Wavelet Variance, Allan Variance, and Leakage

David A. Howe and Donald B. Percival

Abstract— **Wavelets have recently been a subject of great interest in geophysics, mathematics and signal processing. The discrete wavelet transform can be used to decompose a time series with respect to a set of basis functions, each one of which is associated with a particular scale. The properties of a time series at different scales can then be summarized by the wavelet variance, which decomposes the variance of a time series on a scale by scale basis. The wavelet variance corresponding to some of the recently discovered wavelets can provide a more accurate conversion between the time and frequency domains than can be accomplished using the Allan variance. This increase in accuracy is due to the fact that these wavelet variances give better protection against leakage than does the Allan variance.**

I. Introduction and Summary

THE analysis of a time-ordered set of phase measurements $\{x_t\}$ often falls into one of three categories. The first approach treats $\{x_t\}$ as a series to be expressed in terms of global basis functions such as orthogonal polynomials. The second approach uses the mean square of the second difference of $\{x_t\}$ at various sampling intervals τ to define the Allan variance. The third approach is the windowed discrete Fourier transform (DFT), which is used to estimate the power spectrum $S_x(\cdot)$ for $\{x_t\}$.

The problems with the first approach (polynomial fits) are that many coefficients might be needed to adequately represent the phase measurements, and local features in $\{x_t\}$ can be misrepresented because the coefficients are calculated using global basis functions. Potential problems with the second approach include sensitivity to deterministic drifts and leakage because the transfer function for the Allan variance has substantial sidelobes (this leakage can also occur in the unwindowed DFT). The problems with the third approach are dependence on the chosen window and high variability because the windowed DFT is inherently narrowband.

Wavelet analysis tries to address the above problems with one unified approach. First, wavelet analysis is based upon the discrete wavelet transform, a "time and scale" representation of $\{x_t\}$ that is hierarchical rather than global and hence can represent localized features easily. Second, the wavelet transform is narrowband at low frequencies and broadband at high frequencies. The Allan variance of fractional frequency deviates is in fact a wavelet variance corresponding to the Haar wavelet. Wavelet variances based upon higher order wavelets are natural extensions to the Allan variance and have potential advantages over the Allan variance in terms of leakage and insensitivity to deterministic drifts. Plots of the square root of the wavelet variance versus averaging time (or scale) yield curves analogous to the usual "σ/τ" curve for the Allan variance. As is true for the Allan variance, the wavelet variance can be regarded as an octave-band estimate of the spectrum and hence does not suffer from the high variability of the windowed DFT. Because higher order wavelets provide a better approximation to octave-band filters than does the Haar wavelet, it is easier to translate higher order wavelet variances into reasonable spectral estimates.

Even though wavelets are a relatively new topic, there is already an enormous literature about them—see [1] and references therein. In what follows, we merely motivate the use of the wavelet variance, with emphasis on the problem of leakage (see [3] for more details).

II. Power-Law Noise Processes

The phase difference $\{x_t\}$ between two clocks measured as a function of time is generally modeled by a deterministic part (quantified by a time of set, frequency offset, and frequency drift, but ignored in what follows) and a random part. Historically, pure power-law noise processes have played a vital role in characterizing the random part, which means that the power spectrum $S_x(\cdot)$ for $\{x_t\}$ is proportional to f^α for positive Fourier frequencies, f. Correctly determining α is a primary objective of spectral analysis. Once the exponent α has been determined, we derive estimates of how a clock's timekeeping ability might evolve [4].

III. Narrowband versus Broadband Processing

Spectrum analyzers typically compute a power spectrum for a time series $\{x_t\}$ by windowing the series using $\{h_t\}$, and then taking the squared modulus of the DFT of the windowed series $\{h_t x_t\}$. The purpose of windowing is to reduce a potential bias known as leakage, in which power "leaks" from high power into low power portions of the spectrum. The windowed DFT is inherently narrowband and hence highly variable across frequencies, which makes interpretation of DFT-based spectral estimates somewhat problematic for the novice.

Because narrowband processing is not required for broadband processes such as power-law processes, the time and frequency community handles power-law noise processes using the Allan variance [4]. This variance can be interpreted as the variance of a process after being subjected to approximate

Manuscript received July 1, 1994; revised October 15, 1994.

D. A. Howe is with the Time and Frequency Division, National Institute of Standards and Technology, Boulder, CO 80303 USA.

D. B. Percival is with the Applied Physics Laboratory, University of Washington, Seattle, WA 98195 USA.

IEEE Log Number 9408686.

bandpass filters of constant Q. The Allan variance can be used to form a broadband spectral estimate using well-known conversion schemes [4]. However, while broadband processing produces less variable spectral estimates than those of narrowband processing, both types of processing are subject to leakage. While leakage has long been recognized as a concern in spectral analysis, its importance in broadband processing has not received much attention. One rationale for the wavelet variance is that higher order wavelets effectively address the leakage problem.

IV. Wavelets and the "Scale Domain"

Suppose that $x_0, x_1, \cdots, x_{N-1}$ form a sequence $\{x_t\}$ of N time-ordered phase measurements. Let us define $\sum |x_t|^2 \equiv \mathcal{E}_x$ to be the "energy" in our finite set of measurements. We can then trivially regard $|x_t|^2$ as the contribution to the energy $\mathcal{E}_x$ due to the component of $\{x_t\}$ with time index t. We can also regard $\{x_t\}$ as the "time domain" representation of our phase measurements.

Next, consider the DFT of $\{x_t\}$, namely,

$$X_k \equiv \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} x_t e^{-i2\pi f_k t}, \qquad k = 0, 1, \cdots, N-1$$

where X_k is the kth DFT coefficient and is associated with the kth Fourier frequency $f_k \equiv k/N$. Parseval's theorem tells us that $\sum |X_k|^2 = \mathcal{E}_x$. Hence we can regard $|X_k|^2$ as the contribution to the energy $\mathcal{E}_x$ due to the component of $\{X_k\}$ with frequency index k, and we can regard $\{X_k\}$ as the "frequency domain" representation of our phase measurements. The time and frequency domain representations are equivalent because we can recover $\{x_t\}$ from $\{X_k\}$ using the inverse DFT.

As is true for the DFT, the discrete wavelet transform (DWT) of $\{x_t\}$ preserves the energy $\mathcal{E}_x$ in a set of coefficients; however, unlike the DFT, these coefficients are not indexed by frequency, but rather doubly indexed by time shift j and "scale" τ. The DWT is defined in terms of a "mother wavelet" $\psi(\cdot)$ and an associated "scaling function" $\phi(\cdot)$, where $\psi(\cdot)$ can be any member of a large class of functions satisfying certain stringent conditions [1]. Assuming for convenience that $N = 2^p$ for some positive integer p, we define $\psi_{j,\tau}(\cdot)$ as a shifted and scaled version of $\psi(\cdot)$:

$$\psi_{j,\tau}(t) = \frac{1}{\sqrt{2\tau}} \psi\left(\frac{t}{2\tau} - j\tau\right)$$

where $\tau = 1, 2, 4, \cdots, N/2$ indexes a "power of 2" scale, while $j = 0, 2\tau, 4\tau, \cdots, N - 2\tau$ indexes shifts in time commensurate with scale τ. The DWT coefficients are the doubly indexed series $\{d_{j,\tau}\}$ defined by

$$d_{j,\tau} \equiv \sum_t x_t \psi_{j,\tau}(t)$$

along with $c \equiv \sum x_t \phi(t/N)/\sqrt{N}$. Parseval's theorem tells us that $\sum_{\tau, j} |d_{j,\tau}|^2 + |c|^2 = \mathcal{E}_x$. Hence we can regard $|d_{j,\tau}|^2$ as the contribution to the energy $\mathcal{E}_x$ due to the component of $\{d_{j,\tau}\}$ with time shift index j and scale index τ, and we can regard $\{d_{j,\tau}\}$ as the "scale domain" (or "time/scale" domain) representation of $\{x_t\}$. This representation is fully equivalent to the time and frequency domain representations because we can recover $\{x_t\}$ from $\{d_{j,\tau}\}$ and c using the inverse DWT.

As an example of a scale domain representation, let us set our mother wavelet $\psi(\cdot)$ equal to the Haar wavelet $\psi^{(\mathrm{Haar})}(\cdot)$, which we define here as $\psi^{(\mathrm{Haar})}(t) = -1$ for $0 \le t < 1/2$; $\psi^{(\mathrm{Haar})}(t) = 1$ for $1/2 \le t < 1$; and $\psi^{(\mathrm{Haar})}(t) = 0$ otherwise. The corresponding scaling function $\phi(\cdot)$ is given by $|\psi^{(\mathrm{Haar})}(\cdot)|$ (a relationship unique to the Haar wavelet). For the Haar wavelet, we find that

$$d_{j,\tau} = \frac{\sqrt{\tau}}{\sqrt{2}} [\overline{x}_{(2j+2)\tau-1}(\tau) - \overline{x}_{(2j+1)\tau-1}(\tau)]$$

where

$$\overline{x}_t(\tau) \equiv \sum_{j=0}^{\tau-1} x_{t-j}/\tau.$$

Let us now define the wavelet variance for scale τ as $\sigma_x^2(\tau) \equiv \mathrm{var}\{d_{j,\tau}\}/\tau$. Under the assumption that $E\{d_{j,\tau}\} = 0$ so that the variance of $d_{j,\tau}$ is equal to $E\{d_{j,\tau}^2\}$, a natural estimator of this wavelet variance is

$$\tilde{\sigma}_x^2(\tau) = \frac{1}{\tau} \frac{1}{N/2\tau} \sum_{j=0}^{N/2\tau-1} d_{j,\tau}^2 = \frac{2}{N} \sum_{j=0}^{N/2\tau-1} d_{j,\tau}^2.$$

Specializing to the Haar wavelet, we find that

$$\tilde{\sigma}_x^2(\tau) = \frac{\tau}{N} \sum_{j=0}^{N/2\tau-1} [\overline{x}_{(2j+2)\tau-1}(\tau) - \overline{x}_{(2j+1)\tau-1}(\tau)]^2.$$

If the x_t's represented average fractional frequency deviations rather than phase measurements, then the above would be the well-known "nonoverlapped" estimator of the Allan variance. The Allan variance therefore corresponds to a wavelet variance when the Haar wavelet is used with average fractional frequency deviations. When viewed from the perspective of wavelets, the Allan variance is thus not a "time domain" quantity, but rather is a "scale domain" or "time/scale domain" quantity.

V. Determination of Power-Law Noise Types

As a function of time, two-oscillator phase deviations typically resemble a realization of a composite power-law process, whose spectrum can be described as $S_x(f) = \sum_\alpha h_\alpha |f|^\alpha$, where the summation is over a finite number of different α's (usually a subset of $\alpha = 0, -1, -2, -3$, and -4). For pure power-law processes, there are well-known formulas for converting from the Allan variance to the frequency domain [4]. For composite power-law processes, this conversion can become problematic for the Allan variance. To see this, let $\sigma_x^2(\tau)$ represent this variance. We can then write

$$\sigma_x^2(\tau) = \int_{-1/2}^{1/2} \mathcal{F}_\tau(f) S_x(f)\, df$$

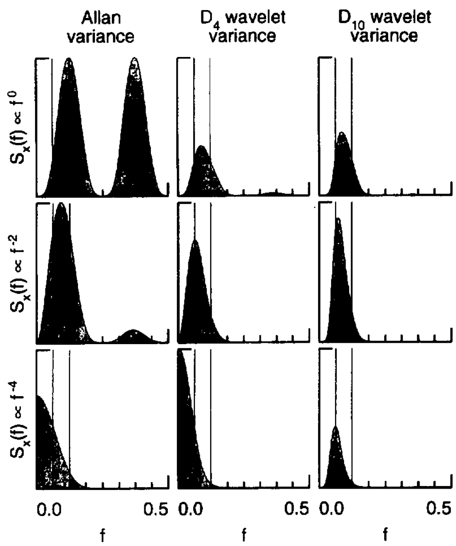

Fig. 1. Modulus squared of the transfer function for the Allan variance (left column) D_4 wavelet variance (middle) and D_{10} wavelet variance (right) times power-law spectra $S_x(\cdot)$ proportional to f^0 (top row), f^{-2} (middle) and f^{-4} (bottom) for scale $\tau = 4$. The integrals of the shaded areas yield the appropriate variance for $\tau = 4$.

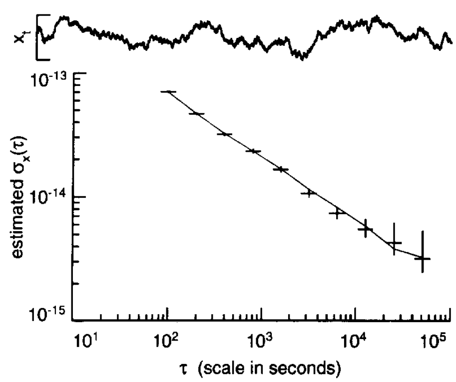

Fig. 2. NIST-7 versus hydrogen maser phase measurements (top plot) and estimated $\sigma_x(\tau)$ versus τ (bottom plot) for the Allan variance (connected curve) and the D_4 wavelet variance (crosses).

where $\mathcal{F}_\tau(\cdot)$ is the modulus squared of the transfer function associated with the Allan variance at scale τ [2]. The shaded areas in the left column of plots in Fig. 1 show the product $\mathcal{F}_\tau(f)S_x(f)$ versus f for the Allan variance for three pure power-law spectra and scale $\tau = 4$. The integral of each shaded area gives $\sigma_x^2(4)$ for the appropriate pure power-law process. In the octave-band interpretation of the Allan variance, $\sigma_x^2(4)$ should roughly reflect the power in the spectrum in the frequency interval $[1/4\tau, 1/2\tau]$. For $\tau = 4$, this interval is $[1/16, 1/8]$ and is delineated on each plot by a pair of thin vertical lines. If the filters associated with the Allan variance were perfect octave-band filters, the shaded area in each plot would be entirely contained between the vertical lines. The amount of the shaded area that lies outside of the vertical lines

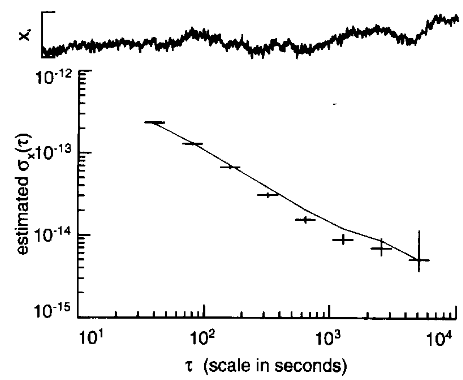

Fig. 3. Time-synchronization phase measurements using the NIST satellite two-way transfer modem configured in an in-cabinet loop test (top plot) and estimated $\sigma_x(\tau)$ versus τ (bottom plot) for the Allan variance (connected curve) and the D_4 wavelet variance (crosses).

represents the contribution to the Allan variance attributable to leakage. These plots indicate that there is substantial leakage for the Allan variance. This leakage is most pronounced for white PM ($S_x(f) \propto f^0$, in the top left plot). If we now consider a composite power-law process dominated between the vertical lines by a power-law with a different exponent than the one displayed in the plots of Fig. 1, we can see the potential problem with leakage, namely, that the integral of $\mathcal{F}_\tau(f)S_x(f)$ (the Allan variance) can be influenced mainly by values of f outside of the vertical lines and hence cannot accurately reflect the values of $S_x(f)$ between the vertical lines.

Fig. 1 also shows corresponding plots for the wavelet variance using the D_4 (middle column) and D_{10} (right column) "extremal phase" wavelets [1]. The D_4 wavelet was chosen because it is "one order up" from the Haar wavelet (and closely mimics the performance of the so-called modified Allan variance [4]), while the D_{10} wavelet is an example of an even higher order wavelet. The D_{10} wavelet variance for scale $\tau = 4$ reflects the spectrum in the passband $[1/16, 1/8]$ to a much better degree than the other variances because the shaded areas are concentrated between the vertical lines to a higher degree for the D_{10} wavelet variance.

VI. Examples

We present two examples using the wavelet variance with phase measurements. The top plot of Fig. 2 shows phase measurements recorded every 100 s over a 3.7 day interval comparing NIST-7 with a hydrogen maser. The bottom plot shows the estimated Allan standard deviation (square root of the Allan variance) versus scale τ (the connected curve) and also the estimated D_4 wavelet standard deviation versus τ (the crosses). The vertical portion of each cross delineates a "one sigma" (68.3%) confidence interval for the true D_4 wavelet standard deviation. The Allan and D_4 wavelet standard deviations agree fairly well here, although there are two scales (τ = 3200 and 6400 s) for which the Allan standard deviation is just inside the confidence limits for the D_4 wavelet standard deviation. Use of the D_4 wavelet here tells us that leakage is not a major problem with the Allan variance for these phase measurements.

The top plot of Fig. 3 shows phase measurements recorded every 40 s over a half day interval reflecting time-synchronization using the NIST satellite two-way transfer modem. The bottom plot here shows the same quantities as in the bottom plot of Fig. 2. While the Allan and D_4 standard deviations agree quite well in the smallest three and largest scales, there is significant difference in the middle three scales; moreover, the difference is consistent with leakage in the Allan variance since the Allan variance is higher than the D_4 variance.

REFERENCES

[1] I. Daubechies, *Ten Lectures on Wavelets.* Philadelphia: SIAM, 1992.
[2] C. A. Greenhall, "A shortcut for computing the modified Allan Variance," in*Proc. 46th Ann. Freq. Control Symp.*, 1992, pp. 262–264.
[3] D. A. Howe and D. B. Percival, "Wavelet analysis for synchronization and timekeeping," *Proc. 48th Ann. Freq. Control Symp.*, submitted for publication..
[4] D. B. Sullivan, D. W. Allan, D. A. Howe, and F. L. Walls, Eds., "Characterization of clocks and oscillators," NIST Tech. Note 1337, 1990.