

Uncertainty of Stability Variances Based on Finite Differences

C. A. Greenhall*
Jet Propulsion Laboratory
California Institute of Technology

W. J. Riley
Symmetricom, Inc.

October 13, 2003

Abstract

This work gives an algorithm for computing the degrees of freedom of estimators of Allan and Hadamard variances, including their modified versions. A consistent approach is used throughout.

1 Introduction

This work gives a method of determining error bars for measurements of frequency stability variances in the presence of power-law phase noises. For Allan variance and modified Allan variance, this subject has been studied extensively for about thirty years ([1]–[13]), although no such work on the Hadamard variance has been published. Recognizing that these variances are all special cases of a single general form, we have found a consistent approach to computing the measurement uncertainty of these variances. The goal is not closed formulas but a readily programmed algorithm that gives numerical results whose accuracy is adequate for the purpose at hand. Although the algorithm given here is divided into cases, all the calculations are based on the same theoretical principles, with no empirical formulas.

The stability variances used in the time and frequency field fall into two categories: unmodified variances, which use d th differences of phase samples for $d = 2$ or 3 , and modified variances, which use d th differences of averaged phase samples. The Allan variances correspond to $d = 2$, the Hadamard variances to $d = 3$; the modified Hadamard variance is also called the pulsar variance [14]. By modeling the differences as an ergodic stationary process, one can define the corresponding variance as a scaling factor times the expected value of the squared differences. One can then obtain unbiased estimates of this variance from available phase data by computing time averages of the squared differences. The usual choices for the estimation stride (the time step) are the sample period τ_0 and the averaging period τ , a multiple of τ_0 . These give respectively the overlapped estimator (OE) and nonoverlapped (NOE) estimator of the stability variance (although “nonoverlapped” is a misnomer; there is always some overlap except in certain studies of Allan variance [1][9] where the stride is 2τ).

We insist on maintaining the distinction between a stability variance and its estimators; for each variance we treat both the OE and NOE. This distinction is not always made in the literature,

*The work of this author was performed at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

where one often sees the modified Allan variance being defined as its OE. Even though the OE of Allan variance has lower uncertainty than the NOE (but not always [15]), the NOE is convenient when phase data are processed in real time or read sequentially from a file.

2 Scope

The algorithm covers the following situations.

- Phase noise spectrum that follows an asymptotic low-frequency power law, $S_x(f) \sim C f^{\alpha-2}$ as $f \rightarrow 0$, for integral α , $2 \geq \alpha \geq -4$ (white FM to random run FM). For each τ one must choose a dominant power law; see [16] for a method of power-law identification. The phase is assumed to be approximately bandlimited to the Nyquist frequency of the sampling period.
- Order of phase differencing $d = 1, 2$, or 3 . (We include $d = 1$ for completeness.) For a given α , d has to be high enough to make the d th phase differences stationary, that is, $2d + \alpha > 1$. Further, the d th phase differences are modeled as a mean-zero Gaussian process.
- For each d , the modified and unmodified variances.
- For each variance, the overlapped and nonoverlapped estimator.

Effects of trend removal, especially drift removal for $d = 2$, are not covered; the d th phase differences are assumed to have mean zero. One can use $d = 3$ to obtain stability results that are invariant to linear frequency drift. Special long-term stability estimators, such as total variance [17] and Theo1 [18], are not covered; these require their own treatments.

3 Theory of operation

Although the presentation of the algorithm is self-contained, here is a brief account of the theory behind it. The algorithm's output is the effective degrees of freedom (edf) of an unbiased estimator V of a stability variance $\sigma^2 = EV$. Define

$$\nu = \text{edf } V = \frac{2(EV)^2}{\text{var } V} = \frac{2\sigma^4}{\text{var } V}; \quad (1)$$

thus, $\text{edf } V$ codes the normalized uncertainty $\sqrt{\text{var } V}/EV$. For these estimators, it has been observed empirically (but not systematically) that $(\nu/\sigma^2)V$ has approximately a χ_ν^2 distribution. Having computed ν and observed V , one can obtain confidence intervals of form $\nu V/x_2 \leq \sigma^2 \leq \nu V/x_1$ from χ_ν^2 levels x_1 and x_2 [7].

The model for phase $x(t)$ is the τ_0 -difference of a pure power-law process:

$$x(t) = \Delta_{\tau_0} w(t), \quad (2)$$

where $w(t)$ is a continuous-time process with spectral density $C f^{\alpha-4}$ for all $f > 0$, and Δ is the backward difference operator. Then $S_x(f)$ is asymptotically proportional to $f^{\alpha-2}$ as $f \rightarrow 0$ and has approximate bandwidth $1/(2\tau_0)$; this is the first reason for using $w(t)$.

Now let

$$z(t) = \Delta_\tau^d \Delta_\varepsilon w(t), \quad (3)$$

where $\varepsilon = \tau_0$ or τ . For $\varepsilon = \tau_0$ we have $z(t) = \Delta_\tau^d x(t)$, which leads to the unmodified variance. For $\varepsilon = \tau = m\tau_0$ we observe from (2) that

$$\Delta_\tau w(t) = w(t) - w(t - m\tau_0) = \sum_{n=0}^{m-1} \Delta_{\tau_0} w(t - n\tau_0) = \sum_{n=0}^{m-1} x(t - n\tau_0) = m\bar{x}(t),$$

where $\bar{x}(t)$ is a τ -average value of the samples of x . Thus, $z(t) = \Delta_\tau^{d+1} w(t)$ leads to the modified variance [13]; this is the second reason for using $w(t)$. In either case we assume that $z(t)$ is a stationary mean-zero Gaussian process.

Ignoring the conventional scaling factors, we define the stability variance and its estimator by

$$\sigma^2 = \text{E}z^2(t), \quad V = \frac{1}{M} \sum_{n=1}^M z^2(n\delta), \quad (4)$$

where the stride δ is τ_0 or τ for the OE and NOE, respectively. The number of terms M depends on the estimator and the number of data. Then $\text{E}V = \sigma^2$. Let $s_z(t) = \text{E}z(u+t)z(u)$. Then $\text{cov}[z^2(t), z^2(u)] = 2s_z^2(u-t)$, and

$$\text{var} V = \frac{2}{M^2} \sum_{n_1, n_2=1}^M s_z^2((n_2 - n_1)\delta). \quad (5)$$

The definition (1), after substitution of (5), simplifies to

$$\frac{1}{\text{edf} V} = \frac{1}{M} \left[1 + \frac{2}{s_z^2(0)} \sum_{j=1}^{M-1} \left(1 - \frac{j}{M}\right) s_z^2(j\delta) \right]. \quad (6)$$

The autocovariance $s_z(t)$ is obtained from (3) by applying a difference operator of order $2d+2$ to the generalized autocovariance (GACV) $s_w(t)$ of the power-law process $w(t)$ [15]:

$$s_z(t) = (\Delta_\tau \Delta_{-\tau})^d \Delta_\varepsilon \Delta_{-\varepsilon} s_w(t).$$

The GACV of $w(t)$ is tabulated below as a function of α .

4 Algorithm for edf calculation

Our purpose is to obtain practical numerical approximations of (6). We give two versions of the algorithm: the simplified version simply truncates the sum in the exact formula; the full version maintains the number of summation terms below a presassigned threshold and avoids catastrophic roundoff errors. They have the same inputs, output, function definitions, and initial step. Some explanation of the approximations is given in appendix A. Because the results are invariant to time scaling, we may set $\tau = 1$, $\tau_0 = 1/m$.

All arithmetic is to be carried out in double-precision floating point. Operations that give signed integers are the floor function $\lfloor x \rfloor$ (greatest integer that is $\leq x$) and ceiling function $\lceil x \rceil = -\lfloor -x \rfloor$ (least integer that is $\geq x$).

4.1 Inputs

α = frequency noise exponent

$\alpha = 2, 1, 0, -1, -2, -3, -4$

Noise type = WHPM, FLPM, WHFM, FLFM, RWFM, FWFM, RRFM

d = order of phase difference

$d = 1$: first-difference variance

$d = 2$: Allan variance

$d = 3$: Hadamard variance

Restriction: $\alpha + 2d > 1$

m = averaging factor τ/τ_0 , positive integer

F = filter factor

$F = 1$: modified variance

$F = m$: unmodified variance

S = stride factor (estimator stride = τ/S)

$S = 1$: nonoverlapped estimator

$S = m$: overlapped estimator

N = number of phase data with sample period τ_0

4.2 Output

edf = effective degrees of freedom of variance estimator

4.3 Constant and function definitions

Set an integer constant J_{\max} (used only in the full version); I suggest $J_{\max} = 100$.

The formal arguments of the following functions have the same names as the input arguments of the main algorithm.

1. Define the function $s_w(t, \alpha)$ as follows:

$$s_w(t, \alpha) \begin{cases} 2 & 1 & 0 & -1 & -2 & -3 & -4 \\ -|t| & t^2 \ln |t| & |t|^3 & -t^4 \ln |t| & -|t|^5 & t^6 \ln |t| & |t|^7 \end{cases} . \quad (7)$$

The entries with $\ln |t|$ must evaluate to 0 when $t = 0$.

2. Define the function

$$\begin{aligned} s_x(t, F, \alpha) &= F^2 \Delta_{1/F} \Delta_{-1/F} s_w(t, \alpha) \\ &= F^2 \left[2s_w(t, \alpha) - s_w\left(t - \frac{1}{F}, \alpha\right) - s_w\left(t + \frac{1}{F}, \alpha\right) \right], \\ s_x(t, \infty, \alpha) &= s_w(t, \alpha + 2), \quad -4 \leq \alpha \leq 0. \end{aligned} \quad (8)$$

3. Define the function

$$s_z(t, F, \alpha, d) = (\Delta_1 \Delta_{-1})^d s_x(t, F, \alpha), \quad d = 1, 2, 3; \quad (9)$$

that is (with dependence on F and α suppressed on the right sides),

$$\begin{aligned} s_z(t, F, \alpha, 1) &= 2s_x(t) - s_x(t-1) - s_x(t+1), \\ s_z(t, F, \alpha, 2) &= 6s_x(t) - 4s_x(t-1) - 4s_x(t+1) + s_x(t-2) + s_x(t+2), \end{aligned}$$

$$s_z(t, F, \alpha, 3) = 20s_x(t) - 15s_x(t-1) - 15s_x(t+1) \\ + 6s_x(t-2) + 6s_x(t+2) - s_x(t-3) - s_x(t+3).$$

4. Define the function

$$\text{BasicSum}(J, M, S, F, \alpha, d) = s_z^2(0, F, \alpha, d) + \left(1 - \frac{J}{M}\right) s_z^2\left(\frac{J}{S}, F, \alpha, d\right) \\ + 2 \sum_{j=1}^{J-1} \left(1 - \frac{j}{M}\right) s_z^2\left(\frac{j}{S}, F, \alpha, d\right). \quad (10)$$

4.4 Initial steps for both versions

1. Compute M , the number of summands in the estimator, as follows:

$$L = m \left(\frac{1}{F} + d \right), \\ M = 1 + \left\lfloor \frac{S(N-L)}{m} \right\rfloor \quad (11)$$

if $N \geq L$, otherwise there are not enough data.

2. Let

$$J = \min(M, (d+1)S). \quad (12)$$

4.5 Main procedure, simplified version

This is just one step:

$$\frac{1}{\text{edf}} = \frac{1}{s_z^2(0, F, \alpha, d) M} \text{BasicSum}(J, M, S, F, \alpha, d). \quad (13)$$

To check the effect of the truncation, one can also try a larger value of J , say $\min(M, 6S)$.

4.6 Main procedure, full version

Let $r = \frac{M}{S}$.

There are four cases. The calculations use coefficients from three numerical tables.

4.6.1 Case 1. Modified variances: $F = 1$, all α

This case also applies to unmodified variances when $F = m = 1$.

If $J \leq J_{\max}$

$$\frac{1}{\text{edf}} = \frac{1}{s_z^2(0, 1, \alpha, d) M} \text{BasicSum}(J, M, S, 1, \alpha, d)$$

Else if $J > J_{\max}$ and $r \geq d+1$, take a_0, a_1 from Table 1; then

$$\frac{1}{\text{edf}} = \frac{1}{r} \left(a_0 - \frac{a_1}{r} \right)$$

Else let $m' = \frac{J_{\max}}{r}$ (not necessarily an integer); then

$$\frac{1}{\text{edf}} = \frac{1}{s_z^2(0, 1, \alpha, d) J_{\max}} \text{BasicSum}(J_{\max}, J_{\max}, m', 1, \alpha, d)$$

4.6.2 Case 2. Unmodified variances, WHFM to RRFM: $F = m$, $\alpha \leq 0$

If $J \leq J_{\max}$

If $m(d+1) \leq J_{\max}$ then let $m' = m$ else let $m' = \infty$. Then

$$\frac{1}{\text{edf}} = \frac{1}{s_z^2(0, m', \alpha, d) M} \text{BasicSum}(J, M, S, m', \alpha, d)$$

Else if $J > J_{\max}$ and $r \geq d+1$, take a_0, a_1 from Table 2; then

$$\frac{1}{\text{edf}} = \frac{1}{r} \left(a_0 - \frac{a_1}{r} \right)$$

Else let $m' = \frac{J_{\max}}{r}$ (not necessarily an integer); then

$$\frac{1}{\text{edf}} = \frac{1}{s_z^2(0, \infty, \alpha, d) J_{\max}} \text{BasicSum}(J_{\max}, J_{\max}, m', \infty, \alpha, d)$$

4.6.3 Case 3. Unmodified variances, FLPM: $F = m$, $\alpha = 1$

If $J \leq J_{\max}$

$$\frac{1}{\text{edf}} = \frac{1}{s_z^2(0, m, 1, d) M} \text{BasicSum}(J, M, S, m, 1, d)$$

Remark: For this case, m must be less than about $1e6$ to avoid roundoff error.

Else if $J > J_{\max}$ and $r \geq d+1$, take a_0, a_1 from Table 2 ($\alpha = 1$), b_0, b_1 from Table 3; then

$$\frac{1}{\text{edf}} = \frac{1}{(b_0 + b_1 \ln m)^2 r} \left(a_0 - \frac{a_1}{r} \right)$$

Else let $m' = \frac{J_{\max}}{r}$ (not necessarily an integer); then

$$\frac{1}{\text{edf}} = \frac{1}{(b_0 + b_1 \ln m)^2 J_{\max}} \text{BasicSum}(J_{\max}, J_{\max}, m', m', 1, d)$$

4.6.4 Case 4. Unmodified variances, WHPM: $F = m$, $\alpha = 2$

This calculation is exact, and can be expressed in closed form. In these formulas, $\binom{n}{k}$ denotes the

binomial coefficient $\frac{n!}{k!(n-k)!}$.

Let $K = \lceil r \rceil$.

If $K \leq d$

$$\frac{1}{\text{edf}} = \frac{1}{M} \left[1 + \frac{2}{\binom{2d}{d}^2} \sum_{k=1}^{K-1} \left(1 - \frac{k}{r} \right) \binom{2d}{d-k}^2 \right]$$

Else

$$\frac{1}{\text{edf}} = \frac{1}{M} \left(a_0 - \frac{a_1}{r} \right)$$

where

$$a_0 = \frac{\binom{4d}{2d}}{\binom{2d}{d}^2}, \quad a_1 = \frac{d}{2},$$

also given in Table 2 ($\alpha = 2$).

4.7 Tables

Table 1. Coefficients for modified variances

α	$d = 1$		$d = 2$		$d = 3$	
	a_0	a_1	a_0	a_1	a_0	a_1
2	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{7}{9}$	$\frac{1}{2}$	$\frac{22}{25}$	$\frac{2}{3}$
1	0.840	0.345	0.997	0.616	1.141	0.843
0	1.079	0.368	1.033	0.607	1.184	0.848
-1			1.048	0.534	1.180	0.816
-2			1.302	0.535	1.175	0.777
-3					1.194	0.703
-4					1.489	0.702

Table 2. Coefficients for unmodified variances

α	$d = 1$		$d = 2$		$d = 3$	
	a_0	a_1	a_0	a_1	a_0	a_1
2	$\frac{3}{2}$	$\frac{1}{2}$	$\frac{35}{18}$	1	$\frac{231}{100}$	$\frac{3}{2}$
1	78.6	25.2	790.	410.	9950.	6520.
0	$\frac{2}{3}$	$\frac{1}{6}$	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{7}{9}$	$\frac{1}{2}$
-1			0.852	0.375	0.997	0.617
-2			1.079	0.368	1.033	0.607
-3					1.053	0.553
-4					1.302	0.535

Table 3. Coefficients for logarithmic denominator, unmodified variances, FLFM ($\alpha = 1$)

$d = 1$		$d = 2$		$d = 3$	
b_0	b_1	b_0	b_1	b_0	b_1
6	4	15.23	12	47.8	40

5 Examples

6 Conclusions

References

- [1] R. Tausworthe, 1972, Convergence of oscillator spectral estimators for counted-frequency measurements, IEEE Transactions on Communications, COM-20, 214–217. [**2-tau stride**]
- [2] P. Lesage and C. Audoin, 1973, Characterization of frequency stability: uncertainty due to the finite number of measurements, IEEE Transactions on Instrumentation and Measurement, IM-22, 157–161.

- [3] P. Lesage and C. Audoin, 1975, Comments on “Characterization of frequency stability: uncertainty due to the number of measurements,” *IEEE Transactions on Instrumentation and Measurement*, IM-24, 86
- [4] P. Lesage and C. Audoin, 1976, Correction to “Characterization of frequency stability: uncertainty due to the number of measurements,” *IEEE Transactions on Instrumentation and Measurement*, IM-25, 270.
- [5] P. Lesage and C. Audoin, 1977, Estimation of the two-sample variance with limited number of data, 31st Annual Symposium on Frequency Control, 311–318.
- [6] K. Yoshimura, 1978, Characterization of frequency stability: uncertainty due to the autocorrelation of the frequency fluctuations, *IEEE Transactions on Instrumentation and Measurement*, IM-27, 1978, 1–7.
- [7] D. Howe, D. Allan, and J. Barnes, 1981, Properties of signal sources and measurement methods, 35th Annual Symposium on Frequency Control, 1-47.
- [8] K. Yoshimura, 1989, Degrees of freedom of the estimate of the two-sample variance in the continuous sampling method, *IEEE Transactions on Instrumentation and Measurement*, 38, 1044–1049.
- [9] E. Masry, 1991, Flicker noise and the estimation of Allan variance, *IEEE Transactions on Information Theory*, 1173–1177. [**2-tau stride**]
- [10] W. Prestwich and T. Kennett, 1991, The statistical properties of Allan variance, *Canadian Journal of Physics*, 69, 1405–1415.
- [11] C. Greenhall, 1991, Recipes for degrees of freedom of frequency stability estimators, *IEEE Transactions on Instrumentation and Measurement*, 40, 994–999.
- [12] T. Walter, 1994, Characterizing frequency stability: a continuous power-law model with discrete sampling, *IEEE Transactions on Instrumentation and Measurement*, 43, 69–79.
- [13] C. Greenhall, 1997, The third-difference approach to modified Allan variance, *IEEE Transactions on Instrumentation and Measurement*, 46, 696–703.
- [14] F. Vernotte, G. Zalamansky, M. McHugh, and E. Lantz, 1996, Estimation of the upper limit on the level of an undetected noise: application to the study of millisecond pulsars, 1996 *IEEE International Frequency Control Symposium*, 875–879.
- [15] C. Greenhall, 1983, A structure function representation theorem with applications to frequency stability estimation, *IEEE Transactions on Instrumentation and Measurement*, IM-32, 364–370.
- [16] D. Howe, R. Beard, C. Greenhall, F. Vernotte, and W. Riley, 2001, A total estimator of the Hadamard function used for GPS operations, *Proceedings of the 32nd Annual Precise Time and Time Interval (PTTI) Meeting*, Nov. 2000, Reston, VA, 255–267.
- [17] D. Howe, 2000, The total deviation approach to long-term characterization of frequency stability, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 47, 1102–1110.
- [18] D. Howe and T. Pepler, 2003, Very long-term frequency stability: estimation using a special-purpose statistic, *Proceedings of the 2003 IEEE International Frequency Control Symposium*, 233–238.

A Explanation of approximations

As is, (6) is unfit for numerical computation. We find empirically that $s_z^2(t)$ tends rapidly to zero as t increases beyond d . For the accuracy needed here (a few percent), there is no point in allowing j/S to be more than $d + 1$. Indeed, for sufficiently large t the calculation of $s_z^2(t)$ blows up from roundoff error, even in double precision, because linear combinations of large s_w values are taken to get small s_z values. At the very least, one should truncate the sum at $j = (d + 1)S$, as in the simplified version of the algorithm.

The full version of the algorithm uses the following general strategy. If $J \leq J_{\max}$ we do the summation (6). When $J > J_{\max}$ there are two cases. First, if $M \geq (d + 1)S$ then $S = m \geq J_{\max}/(d + 1) \gg 1$. We truncate the sum at $(d + 1)S$ and approximate it by an integral; this gives

$$\begin{aligned} \frac{1}{\text{edf } V_d} &\approx \frac{1}{r} \int_0^{d+1} \left(1 - \frac{t}{r}\right) s_z^2(t) dt \\ &= \frac{1}{r} \left(a_0 - \frac{a_1}{r}\right), \end{aligned}$$

where

$$r = \frac{M}{S}, \quad a_0 = 2 \int_0^{d+1} s_z^2(t) dt, \quad a_1 = 2 \int_0^{d+1} s_z^2(t) t dt.$$

These coefficients can be evaluated in advance. Second, if $M < (d + 1)S$ then we do another summation in which J is reduced from M to J_{\max} and S is reduced proportionately from m .

The extra term for $j = J$ in BasicSum makes the sum a trapezoidal approximation to the integral, whether or not the sum is truncated.

This method works straightforwardly for Case 1; indeed, in this setting the modified variances are simpler than the unmodified ones. In Case 2, when m is large we compute $s_z(t)$ using the limiting form $s_x(t, \infty)$, which is actually $-s_w''(t)$. This means that we are treating $x(t)$ as $w'(t)$, the process $w(t)$ being differentiable in the mean-square sense.

The most troublesome case is the short-stride estimators of the unmodified variances for FLPM. As $S = m \rightarrow \infty$, $s_z(t)$ approaches a function with logarithmic singularities. The factor $b_0 + b_1 \ln m$ is an asymptotic form of $s_z(0)$. It would be possible (though inconvenient) to add another large- m subcase as in Case 2, but one does not expect FLPM to be the dominant noise type when m is large.

Case 4 is constructed by knowing that the phase samples are accurately uncorrelated when $w(t)$ is a Wiener process. The simplified computation (13) is correct but wasteful because $s_z(t)$ is a linear combination of hat-shaped peaks of width $2/m$.