with $0 < \alpha < 1$, Eq. (12.6) indicates that the normalized variance will decrease for counting times T near the duration of the recording L.

12.5.1. Use this equation to find the counting time at which the normalized variance achieves a maximum, and express this counting time in terms of the fractal onset time T_F, the duration of the recording L, and the fractal exponent α.

12.5.2. Simulate a fractal-based point process, say the fractal-Gaussian-process-driven Poisson process, and calculate the normalized variance for the largest value of T possible. Plot the results of the simulations (include both mean and mean ± 1-standard deviation values) as well as the predictions of Eqs. (5.44b) and (12.6), choosing a plotting format that highlights the differences between the two predictions.

12.6 *Effect of averaging on spectral estimates* To obtain useful nonparametric estimates of the spectrum (rather than estimates of the fractal exponent α that we have been heretofore pursuing), it is common to average n adjacent values of the estimated spectrum, with n a large number that does not vary with frequency. One can also divide the data into blocks, compute spectrum estimates of each block separately, and then average these estimates across blocks. Explain why we do not use this approach in estimating the fractal exponent.

12.7 *Asymptote subtraction for improved fractal-exponent estimation* All of the count-based measures employed in this chapter attain a constant value in the high-frequency/short-time limit. For example, Eq. (3.59) shows that the point-process spectrum approaches the mean event rate at high frequencies, whereas Eq. (3.42) indicates that both the normalized variance and the normalized Haar-wavelet variance attain a value of unity for short counting times. It is tempting to subtract these limits from the associated measures to extend the useful range of fractal scaling and to thereby improve the performance of the fractal-exponent estimators. We have, in fact, done just that for the periodogram in Sec. 12.3.9, and for the normalized Haar-wavelet variance in Secs. 12.2.3 and 12.3.8. Cite two reasons for caution in applying this approach to real data sets.

12.8 *Fractal behavior in a simulated fractal renewal process* Simulate a number of runs of a fractal renewal process (see Chapter 7) and an equal number of runs of a homogeneous Poisson process (see Sec. 4.1), with the same mean interevent interval (choose $\mathrm{E}[\tau] = \mathrm{E}[\mu] = 1$ for simplicity). Select $\gamma = \frac{3}{2}$, which leads to $\alpha = \frac{1}{2}$ by virtue of Eq. (7.9). This value of γ lies within the range $1 < \gamma < 2$, which limits variation, as described in Sec. 7.1.3. Choose abrupt cutoffs with $B/A = 10^6$ to ensure the presence of fractal behavior over a wide range of times and frequencies; this yields $A = 1.001001/3 = 0.333667$, $B = 0.333667 \times 10^6$, and $L = 10^6$, which gives an expected number of events $\mathrm{E}[N(L)] = 10^6$ [see Eq. (7.2)]. Carry out 100 simulations of this process and average the results to obtain accurate statistics.[9]

Show that the rate periodogram (spectrum estimate) and the normalized Haar-wavelet variance accurately characterize fractal behavior in the fractal renewal process and properly reveal its absence in the homogeneous Poisson process. Also show that the rescaled range statistic, an interval-based measure, cannot distinguish between these two renewal processes and therefore does not reliably detect fractal behavior in general.

[9] Producing a smooth version of this curve that accurately follows its expected value requires inordinate simulation resources (see Prob. 7.6).