

Probability Distributions and Confidence Intervals for Simulated Power Law Noise

Neil Ashby

Abstract—A method for simulating power law noise in clocks and oscillators is presented based on modification of the spectrum of white phase noise, then Fourier transforming to the time domain. Symmetric real matrices are introduced whose traces—the sums of their eigenvalues—are equal to the Allan variances, in overlapping or non-overlapping forms, as well as for the corresponding forms of the modified Allan variance. We show that the standard expressions for spectral densities, and their relations to Allan variance, are obtained with this method. The matrix eigenvalues determine probability distributions for observing a variance at an arbitrary value of the sampling interval τ , and hence for estimating confidence in the measurements. Examples are presented for the common power-law noises. Extension to other variances such as the Hadamard variance, and variances with dead time, are discussed.

I. INTRODUCTION

THE characterization of clock performance by means of average measures such as Allan variance, Hadamard variance, Theo, and modified forms of such variances, is widely applied within the time and frequency community as well as by most clock and oscillator fabricators. Such variances are measured by comparing the times t_k on a device under test, with the times at regular intervals $k\tau_0$ on a perfect reference, or at least on a better reference. Imperfections in performance of the clock under test are studied by analyzing noise in the time deviation sequence $x_k = t_k - k\tau_0$, or the fractional frequency difference during the sampling interval $\tau = s\tau_0$:

$$\Delta_{k,s}^{(1)} = (x_{k+s} - x_k)/(s\tau_0). \quad (1)$$

The frequency spectrum of fractional frequency differences can usually be adequately characterized by linear superposition of a small set of types of power law noise. The frequency spectrum of the fractional frequency differences of a particular noise type is given by a one-sided spectral density [1]

$$S_y(f) = h_\alpha f^\alpha, \quad f > 0. \quad (2)$$

[The units of $S_y(f)$ are Hz^{-1} .] For the common power-law noise types, α varies in integral steps from +2 down to -2, corresponding to white phase modulation, flicker phase

modulation, white frequency modulation, flicker frequency modulation, and random walk of frequency, respectively.

Simulation of clock noise can be extremely useful in testing software algorithms that use various windowing functions and Fourier transform algorithms to extract spectral density and stability information from measured time deviations, and especially in predicting the probability for observing a particular value of some clock stability variance. This paper develops a simple simulation method for a time-difference sequence that guarantees the average spectral density will have some chosen average power law dependence. Expressions for the common variances and their modified forms are derived here which agree with expressions found in the literature, with some exceptions. This approach also leads to predictions of probabilities for observing a variance of a particular type at particular values of the averaging time. A broad class of probability functions naturally arises. These only rarely correspond to chi-squared distributions.

The present approach to noise simulation was motivated by difficulties in applying methods such as those discussed in [2]–[4] to acceleration noise in spacecraft at very low frequencies [5]. Noisy accelerations of spacecraft can arise from causes such as fluctuations in solar radiation pressure, anisotropic thermal radiation, charged particle drag, etc., and after being integrated twice, can give rise to a spectral density of displacement noise that diverges at low frequencies faster than any known power-law clock noise. For example, an acceleration noise amplitude proportional to f^{-1} integrated twice gives rise to a displacement noise spectral density proportional to f^{-6} . Study of such acceleration noise is important in establishing error budgets of space missions in which spacecraft are tracked by means of electromagnetic signals. In [6], the present method was applied to the analysis of active galactic nuclei (AGN) light curves and to improving estimates of errors in period searches by epoch-folding techniques. In [7], it was applied to the analysis of flicker-like phase noise with a spectral density that behaves as $f^{-0.83}$.

Ref. [2] emphasizes the role of causality in the analysis of noise driving a linear system, and presents an algorithm incorporating this feature into the time series itself. The methodology used here is different: the time series produced by the present method is periodic; the frequency spectrum from which it is derived does not exhibit aliasing or leakage, although these effects can be introduced by modifying the time series, e.g., by undersampling or windowing. Various windowing algorithms could be com-

Manuscript received November 14, 2013; accepted October 9, 2014.

N. Ashby is with the University of Colorado, Boulder, CO, and the National Institute of Science and Technology, Boulder, CO (e-mail: ashby@boulder.nist.gov).

DOI <http://dx.doi.org/10.1109/TUFFC.2013.006167>

pared, for example, by generating the periodic time series, applying a window to the data, then transforming back to the frequency domain.

This paper is organized as follows: Section II introduces the basic simulation method and Section III applies the method to the overlapping Allan variance. Section IV shows how diagonalization of the averaged squared second-difference operator, applied to the simulated time series, leads to expressions for the probability of observing a value of the variance for some chosen value of the sampling or averaging time. Expressions for the mean squared deviation of the mean of the variance itself are derived in Section VI. The approach is used to discuss the modified Allan variance in Section VII, and the non-overlapping form of the Allan variance is treated in Section VIII. Appendix A discusses evaluation of a contour integral for the probability, first introduced in Section IV, in the general case.

II. DISCRETE TIME SERIES

We imagine the noise amplitudes at Fourier frequencies f_m are generated by a set of N normally distributed random complex numbers w_n having mean zero and variance $\langle w_m^* w_m \rangle = 2\sigma^2$, that would by themselves generate a simulated spectrum for white phase noise. These random numbers are divided by a function of the frequency, $|f_m|^\lambda$, producing a spectral density that has the desired frequency characteristics. For ordinary power-law noise, the exponent λ is a multiple of $1/2$, but it could be anything. The frequency noise is then transformed to the time domain, producing a time series with the statistical properties of the selected power law noise. The Allan variance, modified Allan variance, Hadamard variance, variances with dead time, and other quantities of interest can be calculated using either the frequency noise or the time series.

In the present paper, we discuss applications to calculation of various versions of the Allan variance. Of considerable interest are results for the probability of observing a value of the Allan variance for particular values of the sampling time τ and time series length N . The derivations in this paper are theoretical predictions. A natural frequency cutoff occurs at $f_h = 1/(2\tau_0)$, where τ_0 is the time between successive time deviations. This number is not necessarily related in an obvious way to some hardware bandwidth. The measurements are assumed to be made at the times $k\tau_0$, and the time errors or residuals relative to the reference clock are denoted by x_k . The averaging or sampling time is denoted by $\tau = s\tau_0$, where s is an integer. The total length of time of the entire measurement series is $T = N\tau_0$. The possible frequencies that occur in the Fourier transform of the time residuals are

$$f_m = \frac{m}{N\tau_0}, \quad -\frac{N}{2} + 1 \leq m \leq \frac{N}{2}. \quad (3)$$

In order that a set of noise amplitudes in the frequency domain represent a real series in the time domain, the amplitudes must satisfy the reality condition

$$w_{-m} = (w_m)^*. \quad (4)$$

N random numbers are placed in $N/2$ real and $N/2$ imaginary parts of the positive and negative frequency spectrum. Thus, if $w_m = u_m + iv_m$, where u_m and v_m are independent uncorrelated random numbers, then $(w_m)^* = u_m - iv_m$. Because the frequencies $\pm 1/2\tau_0$ represent essentially the same contribution, $v_{N/2}$ will not appear. We shall assume the variance of the noise amplitudes is such that

$$\langle (w_m)^* w_n \rangle = \langle u^2 + v^2 \rangle \delta_{mn} = 2\sigma^2 \delta_{mn}; \quad m \neq 0, N/2. \quad (5)$$

Also, $\langle w_m^2 \rangle = \langle u^2 - v^2 + 2iuv \rangle = 0$ for $m \neq N/2$. The index m runs from $-N/2 + 1$ to $N/2$. To avoid division by zero, we shall always assume that the Fourier amplitude corresponding to zero frequency vanishes. This only means that the average of the time residuals in the time series will be zero, and has no effect on any variance that involves time differences.

We perform a discrete Fourier transform of the frequency noise and obtain the amplitude of the k th member of the time series for white PM:

$$x_k = \frac{\tau_0^2}{\sqrt{N}} \sum_{m=-N/2+1}^{N/2} e^{-(2\pi imk)/N} w_m. \quad (6)$$

The factor τ_0^2 is inserted so that the time series will have the physical dimensions of time if w_m has the dimensions of frequency. We then multiply each frequency component by $|f_0/f_m|^\lambda$. This will generate the desired power-law form of the spectral density. The time series will be represented by

$$X_k = \frac{\tau_0^2}{\sqrt{N}} \sum_{m=-N/2+1}^{N/2} \frac{|f_0|^\lambda}{|f_m|^\lambda} e^{-(2\pi imk)/N} w_m. \quad (7)$$

The constant factor $|f_0|^\lambda$ has been inserted to maintain the physical units of the time series. The parameter f_0 determines the level h_α of the noise. For this to correspond to commonly used expressions for the one-sided spectral density, (2), we shall assume that

$$\frac{\tau_0^2 |f_0|^\lambda}{\sqrt{N}} = \sqrt{\frac{h_\alpha}{16\pi^2 \sigma^2 (N\tau_0)}}; \quad (8)$$

we then may show that if $2\lambda = 2 - \alpha$, the correct average spectral density is obtained. The simulated time series is

$$X_k = \sqrt{\frac{h_\alpha}{16\pi^2 \sigma^2 (N\tau_0)}} \sum_m \frac{e^{-(2\pi imk)/N}}{|f_m|^\lambda} w_m. \quad (9)$$

The average (two-sided) spectral density of the time residuals is obtained by squaring a single term in (9) and dividing by the spacing $\Delta f = 1/(N\tau_0)$ between successive allowed frequencies:

$$s_x(f_m) = \frac{h_\alpha}{16\pi^2\sigma^2(N\tau_0)f_m^{2\lambda}} \left\langle \frac{w_m w_m^*}{\Delta f} \right\rangle = \frac{h_\alpha}{8\pi^2 f_m^{2\lambda}}. \quad (10)$$

The average (two-sided) spectral density of fractional frequency fluctuations is given by the well-known relation [8], [9]

$$s_y(f) = (2\pi f)^2 s_x(f), \quad (11)$$

and in agreement with convention, (2), the one-sided spectral density is

$$S_y(f) = 0, f < 0; \quad 2s_y(f) = h_\alpha f^\alpha, f > 0, \quad (12)$$

where $2\lambda = 2 - \alpha$.

III. OVERLAPPING ALLAN VARIANCE

Consider the second-difference operator defined by

$$\Delta_{j,s}^{(2)} = \frac{1}{\sqrt{2\tau^2}} (X_{j+2s} - 2X_{j+s} + X_j). \quad (13)$$

The fully overlapping Allan variance is formed by averaging the square of this quantity over all possible values of j from 1 to $N - 2s$. Thus,

$$\sigma_y^2(\tau) = \left\langle \frac{1}{N - 2s} \sum_{j=1}^{N-2s} (\Delta_{j,s}^{(2)})^2 \right\rangle. \quad (14)$$

In terms of the time series, (9), the second difference can be reduced using elementary trigonometric identities:

$$\begin{aligned} \Delta_{j,s}^{(2)} &= \sqrt{\frac{h_\alpha}{32\pi^2\tau^2\sigma^2N\tau_0}} \sum_m \frac{w_m}{|f_m|^\lambda} e^{-(2\pi i m(j+s))/N} \\ &\quad \times (e^{-(2\pi i m s)/N} - 2 + e^{(2\pi i m s)/N}) \\ &= -\sqrt{\frac{h_\alpha}{2\pi^2\tau^2\sigma^2N\tau_0}} \sum_m \frac{w_m e^{-(2\pi i m(j+s))/N}}{|f_m|^\lambda} \left(\sin \frac{\pi m s}{N} \right)^2. \end{aligned} \quad (15)$$

We form the averaged square of $\Delta_{j,s}^{(2)}$ by multiplying (15) by its complex conjugate, then summing over all possible values of j and averaging:

$$\begin{aligned} \sigma_y^2(\tau) &= \frac{h_\alpha}{2\pi^2\tau^2(N\tau_0)(N-2s)} \sum_{m,n,j} \left\langle \frac{w_m w_n^*}{\sigma^2} \right\rangle \\ &\quad \times \frac{\left(\sin \left(\frac{\pi m s}{N} \right) \sin \left(\frac{\pi n s}{N} \right) \right)^2}{|f_m f_n|^\lambda} e^{-2\pi i(m-n)(j+s)/N}. \end{aligned} \quad (16)$$

The average of the product of random variables only contributes $2\sigma^2$ when $m = n$ [see (5), except when $m = n = N/2$, where $\langle |w_{N/2}|^2 \rangle = \sigma^2$]. The Allan variance reduces to

$$\sigma_y^2(\tau) = \frac{h_\alpha}{\pi^2\tau^2(N\tau_0)} \left(\sum_m \frac{\left(\sin \left(\frac{\pi m s}{N} \right) \right)^4}{|f_m|^{2\lambda}} + \frac{\left(\sin \left(\frac{\pi s}{2} \right) \right)^4}{2(f_{N/2})^{2\lambda}} \right) \quad (17)$$

because every term in the sum over j contributes the same amount. The zero-frequency term is excluded from the sum. For convenience, we introduce the abbreviation

$$K = \frac{2h_\alpha}{\pi^2\tau^2(N\tau_0)}. \quad (18)$$

If we write the sum in terms of positive frequencies only, a factor of 2 comes in except for the most positive frequency, and so

$$\sigma_y^2(\tau) = K \left(\sum_{m>0}^{N/2-1} \frac{\left(\sin \left(\frac{\pi m s}{N} \right) \right)^4}{f_m^{2\lambda}} + \frac{\left(\sin \frac{\pi s}{2} \right)^4}{4(f_{N/2})^{2\lambda}} \right). \quad (19)$$

The influence of the second term in (19) is very small except when s is very small; we shall therefore neglect it in the remainder of this paper. If the frequencies are spaced densely enough to pass from the sum to an integral, then $\Delta f = (N\tau_0)^{-1}$ and

$$\frac{1}{N\tau_0} \sum_m F(|f_m|) \rightarrow \int F(|f|) df, \quad (20)$$

and we obtain the standard result [12]

$$\sigma_y^2(\tau) = 2 \int_0^{f_h} \frac{S_y(f) df}{\pi^2\tau^2 f^2} (\sin(\pi f \tau))^4. \quad (21)$$

Similar arguments lead to known expressions for the non-overlapping version of the Allan variance, the modified Allan variance, and the Hadamard variance; proofs of these statements can be given provided no windowing or under-sampling is applied to the time series. These variances will be discussed in later sections.

IV. EIGENVALUES AND PROBABILITIES

In this section, we shall develop expressions for the probability of observing a particular value A_o for the overlapping Allan variance in a single measurement, or in a single simulation run. A_o is a random variable representing a possible value of the overlapping variance. We use a subscript o to denote the completely overlapping case. To save writing, we introduce the following abbreviations:

$$\begin{aligned}
F_m^j &= \frac{\left(\sin\left(\frac{\pi ms}{N}\right)\right)^2}{|f_m|^\lambda} \cos\left(\frac{2\pi m(j+s)}{N}\right) \\
G_m^j &= \frac{\left(\sin\left(\frac{\pi ms}{N}\right)\right)^2}{|f_m|^\lambda} \sin\left(\frac{2\pi m(j+s)}{N}\right).
\end{aligned} \tag{22}$$

The dependence on s is suppressed, but is to be understood. We write the second difference in terms of a sum over positive frequencies only, keeping in mind that the most positive and the most negative frequencies only contribute a single term because $\sin(\pi(j+s)) = 0$. The imaginary contributions cancel, and from (15) we obtain

$$\Delta_{j,s}^{(2)} = -\sqrt{K} \sum_{m>0} \left(F_m^j \frac{u_m}{\sigma} + G_m^j \frac{v_m}{\sigma} \right). \tag{23}$$

There is no term in $v_{N/2}$. Then, from (17), the overlapping Allan variance is given by

$$\sigma_y^2(\tau) = \frac{K}{N-2s} \sum_j \sum_{m>0} ((F_m^j)^2 + (G_m^j)^2). \tag{24}$$

To compute the probability that a particular value A_o is observed for the Allan variance, given all the possible values that the random variables $u_1, v_1, \dots, u_{N/2}$ can have, we form the integral

$$\begin{aligned}
P(A_o) &= \int \delta\left(A_o - \frac{1}{N-2s} \sum_j (\Delta_{j,s}^{(2)})^2\right) \\
&\times \prod_{m>0} \left(e^{-(u_m^2 + v_m^2)/(2\sigma^2)} \frac{du_m dv_m}{2\pi\sigma^2} \right).
\end{aligned} \tag{25}$$

The Dirac delta function constrains the averaged second difference to the specific value A_o and the normally distributed random variables $u_1, v_1, \dots, u_m, v_m, \dots, u_{N/2}$ range over their values. There is no integral for $v_{N/2}$. Inspecting this probability and (23) for the second difference indicates that we could dispense with the factors of σ^{-1} and work with normally distributed random variables having variance unity. Henceforth, we set $\sigma = 1$.

The exponent involving the random variables is a quadratic form that can be written in matrix form by introducing the $N-1$ dimension column vector U (the zero-frequency component is excluded)

$$U^T = [u_1, v_1, \dots, u_m, v_m, \dots, v_{N/2-1}, u_{N/2}]. \tag{26}$$

Then

$$\frac{1}{2} \sum_{m>0} (u_m^2 + v_m^2) = \frac{1}{2} U^T U = \frac{1}{2} U^T \mathbf{1} U, \tag{27}$$

where $\mathbf{1}$ represents the unit matrix. The delta-function in (25) can be written in exponential form by introducing

one of its well-known representations, an integral over all angular frequencies ω [10]:

$$\begin{aligned}
P(A_o) &= \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{i\omega(A_o - (1/(N-2s)) \sum_j (\Delta_{j,s}^{(2)})^2)} \\
&\times \prod_{m>0} \left(e^{-(u_m^2 + v_m^2)/(2\sigma^2)} \frac{du_m dv_m}{2\pi\sigma^2} \right).
\end{aligned} \tag{28}$$

The contour of integration goes along the real axis in the complex ω plane.

The squared second difference is a complicated quadratic form in the random variables $u_1, v_1, \dots, u_m, v_m, \dots, u_{N/2}$. If this quadratic form could be diagonalized without materially changing the other quadratic terms in the exponent, then the integrals could be performed in spite of the imaginary factor i in the exponent. To accomplish this, we introduce a column vector C^j that depends on j, m, s, N and whose transpose is

$$(C^j)^T = [F_1^j, G_1^j, \dots, F_m^j, G_m^j, \dots, G_{N/2-1}^j, F_{N/2}^j]. \tag{29}$$

Dependence on s is not written explicitly but is understood. The column vector has $N-1$ real components. It contains all the dependence of the second difference on frequency, averaging time τ , and on the particular power law noise. We use indices $\{m, n\}$ as matrix (frequency) indices. The (scalar) second difference operator in (23) can be written very compactly as a matrix product

$$\Delta_{j,s}^{(2)} = -\sqrt{K} (C^j)^T U = -\sqrt{K} U^T C^j. \tag{30}$$

Then, the quantity in (14) to be averaged is

$$\frac{1}{N-2s} \sum_j (\Delta_{j,s}^{(2)})^2 = U^T \left(\frac{K}{N-2s} \sum_j C^j (C^j)^T \right) U. \tag{31}$$

The matrix

$$H_o = \frac{K}{N-2s} \sum_j C^j (C^j)^T \tag{32}$$

is real and symmetric. H_o is also Hermitian and therefore has real eigenvalues. A real symmetric matrix can be diagonalized by an orthogonal transformation [11], [12], which we denote by O . Although we shall not need to determine this orthogonal transformation explicitly, it could be found by first finding the eigenvalues ε and eigenvectors ψ_ε of H_o , by solving the equation

$$H_o \psi_\varepsilon = \varepsilon \psi_\varepsilon. \tag{33}$$

The transformation O is a matrix of dimension $(N-1) \times (N-1)$ consisting of the components of the normalized eigenvectors placed in columns. Then

$$H_o O = O E, \tag{34}$$

where E is a diagonal matrix with entries equal to the eigenvalues of the matrix H_0 . Then, because the transpose of an orthogonal matrix is the inverse of the matrix,

$$O^T H_0 O = E. \quad (35)$$

The matrix H_0 is thus diagonalized, at the cost of introducing a linear transformation of the random variables:

$$\begin{aligned} \frac{1}{N-2s} \sum_j (\Delta_{j,s}^{(2)})^2 &= U^T H_0 U = U^T O O^T H_0 O O^T U \\ &= (U^T O) E (O^T U). \end{aligned} \quad (36)$$

We introduce $N-1$ new random variables by means of the transformation

$$V = O^T U. \quad (37)$$

Then, the term in the exponent representing the Gaussian distributions is

$$\begin{aligned} -\frac{1}{2} U^T \mathbf{1} U &= -\frac{1}{2} U^T O \mathbf{1} O^T U \\ &= -\frac{1}{2} V^T \mathbf{1} V = -\frac{1}{2} \sum_{n=1}^{N-1} V_n^2. \end{aligned} \quad (38)$$

The Gaussian distributions remain basically unchanged.

Further, the determinant of an orthogonal matrix is ± 1 , because the transpose of the matrix is also the inverse:

$$\det(O^{-1}O) = 1 = \det(O^T O) = (\det(O))^2. \quad (39)$$

Therefore, changes in the volume element are simple because the volume element for the new variables is

$$\begin{aligned} dV_1 dV_2 \dots dV_{N-1} &= \left| \det \left(\frac{\partial V_m}{\partial U_n} \right) \right| dU_1 dU_2 \dots dU_{N-1} \\ &= |\det(O)| dU_1 dU_2 \dots dU_{N-1} \\ &= dU_1 dU_2 \dots dU_{N-1}. \end{aligned} \quad (40)$$

After completing the diagonalization,

$$\frac{1}{N-2s} \sum_j (\Delta_{j,s}^{(2)})^2 = \sum_i \varepsilon_i V_i^2. \quad (41)$$

The probability is therefore

$$P(A_0) = \int \frac{d\omega}{2\pi} e^{i\omega(A_0 - \sum_k \varepsilon_k V_k^2)} \prod_i \left(e^{-(V_i^2)/2} \frac{dV_i}{\sqrt{2\pi}} \right). \quad (42)$$

An eigenvalue of zero will not contribute in any way to this probability because the random variable corresponding to a zero eigenvalue just integrates out.

Let the eigenvalue ε_i have multiplicity μ_i , which means that the eigenvalue ε_i is repeated μ_i times. Integration

over the random variables then gives a useful form for the probability:

$$P(A_0) = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \frac{e^{i\omega A_0}}{\prod_i (1 + 2i\varepsilon_i \omega)^{\mu_i/2}}. \quad (43)$$

Finally, the contour integral may be deformed and closed in the upper half complex plane, where it encloses the singularities of the integrand. This is discussed in Appendix A. If $A_0 < 0$ the contour may be closed in the lower half-plane where there are no singularities, so in this case $P(A_0 < 0) = 0$. Knowing the probability, one may integrate with respect to the variance A_0 to find a cumulative distribution:

$$\Phi(A_0) = \int_0^{A_0} P(A) dA. \quad (44)$$

The cumulative distribution determines the values of A_0 corresponding to a specified confidence limits; for example, the median of A_0 occurs when $\Phi(A_0) = 0.5$; values of A_0 occur with probability 0.5 when $\Phi(A_0)$ is within the limits $\Phi(A_0) = 0.5 \pm 0.25$.

A. Properties of the Eigenvalues

First, it is easily checked that the probability is correctly normalized by integrating over all A_0 and using properties of the delta function:

$$\begin{aligned} \int P(A_0) dA_0 &= \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \frac{\int e^{i\omega A_0} dA_0}{\prod_i (1 + 2i\varepsilon_i \omega)^{\mu_i/2}} \\ &= \int_{-\infty}^{+\infty} \frac{\delta(\omega) d\omega}{\prod_i (1 + 2i\varepsilon_i \omega)^{\mu_i/2}} \\ &= \int_{-\infty}^{+\infty} d\omega \delta(\omega) = 1. \end{aligned} \quad (45)$$

Second, the eigenvalues are all either positive or zero. The eigenvalue equation for the eigenvector labeled by ε is

$$\frac{K}{N-2s} \sum_j C^j (C^j)^T \psi_\varepsilon = \varepsilon \psi_\varepsilon. \quad (46)$$

Multiply on the left by ψ_ε^T ; assuming the vector has been normalized, we obtain

$$\varepsilon = \frac{K}{N-2s} \sum_j ((C^j)^T \psi_\varepsilon)^2 \geq 0. \quad (47)$$

Thus, every eigenvalue must be positive or zero.

Next, let us calculate the trace of H_0 . Because the trace is not changed by an orthogonal transformation,

$$\begin{aligned}\text{Trace}(O^T H_o O) &= \text{Trace}(H_o O O^T) = \text{Trace}(H_o O O^{-1}) \\ &= \text{Trace}(H_o) = \sum_i \varepsilon_i.\end{aligned}\quad (48)$$

The sum of the diagonal elements of H_o equals the sum of the eigenvalues of H_o . If we then explicitly evaluate the sum of the diagonal elements of H_o , we find from (48) and (22)

$$\begin{aligned}\sum_i \varepsilon_i &= \frac{K}{N-2s} \sum_j \text{Trace}(C^j (C^j)^T) \\ &= \frac{K}{N-2s} \sum_j \sum_{m>0} ((F_m^j)^2 + (G_m^j)^2) \\ &= K \sum_{m>0} \frac{(\sin \frac{\pi ms}{N})^4}{|f_m|^{2-\alpha}} = \sigma_y^2(\tau).\end{aligned}\quad (49)$$

Every term labeled by j contributes the same amount. We obtain the useful result that *the overlapping Allan variance is equal to the sum of the eigenvalues of the matrix H_o* . Similar results can be established for many of the other types of variances. (See Section VIII.)

B. Distribution of Eigenvalues

The equation for eigenvalues $H_o \psi_\mu = \varepsilon \psi_\mu$ produces many zero eigenvalues, especially when τ is large. The dimension of the matrix H_o is therefore much larger than necessary. We have performed extensive numerical calculations for many different values of N , which indicate that for the completely overlapping Allan variance, the eigenvalue equation has a total of $N-1$ eigenvalues, but only $N-2s$ nonzero eigenvalues; the number of significant eigenvalues is, in fact, equal to the number of terms $N-2s$ in the sum over j in the equations:

$$\frac{K}{N-2s} \sum_n \sum_j^{N-2s} (C_n)^j (C_n)^j \psi_n = \varepsilon \psi_n. \quad (50)$$

The factorized form of H_o , which arises on squaring the difference operator in (32), permits the reduction of the size of the matrix that is to be diagonalized. We introduce the quantities

$$\phi_\mu^j = \sum_n (C_n)^j \psi_{n\mu}. \quad (51)$$

We use the Greek index μ to label a nonzero eigenvalue and the index ν to label a zero eigenvalue. The eigenvalue equation becomes

$$\frac{K}{N-2s} \sum_j (C_m)^j \phi_\mu^j = \varepsilon \psi_{m\mu}. \quad (52)$$

Multiply by $(C_m)^l$ and sum over the frequency index m . Then,

$$\frac{K}{N-2s} \sum_{m,j} (C_m)^l (C_m)^j \phi_\mu^j = \varepsilon \phi_\mu^l. \quad (53)$$

This is an eigenvalue equation with reduced dimension $N-2s$ rather than $N-1$, because the number of possible values of j is $N-2s$. The eigenvalue equation can therefore be written in terms of a reduced matrix H_{red} , given by

$$(H_{\text{red}})^{lj} = \frac{K}{N-2s} \sum_m (C_m)^l (C_m)^j. \quad (54)$$

The indices l, j run from 1 to $N-2s$. Eigenvalues generated by (53) are all nonzero. To prove this, multiply (53) by ϕ_μ^l and sum over l . We obtain

$$\frac{K}{N-2s} \sum_m \left(\sum_l (C_m)^l \phi_\mu^l \right)^2 = \varepsilon \sum_l (\phi_\mu^l)^2. \quad (55)$$

The eigenvalue cannot be zero unless

$$\sum_l (C_m)^l \phi_\mu^l = 0 \quad (56)$$

for every m . The number of such conditions however is larger than the number $N-2s$ of variables, so the only way this can be satisfied is if $\phi_\mu^l = 0$, a trivial solution. Therefore, to obtain normalizable eigenvectors from (53), the corresponding eigenvalues must all be positive. This is true even though some of these conditions may be trivially satisfied if the factor $\sin(\pi ms/N)$ vanishes, which happens sometimes when

$$ms = MN, \quad (57)$$

where M is an integer. Every time a solution of (57) occurs, two equations relating components of ϕ_μ^l are lost. Suppose there were n solutions to (57); then the number of conditions lost would be $2n$. The number of variables is $N-2s$ and the number of conditions left in (56) would be $N-1-2n$. The excess of conditions over variables is thus

$$N-1-2n-(N-2s) = 2(s-n)-1. \quad (58)$$

In Appendix B, we prove that under all circumstances $2(s-n)-1 > 0$.

We temporarily drop the subscript o because the remainder of the results in this section are valid for any of the variances. If the eigenvalues are found and the appropriate matrix is diagonalized, we may compute the probability for observing a value of the variance, denoted by the random variable A , by

$$\begin{aligned}
P(A) &= \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{i\omega(A - V^T E V)} \prod_i \left(\frac{e^{-V_i^2/2} dV_i}{\sqrt{2\pi}} \right) \\
&= \int \frac{d\omega}{2\pi} \frac{e^{i\omega A}}{\prod (1 + 2i\varepsilon_i \omega)^{\mu_i/2}}.
\end{aligned} \tag{59}$$

C. Case of a Single Eigenvalue

If a single eigenvalue occurs once only, the general probability expression, (59), has a single factor in the denominator:

$$P(A) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{d\omega e^{i\omega A}}{\sqrt{1 + 2i\omega\varepsilon}}. \tag{60}$$

The integral is performed by closing the contour in the upper half complex ω plane. There is a branch point on the imaginary axis at $\omega = i/(2\varepsilon)$, and a branch line from that point to infinity. Evaluation of the integral gives

$$P(A) = \frac{1}{\sqrt{2\pi\sigma_y^2(\tau)}} \frac{e^{-A/(2\sigma_y^2(\tau))}}{\sqrt{A}}. \tag{61}$$

This is a chi-squared distribution with exactly one degree of freedom. The computation of the confidence interval for a given $\tau = \tau_0 s$ is simple. The cumulative probability obtained from (61) is

$$\Phi(A) = \int_0^A \frac{1}{\sqrt{2\pi\sigma_y^2(\tau)}} \frac{e^{-x/(2\sigma_y^2(\tau))}}{\sqrt{x}} dx = \operatorname{erf}\left(\frac{A}{\sqrt{2\sigma_y^2(\tau)}}\right). \tag{62}$$

The $\pm 25\%$ limits on the probability of observing a value A are then found to be $1.323\sigma_y^2$ and $0.1015\sigma_y^2$, respectively. An example of this is plotted in Fig. 1. (A is a variance, not a deviation.)

D. Case of Two Distinct Nonzero Eigenvalues

For the overlapping variance, when s has its maximum value $N/2 - 1$ there are two unequal eigenvalues. The probability integral can be performed by closing the contour in the upper half plane, and gives the expression

$$P(A) = \frac{1}{2\sqrt{\varepsilon_1\varepsilon_2}} e^{-(A/4)(1/\varepsilon_1 + 1/\varepsilon_2)} I_0\left(\frac{A}{4}\left(\frac{1}{\varepsilon_2} - \frac{1}{\varepsilon_1}\right)\right), \tag{63}$$

where I_0 is the modified Bessel function of order zero [13], [14]. The probability is correctly normalized. It differs from a chi-squared distribution in that the density does not have a singularity at $A = 0$. This is illustrated in Fig. 2.

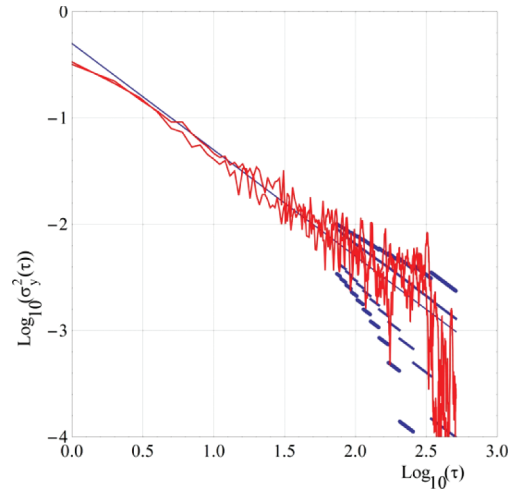


Fig. 1. Comparisons of average non-overlapping Allan variance (light line) with rms deviations from the mean of the variance (heavy lines) and true $\pm 25\%$ limits (medium heavy lines) for $N = 1024$ data items in the time series for the Allan variance for two independent simulation runs are plotted for comparison.

Evaluation of the contour integral when there are more than two distinct eigenvalues is discussed in Appendix A. If an eigenvalue occurs an even number $2n$ times, the corresponding singularity becomes a pole of order n and a chi-squared probability distribution may result; this has only been observed to occur for white PM.

V. CONFIDENCE ESTIMATES

One of the goals of this investigation is to estimate the uncertainty in the variance when some value of the variance is measured. This can be rigorously defined if the probability $P(A_0)$ is available, but it may be difficult to evaluate the integral in (43). In this case, one might be

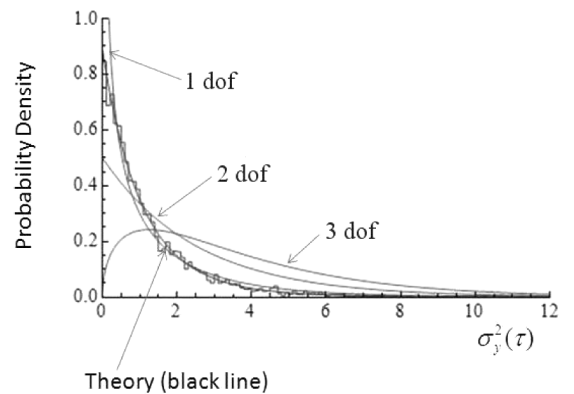


Fig. 2. Simulation of non-overlapping Allan variance for flicker fm with $N = 64$, $s = 19$. For this case, there are two distinct eigenvalues in the matrix for the non-overlapping case. The variance for $s = 19$ was extracted from each of 4000 independent runs and a histogram of the values obtained was constructed for comparison with the probability, (63). Chi-squared distributions with 1, 2, and 3 degrees of freedom are plotted for comparison.

interested in a measure such as the rms value of the variance, measured with respect to the mean variance $\sigma_y^2(\tau)$. One method for obtaining this measure is obtained from the probability without introducing a Fourier integral representation for the delta function:

$$\begin{aligned} \int A_o^2 P(A_o) dA_o &= \int A_o^2 \delta\left(A_o - \sum_i \varepsilon_i V_i^2\right) \\ &\quad \times \prod_m \left(e^{-V_m^2/2} \frac{dV_m}{\sqrt{2\pi}}\right) dA_o \\ &= \int \left(\sum_i \varepsilon_i V_i^2\right)^2 \delta\left(A_o - \sum_i \varepsilon_i V_i^2\right) \\ &\quad \times \prod_m \left(e^{-V_m^2/2} \frac{dV_m}{\sqrt{2\pi}}\right) dA_o \\ &= \int \left(\sum_i \varepsilon_i V_i^2\right)^2 \prod_m \left(e^{-V_m^2/2} \frac{dV_m}{\sqrt{2\pi}}\right). \end{aligned} \quad (64)$$

Expanding the square in the integrand, there are N fourth-order terms and $N(N-1)$ cross terms, so we get

$$\begin{aligned} \int A_o^2 P(A_o) dA_o &= \sum_i \varepsilon_i^2 \langle V^4 \rangle + 2 \sum_{i \neq j} \varepsilon_i \varepsilon_j \langle V^2 \rangle^2 \\ &= 3 \sum_i \varepsilon_i^2 + 2 \sum_{i \neq j} \varepsilon_i \varepsilon_j. \end{aligned} \quad (65)$$

To obtain the variance of the variance, we must subtract

$$\langle A_o \rangle^2 = \sum_i \varepsilon_i^2 + 2 \sum_{i \neq j} \varepsilon_i \varepsilon_j. \quad (66)$$

The result is

$$\langle A_o^2 \rangle - \langle A_o \rangle^2 = 2 \sum_i \varepsilon_i^2. \quad (67)$$

Therefore, the rms deviation of the variance from the mean variance is equal to $\sqrt{2}$ times the square root of the sum of the squared eigenvalues. This result is not very useful if there are only a small number of eigenvalues, because the factor $\sqrt{2}$ can make the rms deviation from the mean of the variance larger than the variance. This is because the probability distribution of the variance is not a normal distribution, and can lead to unreasonably large estimates of confidence intervals. An example is plotted in Fig. 1.

If the eigenvalues or the probabilities are not readily available, a similar confidence estimate can be obtained from an alternative form of (67) by considering the trace of the square of H_o :

$$2 \sum_i \varepsilon_i^2 = 2 \sum_{i,j} \left(\sum_{m>0} (F_m^i F_m^j + G_m^i G_m^j) \right)^2. \quad (68)$$

To prove this, consider diagonalization of H_o by the orthogonal transformation O , and compute

$$\begin{aligned} 2 \sum_i \varepsilon_i^2 &= 2 \text{Trace}(E^2) \\ &= 2 \text{Trace}(O^T H_o O O^T H_o O) = 2 \text{Trace} H_o^2 \\ &= \frac{2K^2}{(N-2s)^2} \sum_{m,n,i,j} (((C_m)^i (C_n)^i) ((C_n)^j (C_m)^j)) \\ &= \frac{2K^2}{(N-2s)^2} \sum_{m,n,i,j} (((C_m)^i (C_m)^j) ((C_n)^j (C_n)^i)) \\ &= \frac{2K^2}{(N-2s)^2} \sum_{i,j} \left(\sum_m (F_m^i F_m^j + G_m^i G_m^j) \right)^2. \end{aligned} \quad (69)$$

If there are not too many terms in the sums over i and j then the sum over the frequency index in (68) can be usefully written as

$$\sum_m (F_m^i F_m^j + G_m^i G_m^j) = \sum_{m>0} \frac{\left(\sin \frac{\pi m s}{N}\right)^4}{f_m^{2\lambda}} \cos \frac{2\pi m(i-j)}{N}. \quad (70)$$

This leads to some useful approximation schemes, but these will not be discussed further in this paper.

Usually the variance of the variance is larger than the range of possible values of the variance computed from the $\pm 25\%$ limits obtained from calculated probabilities. An example is given in Fig. 1, where the confidence intervals are plotted for the non-overlapping variance for $N = 1024$. In Fig. 1, $h_0 = 1$ and $\tau_0 = 1$. For s ranging between 342 and 511 there is only one eigenvalue, so the rms deviation of the variance from the variance is $\sqrt{2}$ times the variance. This gives an upper limit on the confidence interval that is 1.414 times the variance, larger than that obtained from the actual probability distribution. The medium heavy lines in Fig. 1 show the true $\pm 25\%$ probability limits obtained from calculations such as in (107) in Appendix A.

VI. MODIFIED ALLAN VARIANCE

The modified Allan variance is defined so that averages over time are performed before squaring and averaging. We use a subscript m to distinguish this form from the overlapping form of the variance. The definition is

$$\sigma_m^2(\tau) = \left\langle \left(\frac{1}{s} \sum_{j=1}^s \Delta_{j,s}^{(2)} \right)^2 \right\rangle. \quad (71)$$

Summing the expression for the second-difference given in (15),

$$\begin{aligned} \frac{1}{s} \sum_j \Delta_{j,s}^{(2)} &= -\sqrt{\frac{K}{4s^2 \sigma^2 (N\tau_0)}} \sum_m \frac{w_m}{|f_m|^\lambda} \left(\sin \frac{\pi m s}{N} \right)^2 \\ &\quad \times \left(\frac{\sin \frac{\pi m s}{N}}{\sin \frac{\pi m}{N}} \right) e^{-\pi i m (3s+1)/N}. \end{aligned} \quad (72)$$

Squaring this, we write the complex conjugate of one factor and obtain the ensemble average

$$\sigma_m^2(\tau) = \frac{K}{4s^2\sigma^2} \sum_{m,n} \frac{\langle w_m w_n^* \rangle \left(\frac{\sin \frac{\pi ms}{N}}{N} \right)^6}{|f_m f_n|^\lambda \left(\frac{\sin \frac{\pi m}{N}}{N} \right)^2}. \quad (73)$$

Using (5) and writing the result in terms of a sum over positive frequencies only,

$$\sigma_m^2(\tau) = 2 \sum_{m>0} \frac{S_y(f_m)}{\pi^2 s^2 \tau^2 f_m^2 (N\tau_0)} \left(\frac{\sin \frac{\pi ms}{N}}{N} \right)^6 \left(\frac{\sin \frac{\pi m}{N}}{N} \right)^2. \quad (74)$$

Passing to an integral when the frequency spacing is sufficiently small,

$$\sigma_m^2(\tau) = 2 \int_0^{f_h} \frac{df S_y(f)}{\pi^2 s^2 \tau^2 f^2} \frac{(\sin(\pi f \tau))^6}{(\sin(\pi f \tau_0))^2}. \quad (75)$$

This agrees with previously derived expressions [15]. The appearance of the sine function in the denominator of this expression makes the explicit evaluation of the integral difficult. In Table I, we give the leading contributions to the modified Allan variance for very large values of the averaging time τ . In general there are additional oscillatory contributions with small amplitudes that are customarily neglected. These results do not agree, however, with those published in [16]; the leading terms do agree with those published in [17].

For the modified case, instead of (32), there are two summations; we use a subscript m on H to denote the modified form:

$$H_m = \frac{K}{s^2} \sum_j C^j \sum_k (C^k)^T. \quad (76)$$

We then seek solutions of the eigenvalue equation

$$H_m \psi_\varepsilon = \frac{K}{s^2} \sum_j C^j \sum_k (C^k)^T \psi_\varepsilon = \varepsilon \psi_\varepsilon. \quad (77)$$

Multiply by $\sum_l (C^l)^T$ and sum over the frequency index.

$$\frac{K}{s^2} \sum_{l,j,m} ((C_m)^l)^T (C_m)^j \left(\sum_k (C^k)^T \psi_\varepsilon \right) = \varepsilon \left(\sum_l (C^l)^T \psi_\varepsilon \right). \quad (78)$$

The matrix equation has been reduced to a scalar equation for the quantity

$$\phi = \sum_k (C^k)^T \psi_\varepsilon. \quad (79)$$

Here, the matrix product of $(C^k)^T$ with ψ_ε entails a sum over all frequencies. Because we have a scalar eigenvalue equation for ε , there can be one and only one eigenvalue, which is easily seen to be the same as given by (74) for each sampling time τ :

$$\varepsilon = \frac{K}{s^2} \sum_{l,k,m} ((C_m)^l)^T (C_m)^k = \sigma_m^2(\tau). \quad (80)$$

The probability distribution will be of the form of (61), a chi-squared distribution with one degree of freedom.

VII. NON-OVERLAPPING ALLAN VARIANCE

In the non-overlapping form of the Allan variance, the only change with respect to (14) is that the sum over j goes from 1 in steps of s up to $j_{\max} \leq N - 2s$. We denote the number of values of j by n_{\max} . The average non-overlapping variance is the same as that for the overlapping case, but the probability distributions are different. The matrix H_{no} takes the form

$$(H_{\text{no}})_{mn} = \frac{K}{n_{\max}} \sum_{j=1,1+s,\dots}^{j_{\max}} (C_m)^j (C_n)^j. \quad (81)$$

The subscript no labels a non-overlapping variance, and the indices m and n label frequencies. Here, there are only n_{\max} terms in the sum because the values of j skip by s . n_{\max} is given by

TABLE I. ASYMPTOTIC EXPRESSIONS FOR THE MODIFIED ALLAN VARIANCE, IN THE LIMIT OF LARGE SAMPLING TIMES $\tau = s\tau_0$.

Noise type	$S_y(f)$	Mod $\sigma_y^2(\tau)$	λ	α
White PM	$h_2 f^2$	$\frac{3h_2}{\pi^2 \tau^3} \left(1 + \frac{5}{18s^3} \right)$	0	2
Flicker PM	$h_1 f$	$\frac{3h_1 \ln(256/27)}{8\pi^2 \tau^2}$	$\frac{1}{2}$	1
White FM	h_0	$\frac{h_0}{4\tau} \left(1 + \frac{1}{2s^2} \right)$	1	0
Flicker FM	$\frac{h_{-1}}{f}$	$2h_{-1} \ln \left(\frac{3^{27/16}}{4} \right)$	$\frac{3}{2}$	-1
Random walk	$\frac{h_{-2}}{f^2}$	$h_{-2} \pi^2 \tau \left(\frac{11}{20} + \frac{1}{12s^2} + \frac{1}{40s^4} \right)$	2	-2

$$n_{\max} = \left\lfloor \frac{N-1}{s} \right\rfloor - 1, \quad (82)$$

where $\lfloor Q \rfloor$ denotes the largest integer less than or equal to Q . To diagonalize H_{no} and compute probabilities, we look for eigenvalues by seeking solutions of

$$\sum_n (H_{\text{no}})_{mn} \psi_n = \varepsilon \psi_m. \quad (83)$$

We define a reduced eigenvector by

$$\phi_j = \sum_n (C_n)^j \psi_n. \quad (84)$$

The eigenvalue equation reduces to

$$\frac{K}{n_{\max}} \sum_j (C_m)^j \phi_j = \varepsilon \psi_m. \quad (85)$$

Multiply this equation by $(C_m)^l$ and sum over the frequency labels m . Then

$$\frac{K}{n_{\max}} \sum_j \sum_m ((C_m)^l (C_m)^j) \phi_j = \varepsilon \phi_l. \quad (86)$$

Let

$$J_{lj} = \frac{K}{n_{\max}} \sum_m (C_m)^l (C_m)^j = \frac{K}{n_{\max}} (C^l)^T C^j; \quad (87)$$

the sum over frequencies is accomplished by the matrix multiplication. Just as for the overlapping case, if we calculate the trace of the matrix J_{lj} , we find that because each term in the sum over j contributes the same amount,

$$\text{Trace}(J_{lj}) = \frac{K}{n_{\max}} \sum_j (C^j)^T C^j = \sigma_{\text{no}}^2(\tau). \quad (88)$$

Because the trace remains unchanged under an orthogonal transformation, the non-overlapping Allan variance will be equal to the sum of the eigenvalues of (86). The probability functions for a given τ can differ from those for the overlapping or modified cases because the number of eigenvalues and their multiplicities may be different. The eigenvalues are still all greater than or equal to zero.

If the eigenvalues are found and the matrix J_{lj} is diagonalized, we may compute the probability for observing a value of the overlapping variance, denoted by the random variable A_{no} , by

$$\begin{aligned} P(A_{\text{no}}) &= \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{i\omega(A_{\text{no}} - U^T J U)} \prod_i \frac{e^{-U_i^2/2}}{\sqrt{2\pi}} \\ &= \int \frac{d\omega}{2\pi} \frac{e^{i\omega A_{\text{no}}}}{\prod_i (1 + 2i\varepsilon_i \omega)^{\mu_i/2}}. \end{aligned} \quad (89)$$

For example, when there are only two distinct nonzero eigenvalues, the matrix J_{lk} will have elements

$$\begin{aligned} J_{1,1} &= J_{1+s,1+s} = \frac{1}{2} \sigma_{\text{no}}^2; \\ J_{1,1+s} &= J_{1+s,1} \\ &= \int_0^{f_h} \frac{S_y(f) df}{(\pi\tau f)^2} (\sin(\pi\tau f))^4 \cos(2\pi\tau f). \end{aligned} \quad (90)$$

The eigenvalues are then

$$\begin{aligned} \varepsilon_1 &= \frac{1}{2} \sigma_{\text{no}}^2 + |J_{1,1+s}| \\ \varepsilon_2 &= \frac{1}{2} \sigma_{\text{no}}^2 - |J_{1,1+s}|. \end{aligned} \quad (91)$$

The probability is of the form of (63).

Fig. 2 shows an example of this for flicker FM noise, for $N = 64$ items in the time series. The variance corresponding to $\tau = 19\tau_0$ was extracted from each of 4000 independent simulation runs and a histogram of the resulting values was plotted. Good agreement with the predicted probability distribution can be seen.

VIII. OTHER VARIANCES

The analysis methods developed in this paper can be extended to other variances. For example, the Theo variance [18] is defined by an overlapping average,

$$\begin{aligned} \sigma_{\text{Th}}^2(s, \tau_0, N) &= \frac{4}{3\tau^2(N-s)} \left\langle \sum_{i=1}^{N-m} \sum_{\delta=0}^{s/2-1} \frac{1}{s/2-\delta} \right. \\ &\quad \left. \times ((X_i - X_{i-\delta+s/2}) - (X_{i+s} - X_{i+\delta+s/2}))^2 \right\rangle, \end{aligned} \quad (92)$$

where s is assumed to be even, and $\tau = 3s\tau_0/4$. For any of the power law noises, and after passing to an integral, this expression can be transformed with the aid of elementary trigonometric identities to

$$\begin{aligned} \sigma_{\text{Th}}^2(\tau, \tau_0, N) &= \frac{8}{3} \int_0^{f_h} \frac{S_y(f) df}{(\pi\tau f)^2} \\ &\quad \times \sum_{\kappa=1}^{2\tau/3\tau_0} \frac{1}{\kappa} \left(\sin(\pi f \kappa \tau_0) \sin \left(\pi f \left(\frac{4}{3} \tau - \kappa \tau_0 \right) \right) \right)^2. \end{aligned} \quad (93)$$

The integer κ in the denominator makes this variance more difficult to evaluate, but a factorized quadratic form can still be constructed and the eigenvalue structure and resulting probabilities can be analyzed.

The Hadamard variance is defined in terms of a third difference, and is widely used to characterize clock stability when clock drift is a significant issue [19], [20]. A third difference operator may be defined as

$$\Delta_{j,s}^{(3)} = \frac{1}{\sqrt{6\tau^2}}(X_{j+3s} - 3X_{j+2s} + X_{j+s} - X_j). \quad (94)$$

The completely overlapping Hadamard variance is the average of the square of this third difference:

$$\sigma_{\text{H}}^2(\tau) = \frac{1}{N-3s} \left\langle \sum_{j=1}^{N-3s} (\Delta_{j,s}^{(3)})^2 \right\rangle \rightarrow \frac{8}{3} \int_0^{f_h} \frac{S_y(f) df}{(\pi\tau f)^2} (\sin(\pi\tau f))^6. \quad (95)$$

These methods can be applied to cases in which there is dead time between measurements of average frequency during the sampling intervals. Suppose, for example, that the measurements consist of intervals of length $\tau = s\tau_0$ during which an average frequency is measured, separated by dead time intervals of length $D - \tau$ during which no measurements are available. Let the index j label the measurement intervals with $j = 1, 2, \dots, N$. A variance can be defined in terms of the difference between the average frequency in the j th interval and that in the interval labeled by $j + r$:

$$\Delta_{j,r,s}^{(2)} = \frac{1}{\sqrt{2}}(\bar{y}_{j+r,s} - \bar{y}_{j,s}), \quad (96)$$

where $\bar{y}_{j,s}$ is the average frequency in the interval j of length $s\tau_0$. Then, an appropriate variance can be defined as

$$\Psi(\tau, D) = \langle (\Delta_{j,r,s}^{(2)})^2 \rangle. \quad (97)$$

If the measurements are sufficiently densely spaced that it is possible to pass to an integral, this can be shown to reduce to

$$\Psi(\tau, D) = \frac{2}{D\tau} \int_0^{f_h} df \frac{S_y(f)}{(\pi f)^2} (\sin(\pi frD))^2 (\sin(\pi f\tau))^2. \quad (98)$$

When $D = \tau$ and $r = 1$ there is no real dead time and this variance reduces to the ordinary Allan variance.

IX. SUMMARY AND CONCLUSION

In this paper, a method of simulating time series for the common power-law noises has been developed and applied to several variances used to characterize clock stability. These include overlapping and non-overlapping forms of the Allan variance, and the modified Allan variance. It was shown that diagonalization of quadratic forms for the average variances leads to expressions for the probabilities of observing particular values of the variance for a given sampling time $\tau = s\tau_0$. The probabilities are expressed in terms of integrals depending on the eigenvalues of matrices formed from squares of the second differences that are used to define the variance. The eigenvalues are usually distinct; only for white PM have eigenvalues been

observed (after much calculation) to occur with multiplicities other than unity; we do not have a proof of this. Generally speaking, the number of eigenvalues is equal to the number of terms occurring in the sum used to define averages of the second-difference operator. The probability distribution $P(A)$ for some variance A is useful in estimating the $\pm 25\%$ confidence interval about the average variance. The number of eigenvalues gets smaller as the sampling time τ gets larger.

It is well-known that a chi-squared distribution with n degrees of freedom occurs for a variable that is the sum of squares of n normally distributed random variables. It has been shown that, with this method, probabilities for the Allan variance, using the present simulation method, are not always—in fact are rarely—chi-squared distributions. This is because the frequency dependence, inserted to make the time series obey the chosen noise power law, disrupts the distribution of eigenvalues in most cases. Other methods of simulating power law noise have been published; the present approach differs from that of [2]. The present approach respects all the standard expressions for spectral density and the relationships between Allan variance and spectral density for the common power-law noises. It also yields reasonable simulation results for power-law noises that diverge more rapidly than flicker noise at low frequencies.

Application of these methods to simulation of phase noise will be discussed in a future paper.

APPENDIX A EVALUATION OF CONTOUR INTEGRALS FOR PROBABILITY

In almost all cases except for white PM, the eigenvalues are all distinct. We show here how then the probability function, (43), can be reduced to a sum of real integrals. For each of the eigenvalues, we introduce the quantity

$$r_k = 1/(2\varepsilon_k). \quad (99)$$

The integral becomes

$$P(A) = \left(\prod_k \left(\frac{\sqrt{r_k}}{e^{\pi i/4}} \right) \right) \int \frac{d\omega}{2\pi} \frac{e^{i\omega A}}{\prod_k (\omega - ir_k)^{1/2}}. \quad (100)$$

Because $A > 0$, by Jordan's lemma [17], the contour of integration can be deformed into the upper half-plane. The addition of circular arcs at radius $|\omega| = \infty$ contributes nothing. Each of the square root factors in (100) has a branch point at $\omega = ir_k$ with a branch line extending to $+i\infty$ along the imaginary axis. We define the complex argument of each such factor by

$$-\frac{3\pi}{2} < \arg(\omega - ir_k) < \frac{\pi}{2}. \quad (101)$$

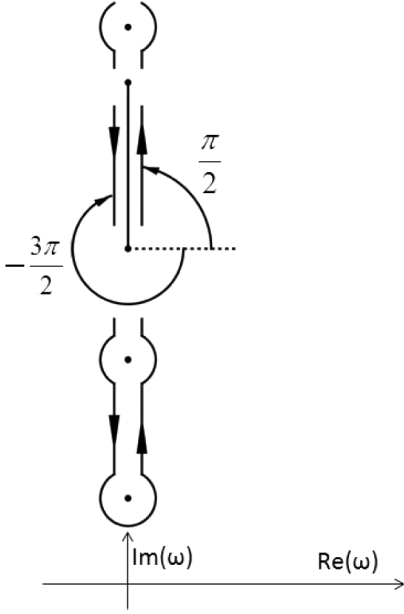


Fig. 3. Contour deformed to run along branch line on the y -axis. A pair of segments is shown with an odd number of singularities below the segments. The branch line required by square root factors of the form $\sqrt{\omega - r_k}$ is defined by the angles—either $\pi/2$ or $-3\pi/2$ —of the segments relative to the $\text{Re}(\omega)$ axis. Contributions from pairs of segments that are above an even number of branch points cancel out.

All branch lines extend along the positive y -axis to infinity. The largest eigenvalues give singularities closest to the real axis. Fig. 3 illustrates the resulting contour. Around each branch point is a circular portion which contributes nothing because as the radius δ of each circle approaches zero, the contribution to the integral approaches zero as $\sqrt{\delta}$. The integral then consists of straight segments where $\omega = iy \pm \delta$, where δ approaches zero. Two such straight segments are illustrated in Fig. 3, one on each side of the branch line along the y -axis. Suppose the interval of interest is placed so that out of a total of M eigenvalues, n of them are below the interval (three are shown in the figure). For the contribution to the integral on the left, y is decreasing, and we can account for this by reversing the limits and introducing a minus sign. The phase factor contributed by n factors with branch points below the interval is

$$\frac{1}{(e^{-(3\pi i)/4})^n} = (e^{(3\pi i)/4})^n. \quad (102)$$

The contribution from branch points above the interval of interest is the same on both sides of the y -axis, and is

$$(e^{(i\pi)/4})^{M-n}. \quad (103)$$

The factor in front of the integral sign in (100) is the same on both sides of the y -axis and is

$$(e^{-(i\pi)/4})^M. \quad (104)$$

The total phase factor of this contribution, including a factor i that comes from setting $\omega = iy$ is thus

$$-i(e^{-(i\pi)/4})^M (e^{(i\pi)/4})^{M-n} (e^{(3\pi i)/4})^n = -i(e^{(\pi i)/2})^n. \quad (105)$$

For the integral along the segment on the positive side of the y -axis, the only difference is that the phase of each of the n contributions from branch points below the interval changes from $3\pi/4$ to $-\pi/4$. The phase factor for this part of the contour is thus

$$+i(e^{-(i\pi)/4})^M (e^{(i\pi)/4})^{M-n} (e^{(-\pi i)/4})^n = +i(e^{(-\pi i)/2})^n. \quad (106)$$

If n is even, the two contributions cancel. If $n = 2m + 1$ is odd, then the contributions add up with a factor $2(-1)^m$. The probability is thus always real and consists of contributions with alternating signs, with every other interval left out.

In summary, the contour integral contributions from portions of the imaginary axis in the complex ω plane that have an even number of branch points below the interval will not contribute to the integral. For example, if there are four distinct eigenvalues, the probability will reduce to

$$P(A) = \frac{1}{\pi} \int_{r_1}^{r_2} \frac{\sqrt{r_1 r_2 r_3 r_4} e^{-yA} dy}{\sqrt{(y - r_1)(r_2 - y)(r_3 - y)(r_4 - y)}} - \frac{1}{\pi} \int_{r_3}^{r_4} \frac{\sqrt{r_1 r_2 r_3 r_4} e^{-yA} dy}{\sqrt{(y - r_1)(y - r_2)(y - r_3)(r_4 - y)}}. \quad (107)$$

Such results have been used to evaluate the probabilities for certain sampling intervals for flicker fm noise in Section IV.

APPENDIX B

PROOF THAT (53) GENERATES POSITIVE EIGENVALUES

In this Appendix, we show that (57) gives rise to positive eigenvalues under all circumstances.

Obviously, if N is a prime number, (57) can never be satisfied ($n = 0$). Consider the case $s = N - 1$. Then, $m = MN$ and there are no solutions because $m \leq N/2$. In general, both m and s may contain factors that divide M or N . Suppose $s = ab$ where at least one of the factors a, b is greater than 1, where a divides M and b divides N . Then let

$$M = am_a, \quad N = bn_b. \quad (108)$$

Then, $m = m_a n_b$. $m = m_a n_b > N/2 = bn_b/2$ cannot happen because $m \leq N/2$. If $m_a n_b = N/2 = bn_b$, then the number of solutions is $n = 1$ and $2(s - n) - 1 = 2(ab - 1) - 1 \geq 2(2 - 1) - 1 > 0$. If $m = m_a n_b < N/2 = bn_b/2$, then there may be solutions $m = m_a n_b$, $m = 2m_a n_b$, $m = 3m_a n_b$, ..., up to $bn_b/2$. The number of such solutions is

$$n = \left\lfloor \frac{bn_b}{m_a n_b} \right\rfloor = \left\lfloor \frac{b}{m_a} \right\rfloor, \quad (109)$$

where $\lfloor x \rfloor$ means the largest integer less than or equal to x . The excess of conditions is then

$$2(s - n) - 1 = 2 \left(ab - \left\lfloor \frac{b}{m_a} \right\rfloor \right) > 1, \quad (110)$$

which proves the assertion.

ACKNOWLEDGMENTS

The author is grateful to D. Howe and J. DeSalvo for discussions and corrections.

REFERENCES

- [1] D. Allan, M. A. Weiss, and J. Jespersen, "A frequency-domain view of time-domain characterization of clocks and time and frequency distribution systems," in *Proc. 45th IEEE Frequency Control Symp.*, 1991, pp. 667–678.
- [2] N. J. Kasdin and T. Walter, "Discrete simulation of power law noise," in *Proc. 1992 IEEE Frequency Control Symp.*, pp. 274–283.
- [3] T. Walter, "Characterizing frequency stability: A continuous power-law model with discrete sampling," *IEEE Trans. Instrum. Meas.*, vol. 43, no. 1, pp. 69–79, 1994.
- [4] N. J. Kasdin, "Discrete simulation of colored noise and stochastic processes and $1/f^\alpha$ power law noise generation," *Proc. IEEE*, vol. 83, no. 5, pp. 802–827, 1995.
- [5] R. T. Stebbins, P. L. Bender, J. Hanson, C. D. Hoyle, B. L. Schumaker, and S. Vitale, "Current error estimates for LISA spurious accelerations," *Class. Quantum Gravity*, vol. 21, no. 5, pp. S653–S660, 2004.
- [6] J. Timmer and M. König, "On generating power law noise," *Astron. Astrophys.*, vol. 300, pp. 707–710, 1995.
- [7] N. Ashby, "Discrete simulation of power law noise," in *Proc. 44th Annu. Precise Time and Time Interval Meeting*, 2012, pp. 289–299.
- [8] J. A. Barnes, A. R. Chi, L. S. Cutler, D. J. Healey, D. B. Leeson, T. E. McGunigal, J. A. Mullen Jr, W. L. Smith, R. L. Sydnor, R. F. C. Vessot, and G. M. R. Winkler, "Characterization of frequency stability," *IEEE Trans. Instrum. Meas.*, vol. IM-20, no. 2, pp. 105–120, May 1971.
- [9] "Characterization of frequency and phase noise," International Radio Consultative Committee (C.C.I.R.), Report 580, 1986, pp 142–150.
- [10] M. J. Lighthill, *Introduction to Fourier Analysis and Generalized Functions*. New York, NY: Cambridge University Press, 1958.

- [11] G. Strang, *Linear Algebra and its Applications*, 3rd ed., San Diego, CA: Harcourt Brace Jovanovich, 1988, sect. 5.5.
- [12] R. Stoll, *Linear Algebra and Matrix Theory*. New York, NY: McGraw-Hill, 1980.
- [13] M. Abramowitz and I. Stegun, Eds. *Handbook of Mathematical Functions* (Applied Mathematics Series 55). 2nd ed., Washington, DC: NBS, 1965, p. 374.
- [14] V. I. Smirnov, *Complex Variables—Special Functions* (A Course of Higher Mathematics, vol. 3, pt. 2). Elmsford, NY: Pergamon Press 1964, sect. 60, pp. 232–233.
- [15] D. B. Sullivan, D. W. Allan, D. A. Howe, and F. L. Walls, Eds. "Characterization of clocks and oscillators," U. S. Government Printing Office, Washington, DC, NIST Technical Note 1337, Mar. 1990, p. TN-9.
- [16] P. Lesage and T. Ayi, "Characterization of frequency stability: Analysis of the modified Allan variance and properties of its estimate," *IEEE Trans. Instrum. Meas.*, vol. 33, no. 4, pp. 332–336, Dec. 1984.
- [17] D. A. Howe and F. Vernotte, "Generalization of the total variance approach to the modified Allan variance," in *Proc. 1999 Precise Time and Time Interval Meeting*, pp. 267–276.
- [18] D. A. Howe and T. Pepler, "Very long-term frequency stability: Estimation using a special-purpose statistic," in *Proc. 2003 Joint Meeting IEEE Int. Frequency Control Symp. and European Frequency and Time Forum Conf.*, pp. 233–238.
- [19] S. Bregni and L. Jmoda, "Improved estimation of the Hurst Parameter of long-range dependent traffic using the modified Hadamard variance," in *Proc. IEEE Int. Conf. Communications*, 2006, pp. 566–572.
- [20] S. A. Hutsell, J. A. Buisson, J. D. Crum, H. S. Mobbs, and W. G. Reid, "Relating the Hadamard variance to MCS Kalman filter clock estimation," in *Proc. 27th Annu. Precise Time and Time Interval Systems and Applications Meeting*, 1995, pp. 291–301.



Neil Ashby received a B.A. degree from the University of Colorado in 1955 and a Ph.D. degree from Harvard University in 1961. He then spent a year in Europe as a Frederick Sheldon Postdoctoral Fellow at the Ecole Normale Supérieure in Paris and at Birmingham University in the UK. He served on the Faculty in the Department of Physics, University of Colorado at Boulder, from 1962 to 2003, and currently is an Affiliate of the National Institute of Standards and Technology in Boulder, CO, where he works on relativistic effects

and noise in clocks, time scales, and navigation. He has served on several international committees and working groups relating to relativistic effects on clocks with applications to timekeeping, geodesy, and navigation, and has contributed numerous studies of relativistic effects in global navigation satellite systems such as the Global Positioning System (GPS). He has received several service awards from the University of Colorado, including the University of Colorado Alumni Foundation Norlin Medal, May 2005. In January 2006, he received the F. K. Richtmyer Award of the American Association of Physics Teachers.