

The Total Deviation Approach to Long-Term Characterization of Frequency Stability

David A. Howe
(Invited Paper)

Abstract—Total variance represents the first advancement in the estimation of an oscillating signal's long-term frequency stability since the conventional Allan variance estimator. It is called "Totvar" (pronounced tōt 'vār) for short. Its commonly reported square-root is the total deviation function called "Totdev" (tōt 'dēv). Totdev efficiently extracts broadband FM oscillator noise levels and types commonly encountered at long-term τ -averaging times, even though these are difficult to measure. The notable advantages of Totdev include: significantly improved confidence for estimating long-term frequency stability, lower sensitivity to the removal of linear frequency drift, and exact decomposition of the sample variance of frequency residuals.

A review of some methods of improving the estimation of long-term stability is presented. Concepts of data extrapolation and smoothing are developed and linked to the definition of Totdev. The approach to interpreting a σ vs. τ stability plot of Totdev is given and is identical to that using the Allan deviation. Formulas are given for computing confidence interval at long term and for computing an easily removed, slight negative bias. The last section shows how Totdev is the first frequency-stability statistic that is unified to the more readily understood, but often inappropriate, standard deviation by an exact decomposition.

I. INTRODUCTION

THIS writing assumes some familiarity with the Allan variance estimator of frequency stability called "Avar," or more conveniently its usually reported square root, "Adev" [1]–[5]. For long averaging times (greater than 10% of a whole data run), Totdev is a recommended substitute for Adev because it is an improved estimator of frequency stability at long-term τ -values while essentially computing the best Allan deviation estimator (the maximum-overlap estimator) at short- and mid-term τ -values [6]. For example, for computations at longest-term averaging time, or 50% of the whole data run, the equivalent degrees of freedom (or edf, used to quantify an estimator's confidence) shows an increase from 1 to 1.5, 2.1, and 3 with commonly encountered FM noises of random walk (RWFm), flicker (FLFM), and white (WHFM) noise [7]–[9]. Totdev has an easily removed slight negative bias associated with RWFm and FLFM noises.

This paper is a quick guide to Totdev's statistical ap-

proach. In Section II, the typical measurement methodology using Adev is described. Looking ahead, Totdev is basically Adev computed on a periodically extended version of a data run. So why is Totdev better than Adev, since its computation simply reuses the same data? I show in Sections III, IV, and V that, in reality, no periodic assumption whatsoever is made about the data in the application of Totdev, but that its apparent periodic methodology "smooths" statistical results by averaging additional numbers of normally unused, yet completely consistent, frequency-stability estimates, called Totdev's "surrogate" estimates. This reduces the dispersion in results.

Totdev's important beneficial properties also have simple forms given in Sections VI and VII. For example, the normally ponderous task of computing confidence intervals is remarkably simple when using Totdev. Its bias relative to the Allan deviation is modest and has a simple formula. Based on these considerations, Totdev can be easily and profitably used to extract additional information about long-term stability, information that is often lost by the use of the normal Allan deviation.

In Section VIII, I describe an important connection between the standard deviation of frequency fluctuations and the usual frequency stability plot using Total deviation, what is called a " σ vs. τ " plot. We can treat Totdev as a full-octave spectrum analyzer whose filter response is given in Section V. We extract exactly $\sqrt{2}$ times the standard deviation by summing values of Totdev in commonly reported dyadic τ -averaging times, meaning increments of 1, 2, 4, \dots , 2^J where positive integer J depends on the length of the data run, along with one additional term that represents variations at all dyadic averaging times greater than the usual last one at 2^J . This feature yields a means of reporting the extent of frequency variations in a data run left unaccounted for in a frequency stability plot using Totdev. This property distinguishes Totdev from Adev, and I explain why it is useful.

II. MEASURING FREQUENCY STABILITY

If the time or the time fluctuations between two oscillators can be measured directly, an advantage is obtained over just measuring frequency fluctuations. The reason is that we can readily see time behavior from actual measurements, and frequency can easily be inferred from time. To avoid measurement-system dead time and simultaneously

Manuscript received April 28, 1999; accepted March 26, 2000.

The author is with the National Institute of Standards and Technology, 325 Broadway, Boulder, CO 80303 (e-mail: dhowe@nist.gov).

measure the underlying frequency stability of the best oscillators often nearly at the same frequency, we use the dual mixer time difference (DMTD) scheme [3]–[5]. Measurement samples of time fluctuations occur at a rate f_s having an interval $\tau_0 = \frac{1}{f_s}$. Given a sequence of time deviates $\{x_n : n = 1, \dots, N_x\}$ with a sampling period between adjacent observations given by τ_0 , we define the $m\tau_0$ -average fractional frequency deviate as:

$$\bar{y}_n(m) \equiv \frac{1}{m} \sum_{j=0}^{m-1} y_{n-j} = \frac{1}{m\tau_0} (x_n - x_{n-m}), \quad (1)$$

since, in practice, $\bar{y}_n(m)$ is computed from x_n and x_{n-m} . Note that $y_n = \frac{1}{\tau_0}(x_n - x_{n-1})$, and to be clear, $\{y_n\} = \{\bar{y}_n(1)\}$, that is, y_n -values are *also* averages. We regard $\{\bar{y}_n(m) : n = m, \dots, N_y\}$ as a finite realization of a stochastic process $\{\bar{Y}_n(m) : n = 0, \pm 1, \pm 2, \dots\}$. If we make $E\{\bar{y}_n(m)\} = 0$ by a proper translation, the standard variance is:

$$\sigma^2(m) \equiv E\{\bar{y}_n^2(m)\}, \quad (2)$$

where E throughout this paper means an expected average or infinite mean. Even in the most stable oscillators, $\{\bar{y}_n(1)\}$ can be a realization of a nonstationary frequency noise such as FLFM and RWFM. Because frequency fluctuations have a nonstationary average, the sample standard variance can be arbitrarily biased dependent on when and how long measurements are taken. Besides being unreliable, the standard variance does not offer a simple way to distinguish oscillator noise types; hence, a way to diagnose the source or cause of frequency instability. Allan [1] devised a convenient variance of oscillator frequency stability that always converges as the first difference of FLFM and RWFM time-series processes becomes stationary in a mean-squared sense and easily distinguishes a common range of noise types by straight-line slopes vs. $m\tau_0$ (on a log-log plot). Consequently, an IEEE subcommittee recommended that the characterization of frequency stability be defined by [2]:

$$\sigma_y^2(m) \equiv \frac{1}{2} E\{[\bar{y}_{n+m}(m) - \bar{y}_n(m)]^2\},$$

$$\text{and } E\{[\bar{y}_{n+m}(m) - \bar{y}_n(m)]\} = \text{Dr}(m), \quad (3)$$

where $\text{Dr}(m)$ is a linear trend or linear frequency drift that is estimated and removed, or, with justification, assumed to be 0. Its estimation and removal is problematic to estimating the underlying noise characteristics in $\sigma_y^2(m)$ at long-term or large m -values, an issue discussed later. If the first difference $\{\bar{y}_n(1) - \bar{y}_{n-1}(1)\}$ is stationary, the stochastic process is such that the expectations above depend on the averaging time index m , but are not dependent on the time index n ; that is, the expectation value does not depend on when we measure. Note that each point estimate of the Allan variance (the first term in curly braces) is computed at m but requires a $2m$ interval.

A hat “ $\hat{\cdot}$ ” denotes a sample estimate of a function. The usual sample Allan variance $\hat{\sigma}_y^2(\tau, T)$ involves averaging time $\tau = m\tau_0$ and sample data run T . As previously mentioned, it is called Avar and is defined by [3], [5]:

$$\hat{\sigma}_y^2(\tau, T) = \text{Avar}(m, \tau_0, N_{x,y}) = \frac{1}{2(N_y - (2m - 1))} \sum_{n=m}^{N_y - m} (\bar{y}_{n+m}(m) - \bar{y}_n(m))^2 \quad (4)$$

and by (1) is equal to

$$\frac{1}{2(m\tau_0)^2 (N_x - 2m)} \sum_{n=m+1}^{N_x - m} (x_{n+m} - 2x_n + x_{n-m})^2, \quad (5)$$

for $1 \leq m \leq \frac{N_x - 1}{2}$. Summand terms in (5) involve a second difference of $\{x_n\}$ known as a central difference when expressed symmetrically over a $2m\tau_0$ span as in (5). Each central difference consists of only a first, middle, and last x_n value, or x_n 's taken in triplet, then subsequently squared and averaged. Having removed drift and other deterministic error sources, an oscillator's random FM noise will have triplets that on average fall on a straight line with a Gaussian distribution, as pointed out by Allan because these central differences now comprise a stationary process of independent components. Their variance thus has a χ^2 (chi-square) distribution and, consequently, becomes skewed low if faced with few samples. This is simply because any one triplet is more likely to fall nearly on a straight line than not. Thus, a one-sample central-difference variance is often uncharacteristically low.

III. WAYS TO INCREASE THE NUMBER OF STATISTICAL SAMPLES

One way to get around the small-sample problem associated with Avar at longest-term τ -values is to call on a useful relationship between the Allan variance and the classical standard variance $\sigma^2(m)$ as:

$$\sigma_y^2(m) = 2|\sigma^2(m) - \sigma^2(2m)|. \quad (6)$$

If $2m$ is the length of the whole data run and if we know nothing more, the standard variance (the so-called biased formulation) is given by $\hat{\sigma}^2(2m) = \bar{y}_n^2(2m)$, which can be computed with good confidence having $2m$ sample values. This global variance can be regarded as the “true” variance, or $\sigma^2(2m)$ in (6). Up to m samples or half the data run, we have $\hat{\sigma}^2(m) = \{\bar{y}_n^2(m)\}$ (and, of course, it is also biased) and (6) yields an estimate of the Allan variance, or $\hat{\sigma}_y^2(m)$ by $\hat{\sigma}_y^2(m) = 2|\hat{\sigma}^2(m) - \sigma^2(2m)|$. A difficulty lies in the fact that, using both $\hat{\sigma}^2(m)$ and the global variance $\bar{y}_n^2(2m)$, usually causes $\hat{\sigma}_y^2(m)$ to be seriously biased low. Because by a time-reversal invariance, $\{\bar{y}_n^2(m)\}$ can be either the first or last half of the whole data run, by extension any $\{\bar{y}_n^2(m)\}$ segment is allowable and the largest possible $\bar{y}_n^2(m)$ is often chosen for $\hat{\sigma}^2(m)$. This gives an

upper-bound value using (6), which, if plausible, is accepted as an estimate of the Allan variance for $\tau = T/2$ or half the full length of the data run and that may offset the negative bias of (6). Use of this ad hoc procedure can be justified when $\sigma_y^2(m)$ appears uncharacteristically low when solely using Avar (4), (5). But this practice is subjective and ultimately flawed because experimenter's expectations govern "plausibility;" however, it hints at a possible alternative to Avar.

Yet another approach is to correct for the bias in a computation of the classical standard variance, because this variance has, by comparison to Avar, many more sample values, hence good confidence at long term. The problem is that its bias is a function of which noise type is present, and Avar, with its relatively poor confidence, must be used to determine this noise type. So the approach ends up actually netting poorer overall confidence compared to using the Avar equations alone.

Another technique is to exploit the time fluctuations in and around Avar's aforementioned central-difference triplets on the belief that we are at liberty to choose a range of neighboring quantities that can serve as alternative or surrogate values and average them to obtain a better estimator $\hat{\sigma}_y^2(\tau, T)$. Recall that Avar measures only a symmetry, or a lack thereof, in equispaced triplet values of $x_{n\pm m}$. The "surrogates" originate in the fact that we can apply a consistency hypothesis to measures of other symmetry in the same span. "Consistency" means that averages of certain other individual estimates, in this case neighboring, similar measurements of symmetry that I will describe next, can equally serve as an estimate. This is the principle behind statistical "smoothing."

For example, an approach that introduces the concept of consistency requires a slightly altered view of Avar, that is, Avar as a measured variance of an idealized, simple linear interpolator. Here we regard the central difference operation mentioned in Section II as a linear predictor, whose prediction uncertainty becomes Adev if we extrapolate the data to expected values outside the actual measurement, namely into the future [10].

To illustrate the concept, note that the form of Avar in (4) originates because values of $\{y_n(m)\}$ are actually measured asymmetrically with respect to $\{x_n\}$ values, that is, post facto, in which $\bar{y}_n(m) = \frac{1}{m\tau_0}(x_n - x_{n-m})$. For symmetry, however, we can substitute:

$$\bar{y}_n^o(m) = \frac{1}{m\tau_0} (x_{n+\frac{m}{2}} - x_{n-\frac{m}{2}}). \quad (7)$$

By a linear interpolation, $\bar{y}_n^o(m) =$ average of:

$$\frac{1}{m\tau_0} (x_{n-m} - x_n) \text{ and } \frac{1}{m\tau_0} (x_n - x_{n+m}), \quad (8)$$

whose result is $\frac{1}{2m\tau_0}(x_{n+m} - x_{n-m})$, or equivalently $\bar{y}_n^o(2m)$. In other words, viewing Avar's central-difference operation as a linear interpolator, the median value x_n drops out in the average $\bar{y}_n^o(m)$, and we can use $\bar{y}_n^o(m)$ defined symmetrically in (7) in place of $\bar{y}_n^o(2m)$ and vice

versa. That is, the usual two asymmetrically computed values of $\bar{y}_n(m)$ in (4) can be substituted with $\bar{y}_n^o(2m) = \bar{y}_n^o(m)$ to also compute (4). The Allan variance $\sigma_y^2(m)$ has an implied use of one linear-interpolation value in a $2m$ interval. We can extend this idea to computations at $\bar{y}_n(2m)$, $\bar{y}_n(4m)$, $\bar{y}_n(8m)$, and so forth.

Consider for example the usual Avar estimator computed at $m = 2$ for a segment of data consisting of a series of five equispaced time-error measurements $\{x_1, x_2, x_3, x_4, x_5\}$. We can compute three (overlapping) average frequency errors as $\{\bar{y}_1(2), \bar{y}_2(2), \bar{y}_3(2)\} = \frac{1}{2\tau_0}\{x_3 - x_1, x_4 - x_2, x_5 - x_3\}$, but $\bar{y}_2(2)$ is not used in (4), (5). Additionally by a linear interpolation, $\bar{y}_2^o(2) = \frac{1}{4\tau_0}[(x_5 - x_3) + (x_3 - x_1)]$, which is also not used, yet it is consistent with the other values. At this point we have four estimates of $\bar{y}(2)$, and Totdev uses all of them instead of only the two used in Adev. To accomplish this, the total deviation approach uses linear extrapolation of the original data run to form the two-sided, mirror-reflection data extension in the precise definition of Totdev given in Section IV. This extension of the whole data run resembles a periodic extension (one which originates from a Fourier-based periodic assumption) and so is confused with statistical treatments that assume Fourier-frequency components as the basis of the data. But actually, the total approach is based on no periodic assumption whatsoever.

The idea of extending an oscillator's measured frequency-difference data in a periodically repeating, or circularized, fashion in the time domain originated because Avar is likely to "collapse" to 0 at long averaging times. Originally total variance started with the idea to simply time-shift the original data by 1/4 of the whole data run, compute a so-called "quadrature" Avar, and average this with the standard Avar [11]. This time shift requires some kind of extension and straight periodicity worked fine. This is justified recalling that in (6), estimate $\hat{\sigma}^2(2m)$ must be regarded as the "true" variance $\sigma^2(2m)$ if the data run is length $2m$. This implies that the data recur infinitely if the ergodic property is assumed to extend outside the data run [12]. Averaging the usual Avar with the quadrature Avar cleverly avoided the possibility of a long-term collapse-to-zero of the usual σ vs. τ frequency stability plot. The improvement in confidence was only slight, but the technique achieved dramatically reduced sensitivity to the removal of an overall drift estimate [13]. Because a quadrature Avar is a consistent estimator of original Avar, it follows that recomputations at every τ_0 increment also are consistent. Long-term estimation of the Allan variance was significantly improved by results using a straight circularization technique on original data $\{x_n\}$, but this technique could not work in the presence of RWFM and/or significant drift because it introduced gross bias [11], [14]. On discovering that the application of a two-sided mirror-reflection extension instead of a straight periodic extension handled RWFM, bias was significantly reduced while still preserving improved confidence of the final estimate, and now a total approach was applicable across Avar's full range of power-law noise types [7]–[9], [13].

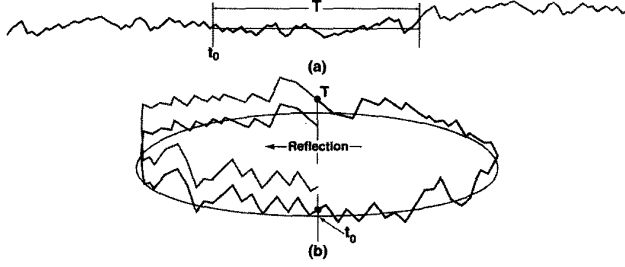


Fig. 1. Circular extension of the original $x(t)$ data set of duration T for computation of Totvar. (a) Extension of a phase record by “reflection” at both ends; (b) circular representation of extended phase record.

IV. DEFINITION OF TOTDEV

Total variance, or Totvar, is defined as:

$$\text{Totvar}(\tau, T) = \text{Totvar}(m, \tau_0, N_x) = \frac{1}{2(m\tau_0)^2(N_x - 2)} \sum_{n=2}^{N_x-1} \left(x_{n-m}^{\#} - 2x_n^{\#} + x_{n+m}^{\#} \right)^2, \quad (9)$$

for $1 \leq m \leq N_x - 1$ where an extended virtual sequence $\{x_n^{\#}\}$ is derived as follows: for $n = 1$ to N_x let $x_n^{\#} = x_n$; for $j = 1$ to $N_x - 2$ let

$$x_{1-j}^{\#} = 2x_1 - x_{1+j}, \quad x_{N_x+j}^{\#} = 2x_{N_x} - x_{N_x-j}. \quad (10)$$

Totvar also can be defined in terms of extended normalized frequency averages by:

$$\text{Totvar}(\tau, T) = \text{Totvar}(m, \tau_0, N_y + 1) = \frac{1}{2(N_y - 1)} \sum_{n=2}^{N_y} \left[\bar{y}_n^{\#}(m) - \bar{y}_{n-m}^{\#}(m) \right]^2, \quad (11)$$

where $\bar{y}_n^{\#}(m) = (x_{n+m}^{\#} - x_n^{\#}) / (m\tau_0)$. Construction of the extended virtual sequence given by (10) is illustrated in Fig. 1 and is called extension by reflection. Totdev values are the square root of Totvar values. Totdev is the current IEEE-recommended statistical test of oscillator frequency stability at long-term averaging times, namely, those τ -values beyond 10% of the length of the whole data run [6].

Rather than extending the original vector $\{x_n\}$ and applying the straight second difference, we can alternatively resample within the original vector. Applied to Totvar, this exercise only points out that the procedure used in the sampling function is intricate and not very intuitive because sampling on $\{x_n\}$ is no longer in terms of equispaced triplets spaced by 2τ [8]. It does, however, show that a large number of surrogate values emerge from Totvar’s data extension. Averaging a larger number of estimators in this fashion is an example of “data smoothing,” which is the key to how Totvar reduces the variability in its result.

V. CONCEPT OF SMOOTHING

All of the aforementioned ideas, including the method of computing Totdev, are ways to “smooth” a statistical result. Smoothing means that we somehow incorporate consistent, neighboring results in a weighted average to reduce the variability of a particular result. For example, other consistent estimators for Avar can be derived from its properties with measurement dead time. We can readily compute two estimates: one in which we intentionally introduce dead time (and positive bias), and another having equal-duration “overlap” time (and negative bias). Then an average of the two estimates is only modestly biased to first approximation. It would be a reasonable and consistent estimate that is not ordinarily considered. This concept links to Totdev in the following way. Using a symmetry argument and assuming that drift has been removed, let us define a useful expectation that is not only independent of values of the time index n , but also nearly independent of averaging time index m . Redefine $\bar{y}_n(m)$ as centered at n by:

$$\bar{y}_n(m) \equiv \frac{1}{m} \left(\sum_{j=0}^{\frac{m}{2}-1} y_{n-j} + \sum_{j=1}^{\frac{m}{2}} y_{n+j} \right), \quad (12)$$

or in terms of $\{x_n\}$ values, $\bar{y}_n(m) = \frac{1}{m\tau_0}(x_{n+\frac{m}{2}} - x_{n-\frac{m}{2}})$, for m even. Random FM noise processes also center before and after n such that,

$$E \{ \bar{y}_{n+\frac{m}{2}}(m) - \bar{y}_n(m) \} = E \{ \bar{y}_{n-\frac{m}{2}}(m) - \bar{y}_n(m) \} = E \{ \bar{y}_{n+\frac{m}{2}}(m) - \bar{y}_{n-\frac{m}{2}}(m) \} = 0, \quad (13)$$

and we can derive the symmetric form of the Allan variance as:

$$\sigma_y^2(m) \equiv \frac{1}{2} E \left\{ \left[\bar{y}_{n+\frac{m}{2}}(m) - \bar{y}_{n-\frac{m}{2}}(m) \right]^2 \right\}. \quad (14)$$

Substitute s for m in (12), and define (14) in terms of $\bar{y}_n(s)$ to obtain:

$$\sigma_{y^o}^2(m, s) \equiv \frac{1}{2} E \left\{ \left[\bar{y}_{n+\frac{s}{2}}(s) - \bar{y}_{n-\frac{s}{2}}(s) \right]^2 \right\}. \quad (15)$$

The separation between any two \bar{y} samples in (15) is still m as in (14), but now we are in a position to positively and negatively vary or symmetrically “modulate” the averaging-time s of each \bar{y} in the neighborhood of its usual value m by a small range, say, $\pm\delta$. This has the effect of smoothing $\sigma_y^2(m)$ by using other nearby consistent values in addition to the usual values of $\bar{y}_{n+\frac{s}{2}}(m) - \bar{y}_{n-\frac{s}{2}}(m)$ as described in Section IV. We increment n , repeat the process, square differences and average to obtain a smoothed value of $\hat{\sigma}_y^2(m)$ denoted by [15]:

$$\overline{\sigma_{y^o}^2(m, s)} = \text{smoothed version of } \sigma_y^2(m). \quad (16)$$

This smoothing operation picks up additional estimates of $\sigma_y^2(m)$ as m increases, improving the usual max-overlap Allan variance estimate, especially if that estimate is uncharacteristically high or low. The (15), and hence $\sigma_{y^o}^2(m, s)$, differ from the Allan variance because each term constituting a $2m$ interval yields its result dependent on m and s . This is because the difference pair of average frequencies $\bar{y}_{n+\frac{m}{2}}(s) - \bar{y}_{n-\frac{m}{2}}(s)$ may be separated when $s < m$ or overlapped when $s > m$. They are conjoined, or adjacent, only when $\delta = 0$, making $s = m$, which is precisely the unsmoothed Allan variance case. With WHFM, $\sigma_{y^o}^2(m, s)$ is unbiased with respect to $\sigma_y^2(m)$ and is biased negatively with FLFM and RWFM. The bias depends on the depth of modulation δ relative to m but is minimized if $\delta = \frac{m}{2}$, or a full range given by no more than m itself. This restriction also serves to control $\sigma_{y^o}^2(m, s)$ in a reasonable manner, but there is a problem with smoothing at the full data run of duration T having all N_x points. The problem is that our ability to smooth becomes more and more restrictive as $m \rightarrow N_x$ (or equivalently, as averaging time $\tau \rightarrow T$, the whole data run) because the extent with which we can modulate m and increment n is bounded by the beginning and end points of a data run of fixed length N_x . To smooth to the end of the data set requires an extension of the data set beyond its normal run. It is here we interpret Avar as the variance of a linear interpolator, as developed in the previous section. In principle, "Total" estimators are based upon the hypothesis that, for segment $\{x(t) : t_0 \leq t \leq T\}$, extensions for $t < t_0$ and $t > T$ can be formed by extrapolating the data, and in the Avar case, tacking on reversed versions of $\{x(t)\}$ at the beginning and end of this part of the function.

As mentioned previously, the original total estimator, and one which did not work well, derived from experiments in which the usual max-overlap Avar estimator was applied to simple periodic extensions of the original data [11]. Totvar's data extension by reflection, or tacking on a reversed version to both ends of the data segment, solved the problem of gross bias with RWFM, even though it was not fully understood why. Part of the confusion is because it was believed that Totvar somehow used the assumption of periodicity in the data coupled with the fact that it is actually a "hybrid" statistic in the sense that it combines the usual sample Allan variance in short- and medium-term with the confidence improvements at long term. More recently, the current understanding of smoothing and extrapolating has led to a new nonhybrid statistic called the modified total variance designed to extract the full range of both PM (phase modulation) and FM noises [16] and is a significant improvement on the modified Allan variance. The total approach is being tried on other time statistics such as classes of structure functions and the Hadamard variance [17].

To illustrate (9) and (10) as a hybrid statistic, consider the case of $m = 1$. The extension of $\{x_n^\# \}$ needs only to be τ_0 longer than $\{x_n\}$ at both ends to compute (9). Thus (9) is essentially the standard Allan estimator (5). As m increases, the extension needs to be longer until at

TABLE I
COEFFICIENTS FOR COMPUTING (17) AND (18), NORMALIZED BIAS
AND EDF OF TOTVAR.

Noise	a	b	c
WHFM	0	1.500	0
FLFM	0.481	1.168	0.222
RWFM	0.750	0.927	0.358

TABLE II
TABULATED QUANTITIES FOR $\tau = T/2$.

Noise	nbias ($T/2$)	edf($T/2$)
WHFM	0	3.000
FLFM	-0.240	2.097
RWFM	-0.375	1.514

$m = \frac{N_x - 1}{2}$, that is, the extensions at each end are length $\frac{N_x - 1}{2}$. Of course, there is no Allan estimator beyond half the length of the data run, namely the region $\tau > \frac{T}{2}$, so if the "hybrid" called Totvar in (9) is allowed to compute values for $m > \frac{N_x - 1}{2}$, it reverts to a region defined by the total variance but not the Allan variance. Computations of Totvar should not extend beyond $\tau = \frac{T}{2}$ to be consistent with the limit of the standard Allan estimator, but these higher-order terms will be considered in Section VIII.

VI. BIAS, EQUIVALENT DEGREES OF FREEDOM, AND FREQUENCY RESPONSE

For accurately estimating the Allan variance, an adjustment must be made to Totvar as defined by (9) and its extension $\{x_n^\#\}$ in (10) to remove a normalized bias (denoted as nbias) that depends on the ratio $\frac{\tau}{T}$ and on whether the noise type in long-term is FLFM or RWFM rather than WHFM. The most notable changes using Totvar compared to Avar involve formulae for nbias as well as increased edf used in Section VII to obtain lower uncertainty using Totvar. Totvar's nbias and edf can be summarized as [9]:

$$\text{nbias}(\tau) = \frac{E\{\text{Totvar}(\tau, T)\}}{\sigma_y^2(\tau)} - 1 = -a \frac{\tau}{T}, \quad (17)$$

$$\text{edf}(\tau) = \text{edf}[\text{Totvar}(\tau, T)] \approx b \frac{T}{\tau} - c, \quad (18)$$

where $0 < \tau \leq \frac{T}{2}$ and a , b , and c are given in Table I. The values of nbias and edf for the important longest-term case $\tau = T/2$ are tabulated in Table II. The edf formula (18) is a convenient, empirical or "fitted" approximation with an observed error below 1.2% of numerically computed exact values derived from Monte-Carlo simulation method; the tabulated values of edf($T/2$) in Table II are exact.

Both Totvar and Avar are invariant to certain manipulations of the vector $\{x_n\}$. The simplest example is that we can reverse and/or invert the sequence $\{x_n\}$ without

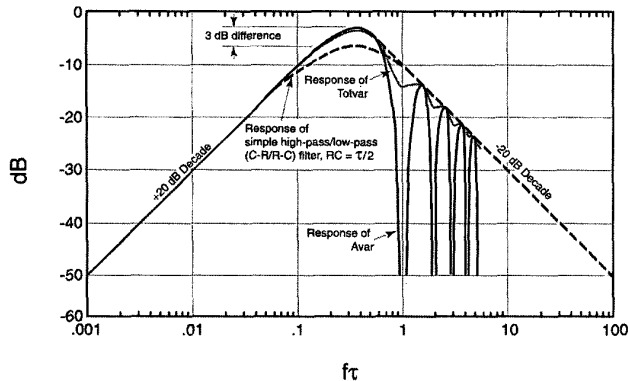


Fig. 2. A comparison of frequency responses of Avar (solid curve), Totvar (shaded curve), and a passband variance consisting of a simple cascade of a single-pole high-pass followed by a low-pass filter with identical break points at $RC = \tau/2$ (dashed curve [19]).

affecting either's result. Unlike Avar's simple sampling function, however, Totvar's data extension causes complicated manipulations of the sampling function itself between $\tau_0 \leq \tau \leq \frac{T}{2}$. These are derived from formulae in [8], but unraveling useful information from them is difficult. It is as informative as, but easier to look at, the frequency-response function associated with Totvar compared to Avar as in Fig. 2 for a comparison of the effect of their sampling functions. The dashed curve in Fig. 2 is a constant- Q , one-octave pass-band filter response considered to be ideal for extracting typical power-law noise levels [18]–[22]. Totvar implements a circular convolution of Avar's frequency response, thus significantly reducing the depth of periodic nulls.

VII. UNCERTAINTY OF TOTDEV

Returning to the topic of characterizing noise, the reason for using Totdev is to very efficiently extract commonly encountered integer power-law noise types and levels of an oscillator's spectral FM noise. This means greater certainty in the extraction of these parameters and others such as drift and quasi-sinusoidal modulation shown in Fig. 3. Assuming statistically independent noise or error sources, Totdev can be expressed as a linear combination of these noises and error sources. The most common measurement-system phase noises (Quantization PM, WHPM, and FLPM) are proportional to τ^{-1} , and the level depends on an upper-bandwidth or cutoff (Fourier) frequency f_h . Frequency noises (WHFM, FLFM, and RWFM) are proportional to $\tau^{-\frac{1}{2}}$, τ^0 , $\tau^{+\frac{1}{2}}$, respectively. Totdev can serve as an estimator of levels of Quasi-Sinusoidal FM (QSFM) and drift ($Dr \propto \tau^{+1}$), although these are usually best analyzed separately, estimated, and removed, so we are dealing only with residuals in the identification of levels and types of underlying power-law noise processes.

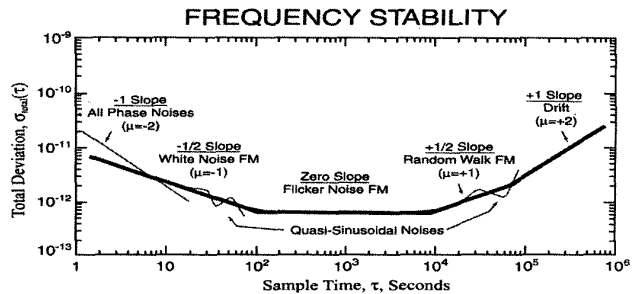


Fig. 3. Total deviation plot (or root Totvar) showing power-law noises as straight lines in addition to other error sources. The goal is to identify noise sources and accurately estimate their levels with this kind of frequency stability plot.

We can write a composite variance as:

$$\sigma_{composite}^2 = \sigma_{QPM}^2 + \sigma_{WH/FLPM}^2 + \sigma_{WHFM}^2 + \sigma_{FLFM}^2 + \sigma_{RWFM}^2 + \sigma_{QSFM}^2 + \sigma_{Drift}^2, \quad (19)$$

and Total deviation as a composite function is then given by:

$$Total\ deviation = \sqrt{\sigma_{composite}^2} =$$

$$f(\sigma_{QPM}, \sigma_{WH/FLPM}, \sigma_{WHFM}, \sigma_{FLFM}, \sigma_{RWFM}, \sigma_{QSFM}, \sigma_{Drift}) = \sum_{\mu=-2}^2 a_{\mu} \tau^{\mu/2}, \quad (20)$$

where the coefficients a_{μ} are usually obtained in a least-mean-squares sense from a Totdev σ vs. τ plot.

The formulation of the two-sample Allan ensemble average will contain $N_{\bar{y}(m)}$ average frequencies and only $\frac{N_{\bar{y}(m)}}{m} - 1$ independent intervals with which to estimate a noise coefficient. This represents the actual number of “degrees of freedom”. In general the 1σ uncertainty for $N_{\bar{y}(m)}$ sets is simply given by:

$$\%error, Allan\ deviation = \frac{100}{\sqrt{2(\frac{N_{\bar{y}(m)}}{m} - 1)}}. \quad (21)$$

This expression is an uncertainty for Adev that is adequate for a quick upper-bound approximation for a confidence interval or error bars above and below each value of a plot of Allan deviation σ vs. τ . We assume that the probability distribution is chi-square, and exact confidence intervals can be determined based on the equivalent degrees of freedom (edf) in overlapping statistical averages for a given noise type, rather than the actual number [3], [23].

Totdev values have edf's that are greater than those obtained using Adev, and significantly greater at long-term τ -values. Chi-square distribution functions are used for estimating confidence of the Allan variance. It turns out that the distribution functions are slightly narrower than chi-square using Totdev (which is another of its benefits) at

long averaging times. Thus a conservative 1σ upper-bound uncertainty on the estimation of a noise coefficient using Totdev is:

$$\%error, Total\ deviation = \frac{100}{\sqrt{2(edf)}}, \quad (22)$$

where edf values are conveniently derived from (18) and Table I based on the computation of the Total deviation as a ratio of τ to the length of the data run T , rather than the Allan deviation's $\frac{N_y(m)}{m}$ independent sets as mentioned earlier.

VIII. ANALYSIS OF VARIANCE

Consider a function of independent variables. In analysis of variance, we explain the total variability of the function in terms of each variable. In the discussion here, we address functionals that depend on a time interval Δt . At this point we can recall a conservation principle regarding the standard sample variance, which states that the mean of the interval variances plus the variance of the interval means equals the standard variance of the entire series. This is true for any process, stationary or not. An analysis of variance in terms of the mean of k interval variances and variance of the k interval means is derived in the Appendix.

The standard variance of finite series $\{X_{ij}\}$ in the Appendix is simply a number, partly due to the variance of k interval means and the remaining part due to the mean of k interval variances. Now consider intervals of duration Δt and a whole data run of length T . The longest possible set of equal-length intervals would be $\Delta t = T/2$; thus there are $k = 2$ consecutive interval means. We recognize that the variance of such two-interval means is the special-case two-sample variance equaling $\frac{1}{2}\hat{\sigma}_y^2(T/2)$, half the sample Allan variance at $\tau = T/2$. But half the sample Allan variance will differ from the standard variance by a remaining portion attributable to the sample variance within each of the two intervals by the conservation principle just stated. By double-sampling at $\Delta t = T/4$, we find the two-sample variance ($k = 2$) now must consist of two nonoverlapped variance estimates whose average, denoted as $\hat{\sigma}_{y,\text{nono}}^2(T/4)$, would be the remaining portion if that were as far as the data could be sampled. Repeating this process until there are no remaining interval variances left unaccounted for, we find that:

$$\frac{1}{2} (\hat{\sigma}_{y,\text{nono}}^2(\tau_0) + \hat{\sigma}_{y,\text{nono}}^2(2\tau_0) + \dots \quad (23)$$

$$+ \hat{\sigma}_{y,\text{nono}}^2(T/4) + \hat{\sigma}_{y,\text{nono}}^2(T/2)) = \hat{\sigma}_{std}^2(T),$$

where $T = m\tau_0$, $m = 2^j$ for $j = 0, 1, \dots, J-1$, and J is a positive integer corresponding to the maximum power-of-2 sample size of y measurements, that is, $N_y = 2^J$. The

nonoverlapping estimator of the Allan variance is:

$$\hat{\sigma}_{y,\text{nono}}^2(2^j) \equiv \frac{2^j}{2N_y} \sum_{k=1}^{\frac{N_y}{2^{j+1}}} [\bar{y}_{2k2^j}(2^j) - \bar{y}_{(2k-1)2^j}(2^j)]^2. \quad (24)$$

The composite in (23) is a common property of what is called a “multiresolution pyramid” [24]. The nonoverlapped $k = 2$ condition requires that the τ -intervals occur in power-of-2 increments. This nonoverlapping sample Allan variance would relate directly to the sample standard variance as in (23) but is inefficient as an estimator [21]. Unfortunately the sample Allan variance from its definition, for example at $\Delta t = T/4$, calls for three variance estimates, not two nonoverlapped, because its definition includes a $\tau = \Delta t$ overlap and the straight-forward relationship to the standard variance is lost rather quickly. In other words, even for a short series,

$$\frac{1}{2} (\hat{\sigma}_y^2(T/4) + \hat{\sigma}_y^2(T/2)) \neq \hat{\sigma}_{std}^2(T), \quad (25)$$

in contrast to (23). Because the definition of the Allan variance contains one τ -overlap, we can admit all possible overlaps to obtain an improved estimator in order to minimize its error bars. Known as the standard “max-overlap” Allan estimator [3] given as (5), it also departs from a tractable connection to the standard variance for the same reason as the original τ -overlap estimator.

Functionals that depend on a time interval Δt have such a strong connection to spectral functions that (23) is a “decomposition” of the sample standard variance and seems appropriate jargon, so it is commonly used. In this regard, decomposition of the standard variance is suited to frequency-domain analysis, and Totvar maintains a straight-forward relationship with the sample standard variance. It abides by the conservation principle if we consider an infinite extension by reflection. This means that the virtual sequence generated by (10) and shown in Fig. 1 recurs indefinitely [9]. Percival and Howe [25] were the first to point out that, for the case in which Totvar is computed in power-of-2 increments above $T/2$ as estimated from data-run T as in (9), a remaining portion, the sum of Totvar terms of all power-of-2 intervals $\tau > T/2$ for $\tau \rightarrow \infty$, adds to the usual multiresolution pyramid to precisely equal the standard sample variance. These leftover higher-order components are never actually reported but are an artifact of infinitely extending the original sequence. They can be regarded as the sum of 0-frequency aliases, a remaining “dc” term to make up $\hat{\sigma}_{std}^2(T)$. Greenhall et al. [9] coined the term Remvar($\frac{T}{2}$) to designate this portion. Totvar beyond T soon drops to nearly zero, so the remaining portion above T is generally very small. Nevertheless, Remvar accounts for this portion and was used in the proof of the decomposition of the standard variance. Summing all the familiar “power-of-2” τ -values in a Totvar plot leads to exactly twice the standard sample variance, much in the same way that integrating an estimate of a spectrum also yields the sample variance.

Knowledge that we can account for all variations in a data-run by its standard variance as “decomposed” in calculations of the sample Total variance is especially useful. For example, at a long-term τ -value of $T/2$, an equal remaining portion ($\text{Totvar}(\frac{T}{2}) = \text{Remvar}(\frac{T}{2})$) would indicate that we have summarized completely the variations at $T/2$.

ACKNOWLEDGMENTS

I am grateful to Chuck Greenhall and Don Percival, who have helped in analyzing many properties of Totvar. I also thank Uwe Radok and Jim Barnes for guidance and support.

REFERENCES

[1] D. W. Allan, “Statistics of atomic frequency standards,” in *Proc. IEEE*, vol. 54, pp. 221–230, Feb. 1966.
 [2] J. A. Barnes, A. R. Chi, L. S. Cutler, D. J. Healy, 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, pp. 105–120, 1971.
 [3] D. A. Howe, D. W. Allan, and J. A. Barnes, “Properties of signal sources and measurement methods,” in *Proc. 35th Ann. Symp. Freq. Contr.*, 1981, pp. 1–47.
 [4] S. R. Stein, “Frequency and time—Their measurement and characterization,” in *Precision Frequency Control*, vol. 2, E. A. Gerber and A. Ballato, Eds. New York: Academic, 1985, pp. 191–232 and 399–416.
 [5] D. B. Sullivan, D. W. Allan, D. A. Howe, and F. L. Walls, Eds. “Characterization of clocks and oscillators,” Natl. Inst. Stand. Technol., Boulder, CO, Technical Note 1337, 1990.
 [6] E. S. Ferre-Pikal, J. R. Vig, J. C. Camparo, L. S. Cutler, L. Maleki, W. J. Riley, S. R. Stein, C. Thomas, F. L. Walls, and J. D. White, *IEEE 1139-1999: Standard Definitions of Physical Quantities for Fundamental Frequency and Time Metrology—Random Instabilities*. New York: Institute of Electrical Engineers.
 [7] D. A. Howe, “Methods of improving the estimation of long-term frequency variance,” in *European Frequency and Time Forum*. Neuchâtel, Switzerland, 1997, pp. 91–99.
 [8] D. A. Howe and C. A. Greenhall, “Total variance: A progress report on a new frequency stability characterization,” in *Proc. 29th Ann. PTTI Systems and Applications Meeting*, Long Beach, CA, 1997, pp. 39–48.
 [9] C. A. Greenhall, D. A. Howe, and D. B. Percival, “Total variance, an estimator of long-term frequency stability,” *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. UFFC-46, pp. 1183–1191, Sep. 1999.
 [10] J. Levine, “Time and frequency metrology,” *Rev. Sci. Instrum.*, vol. 70, pp. 2567–2596, Jun. 1999.
 [11] D. A. Howe, “An extension of the Allan variance with increased confidence at long term,” *Proc. IEEE Int. Freq. Contr. Symp.*, 1995, pp. 321–329.
 [12] —, “Circular representation of infinitely extended sequences,” *Proc. IEEE Int. Freq. Contr. Symp.*, 1995, pp. 337–345.
 [13] D. A. Howe and K. J. Lainson, “Effect of Drift on TOTALEV,” *Proc. IEEE Int. Freq. Contr. Symp.*, 1996, pp. 883–889.
 [14] F. Vernotte, “Stabilité temporelle des oscillateurs: Nouvelles variances, leurs propriétés, leurs applications,” (A circularized Allan variance statistic is defined but not investigated), Ph.D. dissertation, Université de Franche-Comte, Besançon, France, 1991.
 [15] D. C. Fraser and J. E. Potter, “The optimum linear smoother as a combination of two optimum linear filters,” *IEEE Trans. Automat. Contr.*, vol. AC-14, pp. 387–390, Aug. 1969.

[16] D. A. Howe and F. Vernotte, “Generalization of the total variance approach to the modified Allan variance,” in *Proc. 31st Ann. PTTI Systems and Applications Meeting*, Dana Point, CA, Dec. 7–9, 1999.
 [17] F. Vernotte and D. A. Howe, “Generalization of the total variance approach to the different classes of structure functions,” in *Proc. 14th European Frequency and Time Forum*, Torino, Italy, unpublished.
 [18] J. Rutman, “Characterization of frequency stability: A transfer function approach and its application to measurements via filtering of phase noise,” *IEEE Trans. Instrum. Meas.*, vol. IM-23, pp. 40–48, Mar. 1974.
 [19] R. G. Wiley, “A direct time-domain measure of frequency stability: The modified Allan variance,” *IEEE Trans. Instrum. Meas.*, vol. IM-26, pp. 38–41, Mar. 1977.
 [20] D. W. Allan, M. A. Weiss, and J. L. Jespersen, “A frequency-domain view of time-domain characterizations of clocks and time and frequency distribution systems,” in *Proc. 45th Freq. Contr. Symp.*, 1991, pp. 667–678.
 [21] D. B. Percival, “Characterization of frequency stability: Frequency-domain estimation of stability measures,” *Proc. IEEE*, vol. 79, pp. 961–972, 1991.
 [22] D. A. Howe and D. B. Percival, “Wavelet variance, Allan variance, and leakage,” *IEEE Trans. Instrum. Meas.*, vol. IM-44, pp. 94–97, 1995.
 [23] C. A. Greenhall, “Recipes for degrees of freedom of frequency stability estimators,” *IEEE Trans. Instrum. Meas.*, vol. 40, pp. 994–999, 1991.
 [24] O. Rioul and M. Vetterli, “Wavelets and signal processing,” *IEEE Signal Processing Magazine*, vol. 84, pp. 14–38, Oct. 1991.
 [25] D. B. Percival and D. A. Howe, “Total variance as an exact analysis of the sample variance,” in *Proc. 29th Ann. PTTI Systems and Applications Meeting*, Long Beach, CA, 1997, pp. 97–105.

APPENDIX A

Consider a series $\{X_{ij}\}$ with k intervals each having n values and means $m_j, j = 1, \dots, k$. Assume $\sum_1^k X = 0$. Consequently

$$\sum_{j=1}^k m_j = 0, \tag{A1}$$

and make

$$X_{ij} = x_{ij} + m_j, \tag{A2}$$

and

$$\sum_{i=1}^m x_{ij} = 0. \tag{A3}$$

The standard variance V of the data run is denoted $V = \frac{SS}{nk}$ where

$$SS = \sum_{j=1}^k \sum_{i=1}^n (x_{ij} + m_j)^2. \tag{A4}$$

$$SS = \sum_{i=1}^n x_{i1}^2 + \dots + \sum_{i=1}^n x_{ik}^2 + nm_1^2 + \dots + nm_k^2 \tag{A5}$$

plus terms of the form

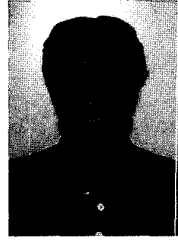
$$\sum_{i=1}^n 2x_{ij}m_j = 2m_j \sum_{i=1}^n x_{ij} = 0.$$

Then

$$V = \frac{1}{nk} SS = \frac{1}{k} \left[\sum_{i=1}^n \frac{x_{i1}^2}{n} + \dots + \sum_{i=1}^n \frac{x_{ik}^2}{n} \right]$$

$$+ \frac{m_1^2}{k} + \dots + \frac{m_k^2}{k} = \frac{1}{k} \sum_{j=1}^k \hat{\varepsilon}_{ss,j}^2(\Delta t) + v_m = \overline{\hat{\varepsilon}_{ss,j}^2(\Delta t)} + v_m, \quad (\text{A6})$$

where $\overline{\hat{\varepsilon}_{ss,j}^2(\Delta t)}$ is the mean of the interval variance $\hat{\varepsilon}_{ss,j}^2(\Delta t)$, and v_m is the variance of the interval means m_j .



David A. Howe has been a physicist with the Time and Frequency Division of NIST, Boulder Laboratories since 1970. He graduated with a B.A. degree from the University of Colorado that year. From 1970-1973, he was with the Dissemination Research Section where he helped coordinate the first lunar-ranging and spacecraft time-synchronization experiment jointly with Jet Propulsion Laboratory. He was with the staff of the Atomic Standards Section from 1973-1984 and did advanced research on cesium and hydrogen maser standards and ruggedized, compact rubidium and ammonia standards. His main activities included spectral estimation using novel digital processing techniques, digital servo design, automated accuracy evaluation of a primary cesium standard, atomic beam analysis, and reduction of oscillator acceleration sensitivity for special applications.

He returned to the Dissemination Research Section in 1984 to lead and implement several global high-accuracy satellite-based time-synchronization experiments with other national laboratories. For this contribution, he was awarded the Commerce Department's Gold Medal in 1990 for advancements in time calibrations among standards laboratories who participate in the maintenance of UTC. Since 1994, he has been a statistical analyst for the Time Scale Section that maintains UTC(NIST) from an ensemble of laboratory atomic frequency standards. He transferred to the Precision Phase Noise Measurement Section in 2000.

He is a member of Sigma Pi Sigma and Phi Beta Kappa societies.