# REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188



**1. AGENCY USE ONLY** *(Leave blank)*

**2. REPORT DATE**
12/31/97

**3. REPORT TYPE AND DATES COVERED**
final report 1/1/97-12/31/97

**4. TITLE AND SUBTITLE**
Analysis, Synthesis, and Estimation of Fractal-Rate Stochastic Point Process

**5. FUNDING NUMBERS**
G "N00014-97-1-0059"

**6. AUTHOR(S)**
Malvin C. Teich

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Department of Electrical and Computer Engineering
Boston University
Boston, MA 02215

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5660

**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**

19980311 033

**11. SUPPLEMENTARY NOTES**

**12a. DISTRIBUTION/AVAILABILITY STATEMENT**
APPROVED FOR PUBLIC RELEASE

**12b. DISTRIBUTION CODE**

**13. ABSTRACT** *(Maximum 200 words)*

Fractal and fractal-rate stochastic point processes (FSPPs and FRSPPs) provide useful models for describing a broad range of diverse phenomena, including electron transport in amorphous semiconductors, computer-network traffic, and sequences of neuronal action potentials. A particularly useful statistic of these processes is the fractal exponent $\alpha$, which may be estimated for any FSPP or FRSPP by using a variety of statistical methods. Simulated FSPPs and FRSPPs consistently exhibit bias in this fractal exponent, however, rendering the study and analysis of these processes non-trivial. We have examined the synthesis and estimation of FRSPPs by carrying out a systematic series of simulations for several different types of FRSPP over a range of design values for $\alpha$. The discrepancy between the desired and achieved values of $\alpha$ is shown to arise from finite data size and from the character of the point-process generation mechanism. In the context of point-process simulation, reduction of this discrepancy requires generating data sets with either a large number of points, or with low jitter in the generation of the points. In the context of fractal data analysis, the results presented here suggest caution when interpreting fractal exponents estimated from experimental data sets.



**14. SUBJECT TERMS**
Fractal, Estimation, Fractal-rate, Stochastic

**15. NUMBER OF PAGES**
7

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| unclassified | unclassified | unclassified | unclassified |

# FINAL REPORT

# ONR GRANT #N00014-97-1-0059

Analysis, Synthesis, and Estimation of
Fractal-Rate Stochastic Point Processes

MALVIN C. TEICH

*Department of Electrical & Computer Engineering*
*Boston University*
*Boston, MA 02215*

December 31, 1997

## Abstract

Fractal and fractal-rate stochastic point processes (FSPPs and FRSPPs) provide useful models for describing a broad range of diverse phenomena, including electron transport in amorphous semiconductors, computer-network traffic, noise in CCD images, and sequences of neuronal action potentials. A particularly useful statistic of these processes is the fractal exponent $\alpha$, which may be estimated for any FSPP or FRSPP by using a variety of statistical methods. Simulated FSPPs and FRSPPs consistently exhibit bias in this fractal exponent, however, rendering the study and analysis of these processes non-trivial. We have examined the synthesis and estimation of FRSPPs by carrying out a systematic series of simulations for several different types of FRSPP over a range of design values for $\alpha$. The discrepancy between the desired and achieved values of $\alpha$ has been found to arise from finite data size and from the character of the point-process generation mechanism. In the context of point-process simulation, reduction of this discrepancy requires generating data sets with either a large number of points, or with low jitter in the generation of the points. In the context of fractal data analysis, the results presented here suggest caution when interpreting fractal exponents estimated from experimental data sets.

# 1 INTRODUCTION

Some random phenomena occur at discrete times or locations, with the individual events largely identical. A stochastic point process [1] is a mathematical construction which represents these events as random points in a space. Fractal stochastic point processes exhibit scaling in all statistics generally considered; fractal-rate stochastic point processes do so only in some of them. We have considered the simulation and estimation of fractal and fractal-rate stochastic point processes on a line, which model a variety of observed phenomena in the physical and biological sciences. It comprises an effort we first began in 1993 [2], reported on in 1995 [3], and concluded in 1997 [4].

# 2 ESTIMATION OF THE ALLAN FACTOR AND POWER SPECTRAL DENSITY

Given a segment of an FRSPP, we seek an estimate $\widehat{\alpha}$ of the true fractal exponent $\alpha$ of the entire process from which the segment was derived. Several effects contribute to estimation error for finite-length segments, regardless of the method used. The fractal exponent provides a measure of the relative strengths of fluctuations over various time scales; for an FRSPP with a relatively large fractal exponent, for example, relatively more energy is concentrated in longer time scale fluctuations than for an FRSPP with a smaller fractal exponent. Variance stems from the inherent randomness of the strengths of these fluctuations. A collection of finite realizations of an FRSPP with the same parameters will exhibit fractal fluctuations of varying strengths, leading to a distribution of estimated fractal exponents.

Bias appears to arise from cutoffs in the time (frequency) domain which give rise to oscillations in the frequency (time) domain, confounding the pure power-law behavior of the fractal periodogram (PG) and Allan factor (AF). In addition, the physical limitations of the measurement process itself impose practical limits on the range of time scales available. Although algorithms exist for accurately compensating for these cutoffs, they presuppose a detailed knowledge of the process *a priori*, which is not in the spirit of estimating an unknown signal. Consequently, we have not attempted to compensate for these cutoffs in this manner. In 1995, we determined [3] the theoretical expected Fano factor (FF) values for an FRSPP with finite duration, and the corresponding expected bias for the corresponding fractal exponent. The FF is the ratio of the count variance to the count mean, as a function of the counting time $T$. Bias for the power spectral density was also considered, but not for a finite data length.

The AF and the PG highlight the fractal behavior of FRSPPs particularly well, and thus prove to be most useful for estimating fractal exponents. We have investigated the bias in these measures, and the effects of this bias on fractal-exponent estimation. The variance of PG-based estimators was also examined in 1995 [3]; that for the AF appears not to be readily amenable to analytical study.

## 2.1 Effects of Finite Data Length on the AF Estimate

Unlike the FF, estimates of the AF do not suffer from bias due to finite data length; thus fractal-exponent estimates based on the AF do not either. Given a particular data set, the estimated AF $\widehat{A}(T)$ at a particular counting time $T$ is given by

$$\widehat{A}(T) = (N-1)^{-1} \sum_{k=0}^{N-2} (Z_{k+1} - Z_k)^2 \Big/ 2N^{-1} \sum_{k=0}^{N-1} Z_k, \tag{1}$$

where $N$ is the number of samples, $\{Z_k\}$ represents the sequence of counts, and the functional dependence of $Z_k$ upon $T$ is suppressed for notational clarity.

This estimate of the AF is simply the estimate of the Allan variance divided by twice the estimate of the mean; however, computing the expected value of this estimate is not straightforward. We have therefore employed the true mean rather than its estimate in Eq. (1). This does not appreciably affect the result, since the error so introduced remains a constant factor for all counting times, and so cancels in power-law slope calculations where logarithms are used. In any case, the estimate of the Allan variance exhibits far larger variations than the estimate of the mean, so the fluctuations in the AF estimate are dominated by the former. Therefore, the expected value of the AF estimate becomes

$$\begin{aligned} \mathrm{E}\left[\widehat{A}(T)\right] &= \mathrm{E}\left\{\widehat{\mathrm{E}}\left[(Z_k - Z_{k+1})^2\right] \Big/ 2\widehat{\mathrm{E}}[Z_k]\right\} \\ &\approx \mathrm{E}\left[(Z_k - Z_{k+1})^2\right] \Big/ 2\mathrm{E}[Z_k] \\ &= \mathrm{E}\left[(Z_0 - Z_1)^2\right] \Big/ 2\mathrm{E}[Z_0] \\ &= (\lambda T)^{-1} \mathrm{E}\left[Z_0^2 - Z_0 Z_1\right], \end{aligned} \tag{2}$$

which is independent of the number of samples, $N$, and therefore of the duration of the recording $L$. In the limit $L \to \infty$, the expected estimated AF for an FRSPP with $0 < \alpha < 3$ follows the generally expected form for an ergodic process. Since Eq. (2) does not depend on $L$, the quantity $\mathrm{E}\left[\widehat{A}(T)\right]$ must assume a constant value independent of $L$; thus the AF does not exhibit bias caused by finite data duration. Finally, since the expected value of the AF estimate is unbiased, the expected value of the estimate of $\alpha$ itself is expected to have negligible bias. This is a simple and important result.

Estimates of the ordinary variance and Fano factor, in contrast, do depend on the duration. In this case we have

$$
\begin{aligned}
\widehat{\mathrm{Var}}[Z] &= (N-1)^{-1} \sum_{k=0}^{N-1} (Z_k - \widehat{\mathrm{E}}[Z])^2 \\
&= (N-1)^{-1} \sum_{k=0}^{N-1} \left( Z_k - N^{-1} \sum_{l=0}^{N-1} Z_l \right)^2 \\
&= (N-1)^{-1} \left[ \sum_{k=0}^{N-1} Z_k^2 - 2N^{-1} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} Z_k Z_l + N^{-2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \sum_{m=0}^{N-1} Z_l Z_m \right] \\
&= (N-1)^{-1} \sum_{k=0}^{N-1} Z_k^2 - N^{-1}(N-1)^{-1} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} Z_k Z_l \\
&= N^{-1} \sum_{k=0}^{N-1} Z_k^2 - N^{-1}(N-1)^{-1} \sum_{k}^{N-1} \sum_{l \neq k} Z_k Z_l. \qquad (3)
\end{aligned}
$$

These cross terms, which do depend on the number of samples, lead to an estimated Fano factor with a confounding linear term

$$\mathrm{E}\left[\widehat{F}(T)\right] \approx 1 + (T/T_0)^{\alpha} - T \Big/ \left(T_0^{\alpha} L^{1-\alpha}\right) \qquad (4)$$

for $0 < \alpha < 1$ [3]. The last term on the right-hand-side of Eq. (4) leads to bias in estimating the fractal exponent; for this and other reasons, we do not employ the FF in fractal-exponent estimation.

## 2.2 Effects of Finite Data Length on the PSD Estimate

Computation of the periodogram (estimated power spectral density, PSD) proves tractable only when obtained from the series of counts $\{Z_k\}$, rather than from the entire point process $N(t)$. This, and more importantly the finite length of the data set, introduce a bias in the estimated fractal exponent as shown below.

We begin by obtaining the discrete-time Fourier transform of the series of counts $\{Z_k\}$, where $0 \leq k < M$:

$$\tilde{Z}_n \equiv \sum_{k=0}^{M-1} Z_k e^{-j2\pi kn/M}. \qquad (5)$$

The periodogram then becomes

$$
\begin{aligned}
\widehat{S}_Z(n) &= M^{-1} \left|\tilde{Z}_n\right|^2 \\
&= M^{-1} \sum_{k=0}^{M-1} \sum_{m=0}^{M-1} Z_k Z_m e^{j2\pi(k-m)n/M}, \qquad (6)
\end{aligned}
$$

with an expected value

$$\begin{aligned}
\mathrm{E}\left[\widehat{S}_Z(n)\right] &= M^{-1}\sum_{k=0}^{M-1}\sum_{m=0}^{M-1} e^{j2\pi(k-m)n/M}\mathrm{E}\left[Z_k Z_m\right] \\
&= M^{-1}\sum_{k=0}^{M-1}\sum_{m=0}^{M-1} e^{j2\pi(k-m)n/M}\mathrm{E}\left[\int_{s=0}^{T}\int_{t=0}^{T} dN(s+kT)dN(t+mT)\right] \\
&= M^{-1}\sum_{k=0}^{M-1}\sum_{m=0}^{M-1} e^{j2\pi(k-m)n/M}\int_{s=0}^{T}\int_{t=0}^{T} G_N[s-t+(k-m)T]\,ds\,dt \\
&= M^{-1}\sum_{k=0}^{M-1}\sum_{m=0}^{M-1} e^{j2\pi(k-m)n/M}\int_{u=-T}^{T}\int_{v=|u|}^{2T-|u|} G_N[u+(k-m)T]\,\frac{du\,dv}{2} \\
&= M^{-1}\sum_{k=0}^{M-1}\sum_{m=0}^{M-1} e^{j2\pi(k-m)n/M}\int_{u=-T}^{T}(T-|u|)\,G_N[u+(k-m)T]\,du \\
&= M^{-1}\sum_{k=0}^{M-1}\sum_{m=0}^{M-1} e^{j2\pi(k-m)n/M}\int_{u=-T}^{T}(T-|u|) \\
&\qquad\times\int_{\omega=-\infty}^{\infty} S_N(\omega)\,e^{j\omega[u+(k-m)T]}\,\frac{d\omega}{2\pi}\,du \\
&= (2\pi M)^{-1}\int_{\omega=-\infty}^{\infty} S_N(\omega)\left|\sum_{k=0}^{M-1} e^{jk(2\pi n/M+\omega T)}\right|^2\int_{u=-T}^{T}(T-|u|)\,e^{j\omega u}\,du\,d\omega \\
&= (2\pi M)^{-1}\int_{\omega=-\infty}^{\infty} S_N(\omega)\frac{\sin^2(\pi n+M\omega T/2)}{\sin^2(\pi n/M+\omega T/2)}\frac{4\sin^2(\omega T/2)}{\omega^2}\,d\omega \\
&= \frac{T}{\pi M}\int_{-\infty}^{\infty} S_N(2x/T)\frac{\sin^2(Mx)\sin^2(x)}{x^2\sin^2(x+\pi n/M)}\,dx. \qquad (7)
\end{aligned}$$

For a general fractal stochastic point process, where the PSD follows its usual power-law form, we therefore have

$$\mathrm{E}\left[\widehat{S}_Z(n)\right] = \frac{\lambda T}{\pi M}\int_{-\infty}^{\infty}\left[1+(\omega_0 T/2)^{\alpha}|x|^{-\alpha}\right]\frac{\sin^2(Mx)\sin^2(x)}{x^2\sin^2(x+\pi n/M)}\,dx. \qquad (8)$$

Focusing on the smaller values of $n$ useful in estimating fractal exponents permits the use of two approximations in Eq. (8). For the values of $\alpha$ of interest ($0<\alpha<2$), the integrand of Eq. (8) will only be significant near $x=-\pi n/M$, yielding

$$\begin{aligned}
\mathrm{E}\left[\widehat{S}_Z(n)\right] &\approx \frac{\lambda T}{\pi M}\int_{-\infty}^{\infty}\left[1+(\omega_0 T/2)^{\alpha}|x|^{-\alpha}\right]\frac{M\pi\delta(x+\pi n/M)\sin^2(x)}{x^2}\,dx \\
&= \lambda T\left[1+(2\pi n/\omega_0 MT)^{-\alpha}\right]\frac{\sin^2(\pi n/M)}{(\pi n/M)^2} \\
&\approx \lambda T\left[1+(2\pi n/\omega_0 MT)^{-\alpha}\right], \qquad (9)
\end{aligned}$$

which is of the expected form. Improvement of this estimation procedure appears to require numerical integration of Eq. (8), which proves nontrivial since the integrand exhibits oscillations with a small period $\pi/M$. Fortuitously, for the parameter values of general interest in our considerations the integrand appears peaked near $x = -\pi n/M$, so that not too many ($\approx 200$) oscillations need be included in the calculations. Indeed, numerical results employing this method, and with $\omega_0 \to 0$ for which Eq. (8) is known to assume the simple value $\lambda T$, agree within 0.1%. These results form a continuation of earlier efforts [2, 3] which ignored the effects of imposing periodic boundary conditions on the Fourier transform, and of binning the events.

Other methods exist for compensating for finite data length in power spectral estimation, such as maximizing the entropy subject to the correlation function over the available range of times (see, e.g., [5]).

# 3 SUMMARY

We have investigated the properties of fractal and fractal-rate stochastic point processes (FSPPs and FRSPPs), focusing on the estimation of measures that reveal their fractal exponents. The fractal-Gaussian-noise driven integrate-and-fire process (FGNIF) is unique as a point process in that it exhibits no short-term randomness; we have developed a generalization which includes jitter, the FGNJIF, for which the short-term randomness is fully adjustable. In addition to the randomness contributed by the fractal nature of the rate, all other FRSPPs appear to exhibit additional randomness associated with point generation, which confounds the accuracy of fractal-exponent estimation. The FGNJIF proved crucial in elucidating the role that such randomness plays in the estimation of fractal exponents in other FRSPPs. We obtained analytical results for the expected biases of the PG- and AF-based fractional exponent estimators due to finite data length. We followed this theoretical work with a series of simulations of FRSPPs for three representative fractal exponents: $\alpha = 0.2$, 0.8, and 1.5 [4]. Using these simulations together with the analytical predictions, we delineated the sources of error involved in estimating fractal exponents. In particular, using the FGNJIF we were able to separate those factors intrinsic to estimation over finite FRSPP data sets from those due to the particular form of FRSPP involved. We conclude that the AF-based estimate proves more reliable in estimating the fractal exponent than the PG-based method, yielding rms errors of 0.06 for data segments of FRSPPs with $10^6$ points over all values of $\alpha$ examined. Finally, we note that wavelet generalizations of the AF appear to yield comparable results [6, 7], suggesting that the AF estimate may be optimal in some applications.

# References


[1] D. R. Cox and V. Isham, *Point Processes* (Chapman and Hall, London, 1980).

[2] S. B. Lowen and M. C. Teich, "Estimating the dimension of a fractal point processes," *Proc. SPIE* **2036**, 64–76 (1993).

[3] S. B. Lowen and M. C. Teich, "Estimation and simulation of fractal stochastic point processes," *Fractals* **3**, 183-210 (1995).

[4] S. Thurner, S. B. Lowen, M. C. Feurstein, C. Heneghan, H. G. Feichtinger, and M. C. Teich, "Analysis, synthesis, and estimation of fractal-rate stochastic point processes," *Fractals* **5**, 565-595 (1997).

[5] J. Skilling, ed., *Maximum Entropy and Bayesian Methods* (Kluwer, Boston, 1988).

[6] M. C. Teich, C. Heneghan, S. B. Lowen, and R. G. Turcott, "Estimating the fractal exponent of point processes in biological systems using wavelet- and Fourier-transform methods," in *Wavelets in Medicine and Biology*, eds. A. Aldroubi and M. Unser (CRC Press, Boca Raton, FL, 1996), pp. 383–412.

[7] C. Heneghan, S. B. Lowen, and M. C. Teich, "Wavelet analysis for estimating the fractal properties of neural firing patterns," in *Computational Neuroscience*, ed. J. M. Bower, (Academic Press, San Diego, 1996), pp. 441–446.