The Third-Difference Approach to Modified Allan Variance

Charles A. Greenhall, *Member, IEEE*

***Abstract*—This study gives strategies for estimating the modified Allan variance (MVAR), and formulas for computing the equivalent degrees of freedom (edf) of the estimators. A third-difference formulation of MVAR leads to a tractable formula for edf in the presence of power-law phase noise. The effect of estimation stride on edf is shown. First-degree rational-function approximations for edf are derived, and their errors tabulated. A theorem allowing conservative estimates of edf in the presence of compound noise processes is given.**

***Index Terms*— Allan variance, clock noise, finite differences, frequency stability, power-law noise, time and frequency.**

I. Introduction

Allan variance (AVAR) and modified Allan variance (MVAR) are statistical measures of fractional frequency instability. They are both used extensively to measure and characterize the stability performance of clocks, oscillators, and systems for disseminating time and frequency [1], [12]–[14]. Let us give brief definitions. The raw data for these measures comprise a sequence x_n of time residuals, say from a comparison of two clocks or a phase comparison of two oscillators. We assume here that the samples x_n are evenly spaced in time, with *sample period* τ_0. Let an *averaging time* $\tau = m\tau_0$ be given, where m is an integer. The Allan variance, denoted by $\sigma_y^2(\tau)$, is defined as $1/(2\tau^2)$ times the time average or mathematical expectation of the squares of second differences, with step m, of the sequence x_n. Modified Allan variance, denoted by mod $\sigma_y^2(\tau)$, is defined in the same way, except that the sequence x_n is replaced by the sequence $\overline{x}_n(m)$ of moving averages

$$\overline{x}_n(m) = \frac{1}{m}\sum_{j=0}^{m-1} x_{n-j}. \tag{1}$$

By virtue of the second difference in their definitions, stable statistical estimates of AVAR and MVAR can be accumulated in the presence of a class of phase noise models, the processes with stationary second increments [12], from which useful fits to the behavior of oscillators, amplifiers, etc., can be selected. Special cases are *power-law* models, associated with spectral densities having the property

$$S_x(f) \sim \text{const} \cdot f^{\beta}$$

as $f \to 0$, where $\beta > -5$. In the usual nomenclature of frequency and time, the noises associated with $\beta = 0, -1, -2, -3, -4$ are called white phase, flicker phase, white frequency, flicker frequency, and random-walk frequency, respectively. Nonintegral values of β are also allowed; the corresponding noises are called "fractional."

A log-log plot of measured $\sigma_y(\tau)$ or mod $\sigma_y(\tau)$ vs τ, the familiar σ-τ plot, often indicates phase noise that can be modeled as a linear combination of uncorrelated power-law components, the component associated with β being identified by a straight-line section with slope $\frac{1}{2}\,(-3-\beta)$. The main advantage of MVAR over AVAR is the increased range of β over which this slope relationship holds: $-5 < \beta < -1$ for AVAR, $-5 < \beta < 1$ for MVAR [3]. In particular, a mod $\sigma_y(\tau)$ plot can easily distinguish white phase ($\sigma \propto \tau^{-3/2}$) from flicker phase ($\sigma \propto \tau^{-1}$). The corresponding asymptotic $\sigma_y(\tau)$ dependencies, τ^{-1} and $\tau^{-1}\sqrt{\log(a\tau)}$ for some a, can barely be distinguished in practice.

It would seem from (1) that the extra averaging operation that gives MVAR its superior power of discrimination also multiplies the amount of calculation by a factor of m. Previous papers [2], [7], which treat the mechanics of MVAR computation, show how to reduce that factor to 4/3, excluding an initial operation on the data set. The approach given in [7] reformulates the definition of MVAR in terms of third differences of the cumulative sum of the time residuals. Here, after restating this formulation, we apply it to the study of the confidence of estimators of MVAR in terms of their equivalent degrees of freedom (edf). Tractable expressions for edf in the presence of power-law noise allow extensive numerical trials of estimator parameters, especially the estimation period, the amount by which the estimator summands are shifted in time. The outcome is a practical guideline for estimator design. Simple approximations to the edf of these estimators are constructed and tested, with the aim of providing a convenient package for computing approximate confidence values for most experimental situations. Finally, we show how to obtain conservative confidence values in the presence of phase noise whose spectrum is a sum of power laws.

II. MVAR and its Estimators

A. Third-Difference Formulation

The definition, calculation, and statistical theory of modified Allan variance are all simplified by an approach that derives MVAR from the cumulative sum of the time residuals x_n. We begin with the standard formulation. Choose an averaging time $\tau = m\tau_0$, and form the time-residual moving averages

Manuscript received April 3, 1995; revised December 1, 1995. This work was performed by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

The author is with the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109 USA (e-mail: cgreen@horology.jpl.nasa.gov).

Publisher Item Identifier S 0018-9456(97)03608-5.

$\overline{x}_n(m)$ from (1). Let Δ_m be the backward difference operator, defined by $\Delta_m f_n = f_n - f_{n-m}$ for any sequence f_n. Use the second-difference operator Δ_m^2 to form the MVAR filter output

$$\begin{aligned} z_n(m) &= \Delta_m^2 \overline{x}_n(m) \\ &= \overline{x}_n(m) - 2\overline{x}_{n-m}(m) + \overline{x}_{n-2m}(m). \end{aligned} \tag{2}$$

By definition

$$\mathrm{mod}\ \sigma_y^2(\tau) = \frac{1}{2\tau^2}\langle z_n^2(m)\rangle \tag{3}$$

where $\langle\ \rangle$ denotes either mathematical expectation E or an infinite time average over n. Although only the variable τ appears, mod $\sigma_y^2(\tau)$ actually depends on both τ and τ_0. For brevity, we shall occasionally suppress the dependence of $z_n(m)$ on the parameter m.

The third-difference formulation expresses $z_n(m)$ in terms of the sequence w_n defined by

$$w_0 = 0, \quad w_n = \sum_{j=1}^{n} x_j. \tag{4}$$

In terms of w_n, the time-residual averages are given by

$$\overline{x}_n(m) = \frac{1}{m}\Delta_m w_n = \frac{1}{m}(w_n - w_{n-m}) \qquad n \geq m$$

which, combined with (2), gives

$$\begin{aligned} z_n(m) &= \frac{1}{m}\Delta_m^3 w_n \\ &= \frac{1}{m}(w_n - 3w_{n-m} + 3w_{n-2m} - w_{n-3m}) \end{aligned} \tag{5}$$

for $n \geq 3m$.

Formula (5) has several advantages over (2) for use in (3). The filter taking w_n to $z_n(m)$ has only four taps; the filter taking x_n to $z_n(m)$ has $3m$ taps. The computation of estimates of mod $\sigma_y^2(\tau)$ from third differences of w_n is like the computation of estimates of $\sigma_y^2(\tau)$ from second differences of x_n, and the computation of strided estimates is simplified. Finally, it is easy to construct useful and tractable stochastic models of the w_n sequence. The cost of these advantages is the computation of w_n from the recursion $w_n = w_{n-1} + x_n$.

B. MVAR Estimator with Variable Stride

To estimate MVAR with limited data, the infinite average in (3) is replaced by a finite average of the $z_n^2(m)$. When computing analogous estimates of AVAR by averaging the squares of $\Delta_m^2 x_n$, it is customary to increase n by either 1 (full overlap) or m (τ overlap). The existing literature on MVAR ([1], for example) usually assumes a step of 1. Here, we allow the step to vary between these extremes. Let us establish some terminology. We specify an *estimation period* $\tau_1 = m_1\tau_0$, where the positive integer m_1 is called the *estimation stride*, and we consider averages over all available values of $z_{3m+km_1}^2(m), k \geq 0$.

Assume that N time residuals $x_1, x_2, \cdots, x_N$ are available. Then there are $N+1$ summed values $w_0, w_1, w_2, \cdots, w_N$. Let M be the number of samples of $z_{3m+km_1}(m)$ obtainable from (5). Then M is the largest integer satisfying $3m + (M-1)m_1 \leq N$, namely,

$$M = \left\lfloor \frac{N - 3m + m_1}{m_1} \right\rfloor \tag{6}$$

where $\lfloor a \rfloor$ denotes the integer part of a. The MVAR estimator to be studied is

$$V = \frac{1}{2\tau^2 M}\sum_{k=0}^{M-1} z_{3m+km_1}^2(m). \tag{7}$$

C. Continuous-Time Analog

A continuous-time analog of this setup yields simple and useful approximations. It is convenient to change the definitions, not only of the underlying noise processes (see below), but also of MVAR and V, by changing discrete-time averages to continuous-time averages. The third-difference approach works here, as well. Let $x(t)$ represent time deviation as a function of time. Write

$$\overline{x}(t;\tau) = \frac{1}{\tau}\int_0^\tau x(t-u)\,du,$$

$$z(t;\tau) = \Delta_\tau^2 \overline{x}(t;\tau), \quad w(t) = \int_0^t x(u)\,du.$$

Then

$$\overline{x}(t;\tau) = \frac{1}{\tau}\Delta_\tau w(t)$$

and hence

$$z(t;\tau) = \frac{1}{\tau}\Delta_\tau^3 w(t). \tag{8}$$

Define the continuous-time analog of mod $\sigma_y^2(\tau)$ by

$$\mathrm{mod}^c\ \sigma_y^2(\tau) = \frac{1}{2\tau^2}\langle z^2(t;\tau)\rangle \tag{9}$$

(identical to Bernier's asymptotic MVAR [3]), and the continuous-averaging estimator V^c by

$$V^c = \frac{1}{2\tau^2 T}\int_0^T z^2(t;\tau)\,dt. \tag{10}$$

Note that if $x(t)$ is available for a duration T_x, then we should let $T = T_x - 3\tau$, the duration of availability of $z(t;\tau)$. Later, to match properties of V^c to those of V, we shall let $T = M\tau_1$, where M is given by (6).

III. Noise Models

The statistical properties of V depend on the random processes chosen to represent the sampled time residuals x_n. Following Walter's treatment of discrete sampling [15], we use an explicit discrete-time power-law model instead of a sampled continuous-time model for our main calculations. This has two advantages. First, we avoid the complications of the interactions among the hardware bandwidth, the sample period, and the averaging time [3], [13]. Second, the discrete-time model works especially well with the third-difference formulation.

TABLE I
GENERALIZED AUTOCOVARIANCE OF w FOR DISCRETE-TIME AND CONTINUOUS-TIME POWER-LAW PHASE NOISE

$L_0 = 0, \quad L_n = \sum_{j=1}^{|n|} \frac{1}{j-1/2}$

noise type	β	$R_w^{\mathrm{d}}(n)$	$R_w^{\mathrm{c}}(t)$
white phase	0	$-\frac{1}{2\tau_0}\,\lvert n\rvert$	$-\frac{1}{2}\,\lvert t\rvert$
flicker phase	-1	$\frac{1}{2\pi\tau_0}\left(n^2-\frac{1}{4}\right)L_n$	$\frac{1}{2\pi}t^2\log\lvert t\rvert$
white freq	-2	$\frac{1}{12\tau_0}\left(n^2-1\right)\lvert n\rvert$	$\frac{1}{12}\,\lvert t\rvert^3$
flicker freq	-3	$-\frac{1}{24\pi\tau_0}\left(n^2-\frac{1}{4}\right)\left(n^2-\frac{9}{4}\right)L_n$	$-\frac{1}{24\pi}t^4\log\lvert t\rvert$
rand-walk freq	-4	$-\frac{1}{240\tau_0}\left(n^2-1\right)\left(n^2-4\right)\lvert n\rvert$	$-\frac{1}{240}\,\lvert t\rvert^5$
	nonintegral	$-\frac{1}{2\tau_0\cos(\pi\beta/2)\Gamma(2-\beta)}\frac{\Gamma(1-\beta/2+n)}{\Gamma(\beta/2+n)}$	$-\frac{1}{2\cos(\pi\beta/2)\Gamma(2-\beta)}\,\lvert t\rvert^{1-\beta}$

Because the measure of estimator confidence to be examined is invariant to scale factors, we use the most convenient scaling for spectral densities to reduce the complexity of constant factors in the generalized autocovariances shown in Table I. Factors for converting to the standard scaling used by the frequency and time community are given below.

The most critical assumption about the models is the absence of linear frequency drift. We assume that the drift rate either is zero or is known from considerations external to the immediate data set. In the latter case, we can assume that the drift has been removed from the data. In particular, x_n has no long-term quadratic component, w_n has no long-term cubic component, and $z_n(m)$ has zero mean. This assumption will later be repeated at the point where it is needed.

A. Discrete-Time Power Laws

Let the two-sided spectral density of the τ_0-sampled sequence x_n be given by

$$S_x^d(f) = |2\,\sin(\pi f\tau_0)|^\beta, \quad |f| \le \frac{1}{2\tau_0}. \tag{11}$$

Then $S_x^d(f) \sim |2\pi f\tau_0|^\beta$ as $\pi f\tau_0 \to 0$. These so-called *fractional-difference processes* were described by Granger and Joyeux [8] and by Hosking [9]. Because the first difference of the process w_n defined by (4) is just x_n, we know that w_n is also a fractional-difference process, with spectral density

$$S_w^d(f) = |2\,\sin(\pi f\tau_0)|^{\beta-2}. \tag{12}$$

This frequency-domain description of w_n has an equivalent time-domain description, the *generalized autocovariance* (GACV) sequence $R_w^d(n)$, where n runs through all the integers. If w_n were stationary, then its ordinary autocovariance (ACV) could be derived as the Fourier transform of (12). For the range of β appropriate to this application (-4 to 0), w_n is not stationary, but does have stationary third increments. With some care, one can extend the notion of ACV to the class of processes with stationary dth increments in such a way that their covariance properties can conveniently be described in terms of a function, the GACV, that still depends on *one* discrete time variable. Although the GACV itself cannot be regarded as a covariance in the usual sense, under certain restrictions it can be *used* like one. GACV's of continuous-time and discrete-time processes have already been used in studies of Allan variance and power-law noise simulation [4]–[6], [11]. A continous-time version of the GACV theory has been published [5]. Here, we can only give hints of the discrete-time theory, which is similar.

Table I gives formulas for $R_w^d(n)$ for the values of β needed in this study. Bear in mind that the noise-type label applies to x_n, a power-law process with exponent β, while $R_w^d(n)$ applies to w_n, a power-law process with exponent $\beta-2$. The formula for nonintegral β in Table I is the same as the one derived for fractional-difference processes by others [8], [9], [11]. It has been verified that this formula actually does extend to the nonstationary situation. Because passage to the limit of the GACV as β approaches an integer is unfortunately not straightforward in general, the formulas for integral β were derived from known ACV's of stationary fractional-difference processes by repeated solution of difference equations of form $-\delta_1^2 R(n;\gamma) = R(n;\gamma+2)$, where $R(n;\gamma)$ is the ACV or GACV of a fractional-difference process with exponent γ, and δ_1^2 is the second-order central difference operator with step 1.

For $\beta > -5$, the process $z_n(m)$ defined by (5) is stationary, and the GACV theory allows its ordinary ACV sequence

$$R_z^d(n;m) = \mathrm{E}[z_{k+n}(m)z_k(m)]$$

to be calculated directly from $R_w^d(n)$ by

$$\begin{aligned} m^2 R_z^d(n;m) &= -\delta_m^6 R_w^d(n) \\ &= -R_w^d(n-3m) + 6R_w^d(n-2m) \\ &\quad - 15R_w^d(n-m) + 20R_w^d(n) \\ &\quad - 15R_w^d(n+m) + 6R_w^d(n+2m) \\ &\quad - R_w^d(n+3m). \end{aligned} \tag{13}$$

The central difference operator $-\delta_m^6$ appears as the operator product $\Delta_m^3\Delta_{-m}^3$.

It is appropriate to note here that (3), (13), and Table I lead to a formula for MVAR in the presence of fractional-difference phase noise, namely,

$$\mathrm{mod}\;\sigma_y^2(\tau) = \frac{1}{2\tau^2}\;\mathrm{E}\;z_n^2(m) = \frac{1}{2\tau^2}R_z^d(0;m) \tag{14}$$

which, when expanded by (13), is equivalent to a formula of Walter [15, eq. (75)] that was derived from a frequency-domain integral.

The standard power-law scaling used by the frequency and time community is based on a one-sided spectral density, $S_y^+(f) \sim h_\alpha f^\alpha$, of fractional frequency $y = dx/dt$, where

$\alpha = \beta + 2$. To convert $R_w^d(n), R_z^d(n;m)$, and mod $\sigma_y^2(\tau)$ to this scaling, multiply them by the factor

$$\frac{h_\alpha}{2}(2\pi)^{-\alpha}\tau_0^{2-\alpha}. \tag{15}$$

B. Continuous-Time Power Laws

Because the continous-time analog given above avoids sampling completely, continuous-time random-process models are appropriate. Let the two-sided spectral density of $x(t)$ be given by

$$S_x^c(f) = |2\pi f|^\beta. \tag{16}$$

Then, since $dw/dt = x$, we know that $w(t)$ is also a power-law process, with spectral density

$$S_w^c(f) = |2\pi f|^{\beta-2}.$$

For $\beta > -5$, the process $w(t)$ has stationary third increments. Moreover, for $\beta < 1, S_w^c(f)$ integrates to a finite value over any frequency range that excludes an interval about zero. Therefore, a high-frequency cutoff is unnecessary. Its GACV function $R_w^c(t)$ [4], [11] is also given in Table I. As with the discrete-time model, the process $z(t)$ given by (8) is stationary, with ACV function

$$R_z^c(t;\tau) = \mathrm{E}[z(u+t;\tau)z(u;\tau)]$$

that can be calculated from $R_w^c(t)$ by

$$\begin{aligned}\tau^2 R_z^c(t;\tau) &= -\delta_\tau^6 R_w^c(t)\\ &= -R_w^c(t-3\tau) + 6R_w^c(t-2\tau)\\ &\quad - 15R_w^c(t-\tau) + 20R_w^c(t)\\ &\quad - 15R_w^c(t+\tau) + 6R_w^c(t+2\tau)\\ &\quad - R_w^c(t+3\tau).\end{aligned} \tag{17}$$

A formula for mod^c $\sigma_y^2(\tau)$, analogous to (14), is

$$\mathrm{mod}^c\ \sigma_y^2(\tau) = \frac{1}{2\tau^2}R_z^c(0;\tau). \tag{18}$$

Substituting $R_w^c(t)$ from Table I into (17), we find from (18) that mod^c $\sigma_y^2(\tau)$ is *exactly* proportional to $\tau^{-3-\beta}$, for $-5 < \beta < 1$. The same result was derived by Bernier [3] from a frequency-domain integral.

The factor for converting $R_w^c(t), R_z^c(t;\tau)$, and mod^c $\sigma_y^2(\tau)$ to standard frequency and time scaling is the same as (15), with τ_0 replaced by 1.

IV. Equivalent Degrees of Freedom

By definition, the equivalent degrees of freedom (edf) of a positive random variable X is defined by

$$\mathrm{edf}\ X = \frac{2(\mathrm{E}[X])^2}{\mathrm{var}\ X} \tag{19}$$

where var X denotes the variance of X. If X is distributed as a constant multiple of a χ_ν^2 random variable, with ν degrees of freedom, then edf $X = \nu$. For example, the sample variance of n independent, identically distributed Gaussians has $n-1$ degrees of freedom. Even if X does not have a chi-squared distribution, edf X can still serve as a convenient dimensionless measure of the confidence of X as an estimator of its mean E$[X]$: we can interpret edf X as the degrees of freedom of the chi-squared distribution that has the same ratio of mean to standard deviation. Since the MVAR estimator V is the sum of squares of *correlated* zero-mean Gaussians, it is reasonable to assume that V is *approximately* distributed as const $\cdot\, \chi^2_{\mathrm{edf}\ V}$, and, on this basis, to construct approximate confidence intervals for mod $\sigma_y(\tau)$ [10], [18].

A. Discrete Time

Let us compute edf V. By (7) and (14)

$$\mathrm{E}[V] = \frac{1}{2\tau^2}R_z^d(0;m) \tag{20}$$

that is, V is unbiased for mod $\sigma_y^2(\tau)$. Also from (7) we have

$$\mathrm{var}\ V = \frac{1}{(2\tau^2 M)^2}\sum_{i=0}^{M-1}\sum_{j=0}^{M-1}\mathrm{cov}(z_{3m+im_1}^2, z_{3m+jm_1}^2) \tag{21}$$

where cov (X,Y) denotes the covariance of the random variables X and Y. To compute the covariances in (21), we assume that $z_n(m)$ is a stationary Gaussian *zero-mean* process. As indicated earlier, the assumption E $[z_n(m)] = 0$ is crucial; in practice, it means that the effect of linear frequency drift on a time scale of order τ is negligible. Since any two jointly Gaussian zero-mean random variables X and Y satisfy cov $(X^2,Y^2) = 2(\mathrm{E}\ [XY])^2$, (21) becomes

$$\mathrm{var}\ V = \frac{2}{(2\tau^2 M)^2}\sum_{i=0}^{M-1}\sum_{j=0}^{M-1}[R_z^d((i-j)m_1;m)]^2. \tag{22}$$

The diagonal $i-j=k$, for $k = 1-M$ to $M-1$, contains $M-|k|$ identical terms. Summing over these diagonals converts the double sum to a single sum, which, combined with (19) and (20), gives the main formula

$$\frac{1}{\mathrm{edf}\ V} = \frac{1}{M}\left[1 + 2\sum_{k=1}^{M-1}\left(1-\frac{k}{M}\right)(\rho_z^d(km_1;m))^2\right] \tag{23}$$

for edf V, where

$$\rho_z^d(n;m) = \frac{R_z^d(n;m)}{R_z^d(0;m)}.$$

Formula (23) is analogous to existing formulas for the edf of AVAR estimators ([6] and references therein). The main difference is that the ACV of z is computed from sixth differences of the GACV of w instead of fourth differences of the GACV of x.

Formula (23) is mathematically equivalent to an earlier formula of Walter [15, (eq. 32)], but requires less computation. Recall from (13) that each value of $R_z^d(n;m)$ needed in (23) is obtained from seven values of $R_w^d(n)$. If no values of $R_w^d(n)$ are stored in advance, it takes $7M$ evaluations of $R_w^d(n)$ to compute (23). Walter's formula for var V is a double sum requiring $5(2m-1)\ (2M-1)$ evaluations of $R_x^d(n)$. In practice, moreover, one can compute and store the values

$R_w^d(n), |n| \le N$, in advance. This shows the advantage of the third-difference approach, which derives MVAR estimator summands from four values of w_n instead of $3m$ values of x_n.

In connection with a recent conference paper [18], tables of edf V for $m_1 = 1$ were generated by the method given here, by Walter's method, and by Monte Carlo simulation. The results of the two theoretical methods agreed within 0.1%; the simulation results agreed with the theoretical results within a few percent.

A note on numerical computation. The ACV $R_z^d(n;m)$ tends to zero as $n \to \infty$, yet is obtained from differences of $R_w^d(n)$, which tends to ∞ with n. Clearly, one should use double precision for evaluating (13). Even so, the computed values of $R_z^d(n;m)$ can deteriorate for large n, especially for nonintegral β, where $R_w^d(n)$ involves Γ functions. I was able to cure this problem by replacing the upper limit $M-1$ of the summation in (23) by $K-1$, where K is the smaller of M and $10m/m_1$. (In all actual computations, m/m_1 is assumed to be an integer.)

B. Continuous Time

The computation of edf V^c follows the same pattern. By (10)

$$\mathrm{E}[V^c] = \frac{1}{2\tau^2} R_z^c(0;\tau)$$

and, with the assumption that $z(t;\tau)$, as a function of t, is a stationary Gaussian zero-mean process,

$$\mathrm{var}\ V^c = \frac{2}{(2\tau^2 T)^2} \int_0^T \int_0^T [R_z^c(t-u;\tau)]^2\, dt\, du.$$

A change of variables converts the double integral to

$$2\int_0^T (T-t)(R_z^c(t;\tau))^2\, dt$$

in which we shall make the further change of variable $t = \tau x$. From Table I and (17), it can be verified that

$$\frac{R_z^c(\tau x;\tau)}{R_z^c(0;\tau)} = \frac{R_z^c(x;1)}{R_z^c(0;1)}.$$

(This is a scaling property of continuous-time power-law noise.) Thus, defining

$$\rho_z^c(x) = \frac{R_z^c(x;1)}{R_z^c(0;1)}$$

we obtain

$$\frac{1}{\mathrm{edf}\ V^c} = \frac{2}{p}\int_0^p \left(1 - \frac{x}{p}\right)(\rho_z^c(x))^2\, dx \tag{24}$$

where $p = T/\tau$.

V. Effect of Estimation Period

Formula (23) was used to generate tables of edf V for combinations of N, m, m_1, and noise exponent β. Recall that

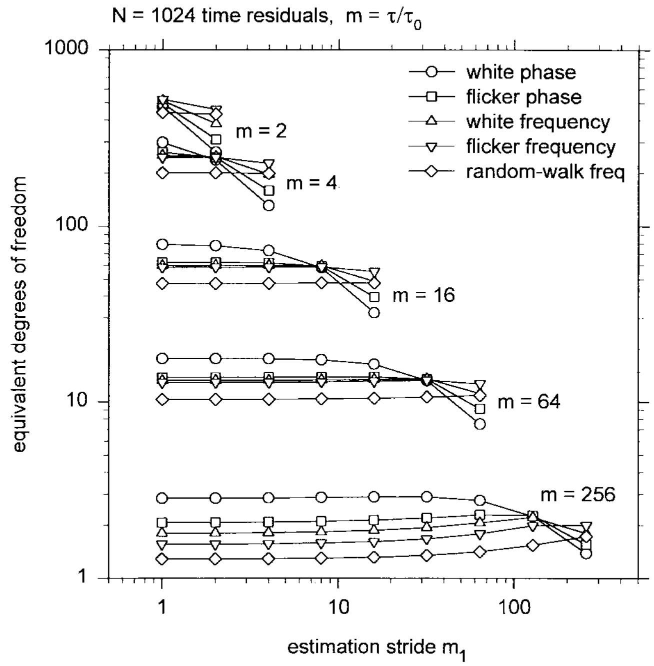


Fig. 1. Equivalent degrees of freedom of MVAR estimators.

N is the number of time residuals in the data set, $m = \tau/\tau_0$, where τ is the averaging time, and $m_1 = \tau_1/\tau_0$, the estimation stride, where τ_1 is the estimation period. From here on, we also assume the *divisibility condition*, which says that the estimation period divides evenly into the averaging time, that is,

$$\frac{\tau}{\tau_1} = \frac{m}{m_1} = r$$

where r is an integer. Thus, the estimation stride m_1 is restricted to divisors of m. This condition allows V and edf V to be calculated from the subsampled arrays w_{jm_1} and $R_w^d(jm_1)$, respectively. For each (N, m, m_1) combination, the number M of estimation summands to be used in (23) is calculated by (6).

A selection of edf values is plotted in Fig. 1 for 1024 time residuals and the five standard phase-noise types. Observe how edf depends on m_1 for fixed m. For each noise type and $m \ge 4$, any m_1 between 1 and $m/4$ gives a value of edf that is nearly maximal for that m. As the two-point curves for $m = 2$ show, we should take $m_1 = 1$ in this case; the same is true for $m = 3$. Here is an empirical result.

Assume an averaging time τ at most 1/4th the duration of the time-deviation record. For each power law between white phase and random-walk frequency, any estimation period τ_1 between τ_0 and $\max(\tau_0, \tau/4)$ that divides evenly into τ gives an MVAR *estimator V whose edf is within 8% of the maximal value for τ.*

Fig. 1 shows that the variation of edf V with m_1 is greatest for white phase. Also, it turns out that the quantity

$$p = \frac{M}{r} = \frac{Mm_1}{m} = \frac{M\tau_1}{\tau} \tag{25}$$

is a rough estimate of edf V, especially for m_1 in the recommended range $1 \leq m_1 \leq \max(1, m/4)$.

The choice of estimation period τ_1 might depend on a tradeoff between convenience and computational effort. For simplicity, one can always choose $\tau_1 = \tau_0$. If the data set is large, one can choose the largest acceptable value, $\tau_1 = \tau/4$, to minimize the number M of terms needed to calculate V from (7).

VI. Lower Bounds for MVAR EDF

The aim of this section is to uncover simple approximation formulas for edf V that can be used in practice in place of the exact summation (23). There are two rigorous lower-bound formulas that can serve this purpose.

A. Discrete Time

Up to now, we have concentrated on a time-domain formulation of edf V. The following result is proved by a frequency-domain argument, which is not given here.

Theorem 1: Assume that the time residuals x_n, with sample period τ_0, are a process with stationary Gaussian zero-mean second increments. Let x_n have the fractional-difference spectrum (11), where $-9/2 < \beta \leq 0$. Let $m = m_1 r$, where m_1 and r are positive integers. Using (4), (5), (7), and any positive integer M, form the MVAR estimator V with averaging time $\tau = m\tau_0$ and estimation period $\tau_1 = m_1\tau_0$. Then

$$\text{edf } V \geq \frac{M\tau_1}{\tau}\,\frac{2I^2}{J} \tag{26}$$

where

$$I = \int_0^{m/2} \frac{\sin^6(\pi x)}{[m\,\sin(\pi x/m)]^{2-\beta}}\,dx,$$
$$J = \int_0^{r/2} \frac{\sin^{12}(\pi x)}{[r\,\sin(\pi x/r)]^{4-2\beta}}\,dx.$$

In other words, we have a bound of form edf $V \geq ap$, where $p = M/r$ as given above in (25). Tables of a versus β, m, and r can be generated by numerical integration.

B. Continuous Time

It is much easier to derive a useful lower bound for edf V^c. Let $p \geq 2$. From (24) we have

$$\begin{aligned}\frac{1}{\text{edf } V^c} &= \frac{2}{p}\left[\int_0^p (\rho_z^c(x))^2\,dx - \frac{1}{p}\int_0^p x(\rho_z^c(x))^2\,dx\right]\\ &\leq \frac{2}{p}\left[\int_0^\infty (\rho_z^c(x))^2\,dx - \frac{1}{p}\int_0^2 x(\rho_z^c(x))^2\,dx\right].\end{aligned}$$

This gives a bound of form

$$\text{edf } V^c \geq \frac{a_0 p}{1 - \dfrac{a_1}{p}}, \qquad p \geq 2. \tag{27}$$

TABLE II
Coefficients for Approximating the EDF of MVAR Estimators

noise type	β	$m=1$: a_0	$m=1$: a_1	$m=2$: a_0	$m=2$: a_1	$m>2$: a_0	$m>2$: a_1
white phase	0	.514	0	.935	0	1.225	.589
	$-1/2$	.543		.954		1.074	.596
flicker phase	-1	.576		.973		1.003	.602
	$-3/2$	.617		.992		.977	.598
white freq	-2	.667		1.010		.968	.571
	$-5/2$	.729		1.024		.961	.510
flicker freq	-3	.811		1.027		.947	.416
	$-7/2$	.914		1.000		.906	.343
rand-walk freq	-4	1.000		.866		.768	.411

The constants a_0 and a_1, which depend only on β, are computed by numerical integration. To use this expression as an approximation to edf V, we again let $p = M/r$.

C. An EDF Approximation Strategy

The right sides of (26) and (27) can be regarded as candidate approximations for edf V. To assess their quality and to choose between them, tables were generated for a selection of N, m, m_1, and β. The following empirical strategy and error statement emerged.

Assume fractional-difference phase noise with power-law exponent between -4 (random-walk frequency) and 0 (white phase), at least 16 time-residual points, an averaging time τ at most 1/5th the duration of the measurement, and an estimation period τ_1 between τ_0 and $\max(\tau_0, \tau/4)$ that divides evenly into τ. In our notation, $-4 \leq \beta \leq 0, N \geq 16, \leq N/5$, and $m = rm_1$, where r is an integer, and $1 \leq m_1 \leq \max(1, m/4)$.

For $m = 1$ or 2, the discrete-time lower bound (26) is used as an approximation for edf V. In all other cases, the continuous-time lower bound (27) is used. The relative error of this strategy is observed to be at most ±11.1%.

To implement this approximation in practice, use the formula

$$\text{edf } V \approx \frac{a_0 p}{1 - \dfrac{a_1}{p}} \tag{28}$$

where $p = Mm_1/m, M$ is obtained from (6), and the coefficients a_0, a_1, as functions of m and β, are drawn from Table II.

Table III shows the percentage errors of this approximation (100(approx/exact -1)) for a selection of N, m, m_1, and β. The full range of observed errors is represented. To balance the errors, it was found expedient to reduce the continuous-time edf approximation, for white phase only, by 5%. Tables II and III include this adjustment.

VII. Compound Noise Spectra

The previous results and methods assume a power-law phase noise spectrum proportional to (11), for some fixed exponent β. If that were indeed the case, our statistical efforts ought to be directed toward estimating the two-parameter set consisting

TABLE III
PERCENTAGE ERRORS OF MVAR ESTIMATOR edf APPROXIMATION

			β				
N	m	m_1	0	-1	-2	-3	-4
1024	1	1	$+0.0$	$+0.0$	$+0.0$	$+0.0$	$+0.0$
	2	1	-0.1	-0.1	-0.1	-0.1	-0.1
	3	1	$+11.1$	-2.8	-3.9	-4.1	-5.0
	16	1	-4.3	-0.7	-0.2	-0.3	-0.2
		2	-2.6	-0.6	-0.2	-0.3	-0.2
		4	$+4.1$	$+0.1$	-0.2	-0.3	-0.2
	128	1	-5.9	-0.4	-0.4	-2.6	$+0.0$
		2	-5.9	-0.4	-0.4	-2.6	$+0.0$
		4	-5.8	-0.4	-0.4	-2.6	$+0.0$
		8	-5.3	-0.4	-0.4	-2.6	$+0.0$
		16	-3.6	-0.3	-0.3	-2.5	$+0.0$
		32	$+3.4$	$+0.4$	-0.1	-2.3	$+0.2$
16	1	1	-3.7	-3.1	-2.4	-1.4	$+0.0$
	2	1	-10.6	-10.0	-9.2	-8.1	-5.9
	3	1	$+9.9$	-2.0	-3.0	-7.2	-3.5

of β and the constant of proportionality. Instead, as usual, we find ourselves using parametric tools to evaluate the confidence of a nonparametric statistic. The value of edf V depends on β. What can we do in the presence of a polynomial phase noise model

$$S_x(f) = \sum_{\beta} g_\beta |\sin(2\pi f \tau_0)|^\beta \tag{29}$$

a finite sum of power-law spectra? Some help is given by the following theorem, which, although weak and perhaps obvious, is better than no knowledge at all about the situation.

Theorem 2: Let the phase noise be a finite sum of independent component noises with stationary Gaussian zero-mean second increments. Form an MVAR *estimator* V from the given phase noise, and corresponding estimators V_k from the components. Then

$$\text{edf } V \geq \min_k \text{ edf } V_k.$$

In other words, we never do worse than the worst component.

To apply this theorem to the situation (29), assume that the component β values are all in some subinterval of [-4, 0] (the whole range, perhaps). Use (28) and Table II to compute edf V_β for each tabulated β in the subinterval, and take the smallest value as a *conservative* estimate of edf V. For example, if one believes that the noise has components between white phase and flicker phase, perhaps from prior knowledge, perhaps as evidenced by a log-log plot of mod σ versus τ with slopes between $-3/2$ and -1, then one can minimize (28) over the first three rows of Table II.

The proof of Theorem 2, although not difficult, is not given here. It can be generalized to AVAR estimators and other situations involving averages of the square of a stationary Gaussian zero-mean process. Its usefulness for MVAR is enhanced by the relatively weak dependence of estimator edf on β, as can be seen from Fig. 1. An inspection of edf tables for fully overlapped AVAR estimators [6], [14] shows a much sharper dependence on β, especially for large τ/τ_0. Thus, minimizing over a set of β in the computation of estimator edf causes a smaller loss of accuracy for MVAR than for AVAR.

VIII. CONCLUSIONS

Although the overall problem of estimating modified Allan variance MVAR may appear to be more difficult than the same problem for conventional Allan variance AVAR, theoretical and numerical results calculated here from the third-difference approach show that in some ways the situation is actually reversed. A tractable expression for the edf of MVAR estimators in the presence of power-law phase noise was derived, and simple approximations constructed. Numerical computations of edf yielded a rationale for choosing the estimation period or stride: it was found empirically that the use of an estimation period up to one-fourth the averaging time does not appreciably degrade the confidence of the estimator below that of the fully overlapped estimator. Often, in fact, there is no degradation. The computations also revealed that the extra filtering inherent in MVAR causes the edf of an estimator to be less sensitive to the power-law exponent than the edf of a typical AVAR estimator. Consequently, MVAR error bars can be more robust against spectrum uncertainties than AVAR error bars.

The most important limitation on these results, especially for long tests of oscillators, is that linear frequency drift must be negligible. If a drift rate is known from considerations external to the immediate data set, then one can remove it from the phase data, and we are back to the case of zero drift. For AVAR, it is known that estimation of drift from the data themselves, and removal therefrom, cause negative AVAR estimator biases that worsen as averaging time τ increases. The use of three-point [16], [17] or four-point [4] drift estimators, which extract a quadratic component of the time-residual sequence x_n, simplifies calculations of the mean and variance of estimators of AVAR with drift removed. I have no doubt that similar calculations for MVAR estimators can be made on the basis of four-point drift estimators that extract a *cubic* component of the sequence w_n of cumulative sums of x_n.

REFERENCES

[1] D. W. Allan, "Time and frequency (time-domain) characterization, estimation, and prediction of precision clocks and oscillators," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. UFFC-34, pp. 647–654, 1987.

[2] ——, "Time and frequency metrology: current status and future considerations," in *Proc. 5th Europ. Frequency Time Forum*, 1991, pp. 1–9.

[3] L. G. Bernier, "Theoretical analysis of the modified Allan variance," in *Proc. 41st Annu. Frequency Control Symp.*, 1987, pp. 116–121.

[4] C. A. Greenhall, "The fundamental structure function of oscillator noise models," in *Proc. 14th Precise Time, Time Interval Applications Planning Meet.*, 1982, pp. 281–294.

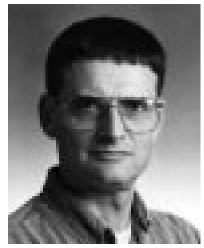
[5] ——, "A structure function representation theorem with applications to frequency stability estimation," *IEEE Trans. Instrum. Meas.*, vol IM-32, pp. 364–370, 1983.

[6] ——, "Recipes for degrees of freedom of frequency stability estimators," *IEEE Trans. Instrum. Meas.*, vol 40, pp. 994–999, 1991.

[7] ——, "A shortcut for estimating the modified Allan variance," in *Proc. 1992 IEEE Frequency Control Symp.*, pp. 262–264.

[8] C. W. J. Granger and R. Joyeux, "An introduction to long-memory time series models and fractional differencing," *J. Time Series Anal.*, vol. 1, pp. 15–29, 1980.

[9] J. R. M. Hosking, "Fractional differencing," *Biometrika*, vol 68, pp. 165–176, 1981.
[10] D. A. Howe, D. W. Allan, and J. A. Barnes, "Properties of signal sources and measurement methods," in *Proc. 35th Annu. Frequency Control Symp.*, 1981, pp 1–47.
[11] N. J. Kasdin and T. Walter, "Discrete simulation of power law noise," in *Proc. 1992 IEEE Frequency Control Symp.*, pp. 274–283.
[12] P. Lesage and C. Audoin, "Characterization and measurement of time and frequency stability," *Radio Sci.*, vol. 14, pp. 521–539, 1979.
[13] J. Rutman and F. L. Walls, "Characterization of frequency stability in precision frequency sources," *Proc. IEEE*, vol. 79, pp. 952–960, 1991.
[14] S. R. Stein, "Frequency and time—their measurement and characterization," in *Precision Frequency Control*, E. A. Gerber and A. Ballato, Eds. New York: Academic, 1985, vol. 2, pp. 191–232, 399–416.
[15] T. Walter, "Characterizing frequency stability: a continuous power-law model with discrete sampling," *IEEE Trans. Instrum. Meas.*, vol 43, pp. 69–79, 1994.
[16] M. A. Weiss, D. W. Allan, and D. A. Howe, "Confidence on the second difference estimation of frequency drift," in *Proc. 1992 IEEE Frequency Control Symp.*, pp. 300–305.
[17] M. A. Weiss and C. Hackman, "Confidence on the three-point estimator of frequency drift," in *Proc. 24th Precise Time, Time Interval Applications Planning Meet.*, 1992, pp. 451–460.
[18] M. A. Weiss, F. L. Walls, C. A. Greenhall, and T. Walter, "Confidence on the modified Allan variance and the time variance," *9th Europ. Frequency Time Forum*, 1995.



Charles A. Greenhall (M'92) was born in New York City in 1939. He received the B.S. degree in physics from Pomona College, Claremont, CA, and the Ph.D. degree in mathematics from the California Institute of Technology, Pasadena, in 1966 with a dissertation on Fourier series.

He spent five years as an Assistant Professor of Mathematics at the University of Southern California, Los Angeles. Since 1973, he has worked at the Jet Propulsion Laboratory (JPL), California Institute of Technology, where he has engaged in the study and practice of clock measurement and statistics at the JPL Frequency Standards Laboratory since 1981.

Dr. Greenhall belongs to the American Mathematical Society.