INVESTIGATING THE BIASES IN ALLAN AND HADAMARD VARIANCES AS MEASURES OF MTH ORDER RANDOM STABILITY

Victor R. Reinhardt
Raytheon Space and Airborne Systems, El Segundo, CA, USA

Abstract

Allan and Hadamard variances (AHVs) generated from data consisting of purely random noise are well-known as mean square measures of M^{th} order random stability (MORS) over the difference interval τ. When data contain deterministic drift (intrinsic aging plus environmentally induced temporal changes) as well as noise, however, it is also well known that AHVs can be biased measures of MORS. In such cases, one generally minimizes this bias by "removing" the drift from the data in question using auxiliary data fitting methods and then generating the AHV using the drift-removed residuals. This paper investigates the nature of one aspect of the residual AHV bias that remains after such drift-removal: kernel or $K_{stat}(f)$-bias. This bias exhibits itself in the spectral integral representation of a drift-removed AHV as an alteration of the spectral kernel $K_{stat}(f)$ that relates the AHV to S(f), the power spectral density of the random data component from the non-drift removed AHV. This $K_{stat}(f)$-bias occurs because a fitting process, by its very nature, removes some of the noise along with the drift in the fit residuals, especially when negative power law noise is present. In this paper, charts of $K_{stat}(f)$-bias are generated as a function of τ/T for various AHV statistics (overlapping, total, and modified), various drift-removal methods (Greenhall, various polynomial order least squares fits, and environmental drift removal), and various power law noise processes. Selected drift-removed $K_{stat}(f)$ charts are also examined to provide intuitive insight into why drift-removal $K_{stat}(f)$-bias behaves as it does. To aid in the generation of the numerous drift-removed bias charts in this paper, an efficient numerical technique is introduced for numerically computing drift-removed $K_{stat}(f)$ directly in the time-domain from simple t-domain definitions of AHV statistics and drift-removal methods. This technique avoids the need for computationally intensive phase randomization of a t-domain input in order to eliminate errors that occur when the t-domain input is not wide-sense stationary. Finally, the paper demonstrates two important results when higher-order negative power law noise is present: (a) that noise whitening (increasing the order of a polynomial drift-removal fit until the data residuals appear uncorrelated) greatly increases the $K_{stat}(f)$-bias in drift-removed AHVs and leads to misleadingly-low estimates of the MORS, and (b) that the removal of temporally complex environmental drift can also generate significant $K_{stat}(f)$-bias.

INTRODUCTION

ALLAN AND HADAMARD VARIANCES (AHVS)

Allan and Hadamard Variances (AHVs) can be written as mean square (MS) averages of the M^{th} order Δ-measure $\Delta^M(\tau)x(t)$ for $M = 2$ and $M = 3$ respectively **[1]**. Here, $\Delta^M(\tau)x(t)$ is the M^{th} application of the 1^{st}-order forward-difference $\Delta(\tau)x(t) = x(t+\tau) - x(t)$, where $x(t)$ is the time error, and t is an error-free observation time. Figure 1 illustrates $\Delta^M(\tau)x(t_n)$ for $M = 1$ and $M = 2$, where $x(t_n)$ represents N uniformly-spaced data samples of $x(t)$ over a total observation time $T = N\tau_o$, τ_o is the sampling interval, and τ is

assumed to be an integer multiple of τ_o **[1]**.

In this paper, $\sigma_y(\tau)^2$ is the fractional frequency error Allan variance, whether it is expressed in terms of $y(t)$, the fractional frequency error or $x(t)$, and $\sigma_x(\tau)^2 = (\tau^2/3)\sigma_y(\tau)^2$ is the time error Allan variance, where $\sigma_y(\tau)^2$ can be any statistic of the Allan variance, rather than just the modified variance **[1]**. When used alone, as in $\sigma_y(\tau)^2$ or $\sigma_x(\tau)^2$, these variances will indicate the theoretical variance, as in **[1]**

$$\sigma_x(\tau)^2 = 6^{-1}\mathcal{E}[x(t_{n+2m}) - 2x(t_{n+m}) + x(t_n)]^2 \tag{1}$$

where $\mathcal{E}$ indicates an ensemble average **[2]**. Multiple-sample Allan variance statistics, such as overlapping, total, and modified statistics **[1]**, are designated by Ov $\sigma_x(\tau)^2$, Tot $\sigma_x(\tau)^2$, Mod $\sigma_x(\tau)^2$, etc., or by unabbreviated prefixes. For example, the overlapping $\sigma_x(\tau)^2$ is given here by

$$\mathrm{Ov}\,\sigma_x(\tau)^2 = 6^{-1}(N-2m)^{-1}\Sigma_{n=1}^{N-2m}\mathcal{E}[x(t_{n+2m}) - 2x(t_{n+m}) + x(t_n)]^2 \tag{2}$$

In Equation 2, note that the $\mathcal{E}$-averaged AHV form is used for multiple-sample statistics as well as theoretical ones, so that we can write any AHV statistic $\sigma_{stat}(\tau)^2$ in terms of the spectral integral **[1, 3, 4]**

$$\sigma_{stat}(\tau)^2 = \int_0^{f_h} df\, K_{stat}(f)\, S_x(f) \tag{3}$$

$K_{stat}(f)$ in Equation 3 is a spectral kernel that relates $\sigma_{stat}(\tau)^2$ to $S_x(f)$, the double sideband power spectral density of $x(t)$ **[1]**. An example of such an AHV kernel is $K_{stat}(f) = (8/3)\sin^4(\pi f\tau)$, the kernel for the theoretical $\sigma_x(\tau)^2$ **[1, 5]**. Note that all kernels in this paper relate $\sigma_{stat}(\tau)^2$ to $S_x(f)$, rather than to $S_y(f)$. In Equation 3, f_h is an infinitely sharp high-pass frequency cut-off used to approximate the effect of system filtering on the AHV **[1]**. This approximation is suitable for the purposes of this paper.

The theoretical fractional frequency "Hadamard" **[6]** or Picinbono **[7]** variance is designated in the paper by ${}_H\sigma_y(\tau)^2 = (6\tau^2)^{-1}\mathcal{E}\Delta^3(\tau)x(t)$ **[6]**. Similarly, ${}_H\sigma_x(\tau)^2$ indicates the time error Hadamard variance given by ${}_H\sigma_x(\tau)^2 = (3\tau^2/10){}_H\sigma_y(\tau)^2$. This makes ${}_H\sigma_x(\tau)^2$ equal to the standard variance of $x(t)$ when $x(t)$ consists of uncorrelated (white) noise. The term Hadamard is in quotes here, because the form of the "Hadamard" variance used here **[6]** is proportional to that introduced by Picinbono **[7]** and is not the same as the variance first introduced by Hadamard **[8]**. For more detail on the Hadamard variance form as used here see **[6]**. Again, Ov, Tot, or Mod (or their unabbreviated prefixes) in front of ${}_H\sigma_y(\tau)^2$ or ${}_H\sigma_x(\tau)^2$ indicate these statistics for the Hadamard variance.

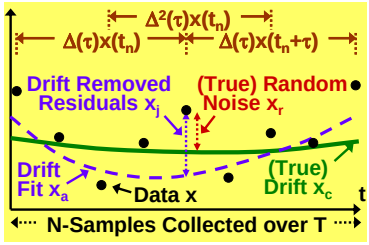

Figure 1. Δ-measures, random error, and deterministic drift.

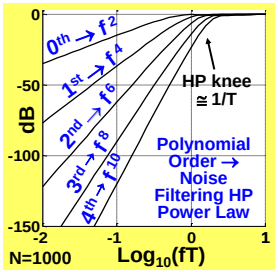

Figure 2. Noise highpass (HP) filtering in the residuals from polynomial ULSF drift removal.

MS M^TH ORDER RANDOM STABILITY (MORS) AND AHVS

In this paper, we define MS M^{th} order random stability (MORS) of a random noise variable $x_r(t)$ simply as a normalized version of $\mathcal{E}[\Delta^M(\tau)x_r(t_n)]^2$, where $x_r(t_n)$ is a zero-mean random process (see Figure 1) **[2]**. We also assume here that $x_r(t_n)$ is wide-sense stationary (WSS) **[2]**. One can generalize this definition **[5]**, but we will not need this more general definition in this paper. Thus, in this paper, Allan and Hadamard variance statistics generated from $x_r(t_n)$ are perfect unbiased measures of 2nd and 3rd order $x(t)$-MORS respectively.

In Figure 1, we note that a typical data set $x(t_n)$ is composed of two different data components: the (true) random noise $x_r(t_n)$ and the (true) deterministic drift $x_c(t_n)$, which we define as intrinsic temporal aging plus environmentally induced temporal changes **[9]**. It is well-known that Allan variances are insensitive to 1st order $x(t)$ polynomial drift and Hadamard variances are insensitive 2nd order $x(t)$ polynomial drift **[5]**. In such cases, AHVs generated from raw drift-containing data are unbiased measures of MORS. However, when data does not contain such drift, it is also well known that AHVs generated from raw $x(t_n)$ drift-containing data will be biased as measures of MORS because of drift contamination **[5].** That is, AHVs generated from such raw $x(t_n)$ data will not be ensemble-mean equivalent to those generated from purely random noise. Finally, we note that the term *bias* in this paper will refer to a multiplicative deviate bias, that is, the ratio of the square root of an $\mathcal{E}$-averaged AHV statistic generated from the data under consideration to that generated from its true noise component alone (the true MORS).

In such drift-sensitive situations, one generally attempts to minimize this bias by "removing" the drift with an auxiliary data fitting method ($x_a(t_n)$ in Figure 1). Then the AHV is generated using the drift-removed residuals ($x_j(t_n)$ in Figure 1) rather than the raw data. Examples of such auxiliary data fitting methods are

- unweighted least square fits (ULSFs) **[10]** using polynomial fit model functions **[3]**, logarithmic models for crystal oscillators **[13-14]**, or environmental data **[9],**
- the Greenhall frequency drift estimator **[11]**,
- Kalman filters **[12]**,
- and when the functional form of the drift is unknown, noise whitening (increasing the order of a polynomial ULSF until the data residuals appear uncorrelated).

It is known, however, that such drift removal methods are imperfect and leave a residual bias, especially in the presence of negative power law (neg-p) noise **[3, 11]**. This paper investigates the nature of this residual drift-removed AHV bias. It is finally noted that we will not address Kalman drift removal here, though this is an important subject in and of itself **[4]**.

NOISE OR KERNEL RESIDUAL BIAS

Residual drift-removal bias can be separated into two components: model error bias and noise or kernel bias **[5]**. Model error bias occurs when the model function used in the fit incorrectly models the functionality of the true drift over T. That is, there is no parameter adjustment for the fit model that allows the fit $x_a(t_n)$ to exactly reproduce the true drift $x_c(t)$ over T. In this case, some of the drift remains in the residuals $x_j(t_n)$ after the fit is performed and this contaminates the drift-removed AHV as a measure of MORS. If the drift model can reproduce $x_c(t_n)$ for some parameter adjustment, the model is linear in the adjustable parameters, and the fitting process is well-behaved, all of the true drift $x_c(t_n)$ will be removed in the residuals, regardless of effect of the noise in the data on $x_a(t_n)$. This is because a parameter-linear fitting process obeys the Superposition Principal, and thus $x_a(t_n)$ can be separated into a perfect noiseless solution and a noise-only solution **[3]**. For the remainder of this paper, we will assume that this model error bias is negligible and focus on the second bias component, noise or kernel bias.

Noise or kernel bias occurs because data-fitting methods, by their very nature, can't completely separate drift from noise in a finite length data set **[2-3, 10]**. This is especially true when highly-correlated neg-p noise is present **[3]**. To understand this, we note that data-fitting methods are in reality filters that rely on the orthogonality or lack or correlation between the drift and noise in the N-dimensional space of the data to separate the drift from the noise **[10]**. Thus, noise will contaminate the fit solution to the extent that it is correlated with the drift in a data set **[3, 10]**. The consequence of this is that the drift-correlated portion of the noise is removed from the residuals, biasing the drift-removed AHV as a measure of MORS **[5]**. In the variance of the residuals, this noise removal can be expressed as a kernel $K_{res}(f)$ that filters $S_x(f)$ in a spectral integral similar to Equation 3 **[3]**. Figure 2 plots $K_{res}(f)$ generated by M^{th} order polynomial ULSF drift removal for various orders **[3]**. Note that $K_{res}(f)$ for M^{th} order polynomial ULSF drift removal acts like a $2M^{th}$ order highpass filter with a knee frequency on the order of $1/T$ **[3]**.

For a drift-removed $\sigma_{stat}(\tau)^2$, $K_{res}(f)$ modifies $K_{stat}(f)$ from its form without drift removal to create bias. We will call this noise, kernel, or $K_{stat}(f)$-bias. The nature of this $K_{stat}(f)$-bias for various AHV statistics, drift removal methods, and power law $S_x(f)$ **[1]** is the subject of this paper.

AHV DRIFT-REMOVED KERNEL BIASES TO BE INVESTIGATED IN THE PAPER

In the paper, we investigate $K_{stat}(f)$-bias as a function of τ/T for

- Overlapping, total, and modified Allan variance statistics;
- The overlapping Hadamard variance statistic;
- Greenhall frequency drift removal;
- 2^{nd} and 6^{th} order polynomial ULSF phase drift removal;
- Environmental frequency drift removal;
- The integer (p) power law noise orders p = 0, -1, -2, -3 in $S_x(f) = f^p$.
-

Also note that **[1]**

- p = 0 is associated with white phase noise,
- p = -1 is associated with flicker of phase noise,
- p = -2 is associated with random walk of phase noise,
- p = -3 is associated with flicker of frequency noise, and
- $p < 0$ defines negative power law (neg-p) noise (integer neg-p noise if p is a negative integer).

We further note that 6^{th} order polynomial drift removal is used here to represent the $K_{stat}(f)$-bias that would occur when noise whitening techniques are used to identify unknown temporally complex deterministic drift in data. In noise whitening, one increases the order of a polynomial fit until the residuals become uncorrelated. This based on the assumption that the true noise is sample-uncorrelated, which is of course not true for negative power law noise **[3-4]**. For temporally-complex drift, polynomial orders of greater than 3 are typically required to achieve such uncorrelated residuals, and one can show that 6^{th} order polynomial drift-removal $K_{stat}(f)$-bias is typical of orders greater than 3.

AN EFFICIENT TECHNIQUE FOR NUMERCIALLY COMPUTING DRIFT-REMOVED KERNELS

Once a drift-removed $K_{stat}(f)$ is known, computing AHV kernel-bias for a given $S_x(f)$ using Equation 3 is straightforward. However, generating such $K_{stat}(f)$ in the frequency domain, while theoretically possible,

can be mathematically cumbersome **[3]**. Since both AHV and drift-fitting definitions are simple to represent in the time-domain, a technique that generates f-domain drift-removed $K_{stat}(f)$ using t-domain techniques is desirable.

The principal behind t-domain generation of such drift-removed $K_{stat}(f)$ is straightforward. Let us assume we have a wide-sense stationary (WSS) **[2]** ensemble of discrete t-domain data sets that has a discrete f-domain PSD matrix equivalent to the continuous PSD $S_x(f) = \delta(f-f_o)$ for $0 < f_o < f_h$. Observing Equation 3, one can see that numerically computing $\sigma_{stat}(\tau)^2$ in the t-domain using this ensemble and t-domain definitions of the AHV and drift-removal fitting method in question yields the discrete equivalent of the desired drift-removed $K_{stat}(f)$.

To find such an ensemble, let us start with the simple real harmonic data set

$$x(t_n) = 2\cos(\omega_o t_n + \phi) \tag{4}$$

where ϕ is some arbitrary phase offset. (Note that AHV definitions assume $x(t)$ and $y(t)$ are real quantities **[1]**.) We further note that Equation 4 has a discrete autocorrelation matrix equivalent to the continuous autocorrelation function

$$R_x(t+0.5\tau,\, t-0.5\tau) = \mathcal{E}\, 4\cos(\omega_o(t+0.5\tau+\phi))\cos(\omega_o(t-0.5\tau+\phi)) = 2\cos(\omega_o(2t+2\phi)) + 2\cos(\omega_o\tau) \tag{5}$$

which is <u>not</u> WSS, but is non-stationary (NS) **[2]**. The NS nature of R_x in Equation 5 is indicated by its extra t-dependent term $2\cos(\omega_o(2t+2\phi))$ **[2]**. Because of this non-stationarity, a $\sigma_{stat}(\tau)^2$ computed using Equation 4 will deviate from the true WSS $K_{stat}(f_o)$ for the AHV in question (for both drift removed and non-drift removed cases). Figure 3 shows $K_{stat}(f)$ for the non-drift removed Ov $\sigma_x(\tau)^2$ computed in the t-domain from $\sigma_{stat}(\tau)^2$ using $x(t_n) = 2\cos(\omega t_n)$ and $2\cos(\omega t_n - \pi/2)$ (or $2\sin(\omega t_n)$). Here, the computed results are compared against $K_{stat}(f)$ generated from its analytical formula labeled "Theoretical" **[1, 3]**.

One way to correct this problem is to average such $\sigma_{stat}(\tau)^2$ over an ensemble of data sets with the phase variable ϕ ranging from 0 to 2π (or its equivalent, to average over an ensemble with start times t_0 varying over a full cosine period). Averaging over this ensemble causes the 1st NS term in Equation 5 to become zero leaving only the 2nd WSS term $\cos(\omega_o\tau)$. When this ensemble is used, one finds that the averaged $\sigma_{stat}(\tau)^2$ is equal the true WSS $K_{stat}(f_o)$. However, such a complete ϕ-randomization is computationally intensive because of the large number of ensemble members involved.

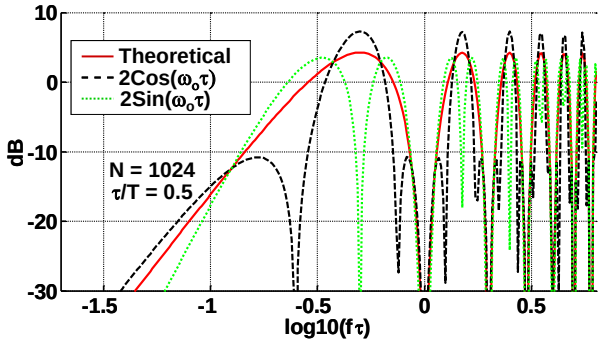

Figure 3. Deviations from the theoretical (analytical) WSS $K_{stat}(f)$ for the non-drift removed Ov $\sigma_x(\tau)^2$ generated in the t-domain using $x(t_n) = 2\cos(\omega t_n)$ and $2\sin(\omega t_n)$.

To reduce the computational load, we note that a simple two-data-set ensemble with $\phi = 0$ and $\phi = -\pi/2$ will also average the first non-WSS term in Equation 5 to zero. This is equivalent to averaging $\sigma_{stat}(\tau)^2$ over

the two ensemble members $x(t_n) = 2\cos(\omega_o t_n)$ and $x(t_n) = 2\sin(\omega_o t_n)$. Using this ensemble, one can show that all non-drift removed AHV $K_{stat}(f)$ are equal to their correct analytically computed theoretical counterparts (see Figure 11 as an example).

A modification of this technique that can be used when the fit model and method are analytically well-behaved is as follows: (a) set $x(t_n)$ to $\exp(j\omega_o t_n)$, (b) replace real squares in the t-domain definitions of the AHV in question with absolute value squares, and (c) use complex arithmetic in the computation of $\sigma_{stat}(\tau)^2$. This is equivalent to averaging over the real ensemble members $2\cos(\omega_o t_n)$ and $2\sin(\omega_o t_n)$ using real arithmetic. When using simulation tools that utilize complex notation, this modification requires less code and runs faster than the real method. Note, however, that this complex method cannot be used for generating drift-removed $K_{stat}(f)$ when the drift-model contains functions such as $A\cdot\ln(B\cdot t+1)$ (as in long-term crystal oscillator drift models **[13-14]**) which are not analytically well-behaved (the log function has a complex branch cut).

DRIFT-REMOVED AHV KERNEL-BIAS

DRIFT-REMOVED ALLAN DEVIATE BIAS

Modified Allan Deviate Bias. Figure 4(a) shows Greenhall drift-removed Mod $\sigma_x(\tau)$ or Mod $\sigma_y(\tau)$ bias (curves are the same) for $r \cong 0.1$ **[11]** plotted versus τ/T. Shown are bias curves generated using the paper's numerical technique for $p = 0$, -1, -2, and -3, for all possible τ values, and for $N = 1018$ to 1036. Discrete bias points published by Greenhall in 1997 for $N \rightarrow \infty$ **[11]** are also shown. The ability to compute such a large number of $K_{stat}(f)$ and their associated bias curves demonstrates the efficiency of the paper's numerical computation technique.

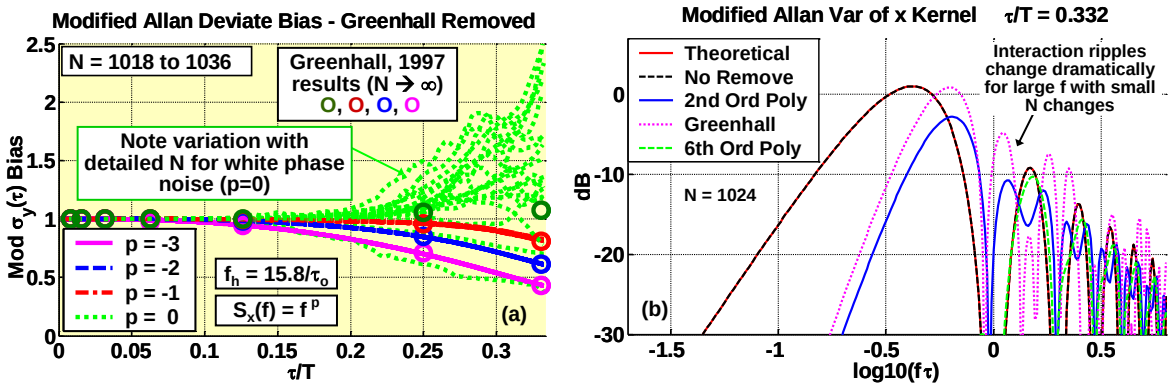

Figure 4. Greenhall drift-removed Mod $\sigma_x(\tau)$ and $\sigma_y(\tau)$ bias vs τ/T and $K_{stat}(f)$ for Mod $\sigma_x(\tau)^2$.

From Figure 4(a), one can see that the paper's results are in agreement with those of Greenhall for $p = -1$, -2, and -3, but not for $p = 0$. This fact at first puzzled the author until N values from 1018 to 1036 were run. This showed that $p = 0$ bias at appreciable τ/T fluctuates dramatically as N changes by small amounts, as can be observed in Figure 4(a). These large fluctuations can be explained by observing the Figure 4(b) graph of $K_{stat}(f)$ for Greenhall drift-removed Mod $\sigma_x(\tau)^2$ with $\tau/T = 0.332$ and $N = 1024$. (Figure 4(b) also shows Mod $\sigma_x(\tau)^2$ $K_{stat}(f)$ for 2nd and 6th order polynomial ULSF drift removal, which will be discussed later.) Note from the figure that the nodes and anti-nodes in the Greenhall drift-removed $K_{stat}(f)$ at higher f vary quite a bit from those of the non-drift-removed $K_{stat}(f)$. On generating similar graphs for various N, one observes these nodes and anti-nodes change dramatically as N changes. Thus, the bias changes so much with N in Figure 4(a), because the spectral $\sigma_{stat}(\tau)^2$ integral is sensitive to these higher f fluctuations

when integrating spectrally flat $p = 0$ $S_x(f)$. On the other hand, neg-p $S_x(f)$, which spectrally roll off, will have spectral $\sigma_{stat}(\tau)^2$ integrals that are less sensitive to these higher-f ripple fluctuations, and thus do not show such bias variations with N. Finally however, it is noted that that this N sensitivity for $p = 0$ bias, while interesting from a theoretical point of view, is somewhat moot, since neg-p noise usually dominates $\sigma_{stat}(\tau)^2$ for appreciable τ/T.

Figures 5(a) and 5(b) show 2nd order and 6th order polynomial ULSF drift-removed Mod $\sigma_x(\tau)$ and Mod $\sigma_y(\tau)$ bias vs τ/T respectively. Again for $p = 0$, one can see that there is bias variation with N as there was for Greenhall removal. For $p = -2$ and -3 noise, we also note that the bias for 2nd order poly drift removal is comparable but slightly worse than that for Greenhall drift removal and, for 6th order poly drift removal, that the bias is larger than those for the other drift removal methods. The 6th order removal bias also becomes appreciable for much lower values of τ/T. This 6th order polynomial ULSF drift removal bias is representative of what could occur if noise whitening were used to remove the drift. Finally, the paper's results confirm Greenhall's conclusions that his drift removal method generates the least Mod $\sigma_y(\tau)$ bias for higher neg-p orders **[11]**.

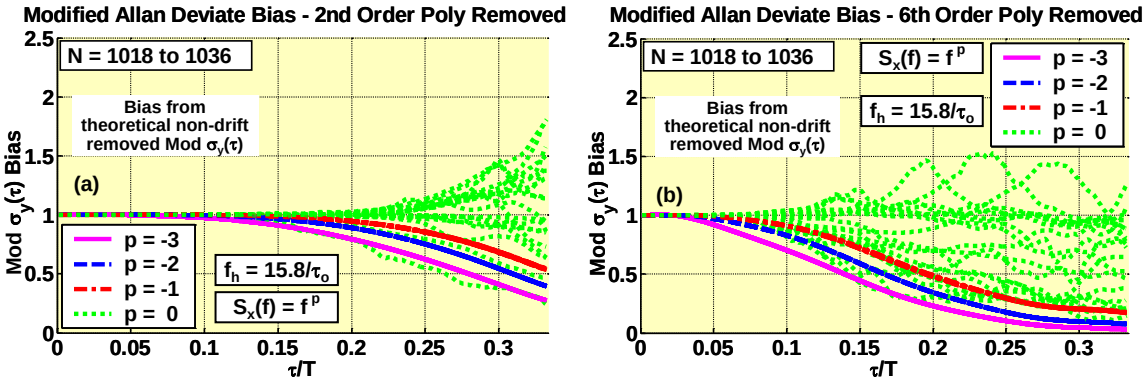

Figure 5. The 2nd and 6th order drift-removed Mod Allan bias vs τ/T.

Total and Overlapping Allan Deviate Bias. Figure 6(a) shows the bias in the non-drift-removed total Allan deviate vs τ/T, where the bias here is defined as the ratio between the total Allan deviate and theoretical non-total Allan deviate. The fact that the non-drift-removed total Allan deviate has a small bias may be surprising to some, but this bias is generated by the data folding process used in the total statistic **[1]**. From Figure 6(b), which graphs $K_{stat}(f)$ for the non-drift-removed Tot $\sigma_x(\tau)$ at its maximum τ/T (along with other kernels), one can explain the mechanics of the bias caused by this data folding process. In the folding process, the number N of data points is artificially extended by adding two order-reflected versions of the original data set on either side of the original measured data set **[1]**. The Allan (or Hadamard) statistic is then averaged over this extended data set. For appreciable τ/T, this data folding process has the effect of modulating the sample spacing relative to the original set of the individual M^{th} order differences in the statistic. One effect of this modulation is that the average sample spacing over the data set is smaller than the specified τ at appreciable τ/T. From Figure 6(b), one can see this effect in the shift in position of the total $K_{stat}(f)$ f^4 slope at low fT from that of the theoretical non-total $K_{stat}(f)$. This generates some bias. Second at higher f, one can see from Figure 6(b) that this τ-modulation also fills in the nulls of the theoretical non-total $K_{stat}(f)$, which generates more bias. Finally, note that the Tot $\sigma_x(\tau)$ bias for $p = 0$ noise exhibits fluctuations with N similar to those in Mod $\sigma_x(\tau)$ though smaller in size.

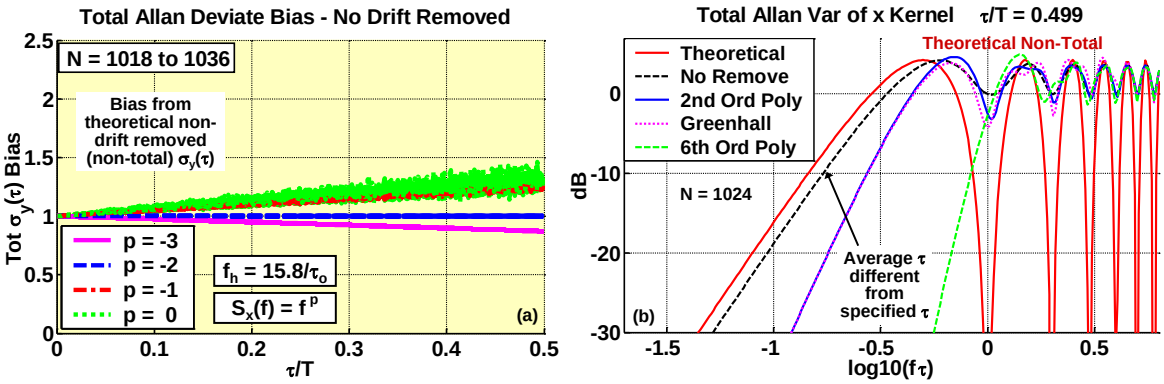

Figure 6. Non-drift-removed total Allan bias vs τ/T and total $K_{stat}(f)$.

Figure 7, Figure 8, and Figure 9 show $K_{stat}(f)$-bias vs τ/T for the total and overlapping Allan deviates with Greenhall, 2nd order polynomial ULSF, and 6th order polynomial ULSF drift removal. Here, for appreciable τ/T, we note that Greenhall drift removal also produces a lower bias than that for 2nd order polynomial drift removal when p = -2 and -3.

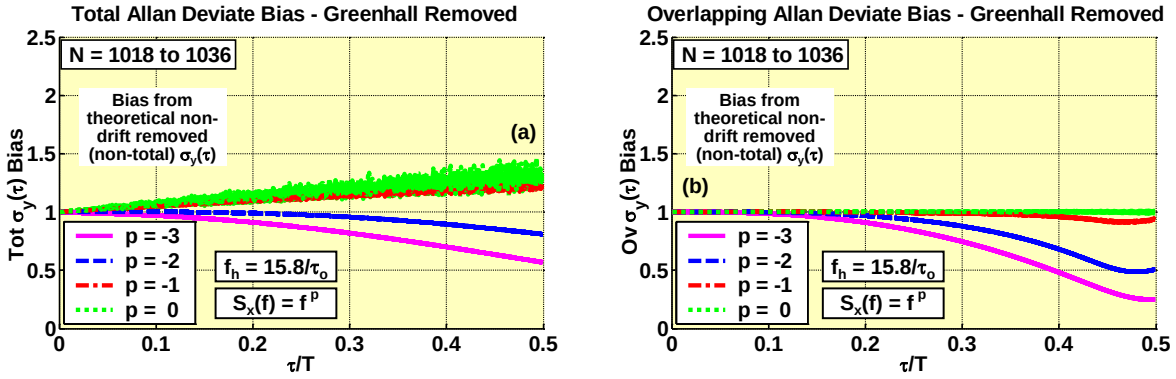

Figure 7. Greenhall drift-removed total and overlapping Allan bias vs τ/T.

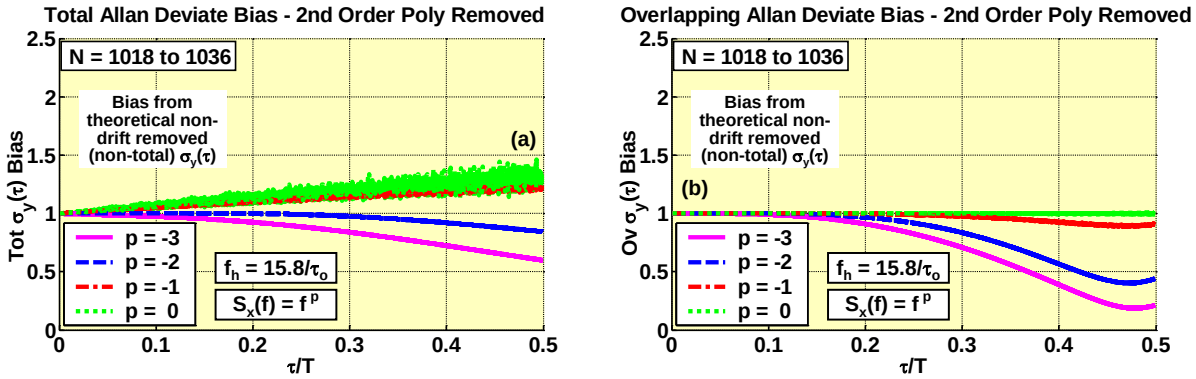

Figure 8. The 2nd order polynomial drift-removed total and overlapping Allan bias vs τ/T.

The Effects of Noise Whitening and Complex Aging and Environmental Drift Removal. From Figure 5 and Figure 9, one can observe that 6th order polynomial drift removal produces significantly more bias

and becomes appreciable at significantly lower values of τ/T than that for the other drift-removal methods. Also observe that the total deviate is not as effective in reducing 6th order $K_{stat}(f)$-bias for p = -2 and -3 noise as it is for the other drift-removal methods. Again, 6th order polynomial removal is representative of the order one would need to whiten the data residuals for some unknown temporally complex deterministic drift. It is noted that polynomial orders greater than 3 produce similar bias results. In noise whitening, one increases the polynomial fit order until the data residuals $x_j(t_n)$ appear uncorrelated. Thus, the noise whitening process adds the part of the noise that is correlated with the fit model to the deterministic drift in the fit solution (the Orthogonality Principle for ULSF residuals **[10]**). As one increases the absolute value of p, one can show that neg-p noise becomes more and more correlated with polynomial fit models **[3-4]**, and the whitening process thus removes more and more of the noise from the data residuals. This leads to an increase in AHV $K_{stat}(f)$-bias. The p = -1 noise is a marginal case because of its weaker correlation properties **[3-4]**. Finally, this correlation also produces misleading deterministic drift estimates. It is known in the literature (but sometimes ignored in practice) that noise whitening or similar fitting methods are not valid when the error correlates with the fit model **[2-4, 10]**. The conundrum here for neg-p noise is that such correlation cannot be reduced by extending T (as is the case for more well-behaved correlated noise processes), because neg-p noise has an infinite correlation time **[3-4]**.

$K_{stat}(f)$-bias effects can also occur when attempting to remove functionally complex aging, such as long-term crystal oscillator frequency aging, with higher order polynomials. Direct modeling of crystal oscillator long-term frequency $A{\cdot}\ln(B{\cdot}t{+}1)$ functional terms **[13-14]** has been investigated by the author. Unfortunately, this investigation was not completed in time for this paper, because of difficulties associated with the non-analytic nature of the $A{\cdot}\ln(B{\cdot}t{+}1)$ function. Thus at this point, one can only surmise the bias impact of direct $A{\cdot}\ln(B{\cdot}t{+}1)$ drift removal from the 6th order polynomial results, which is representative of the polynomial order required to reduce model error in long-term crystal oscillator frequency aging removal when using a polynomial drift model.

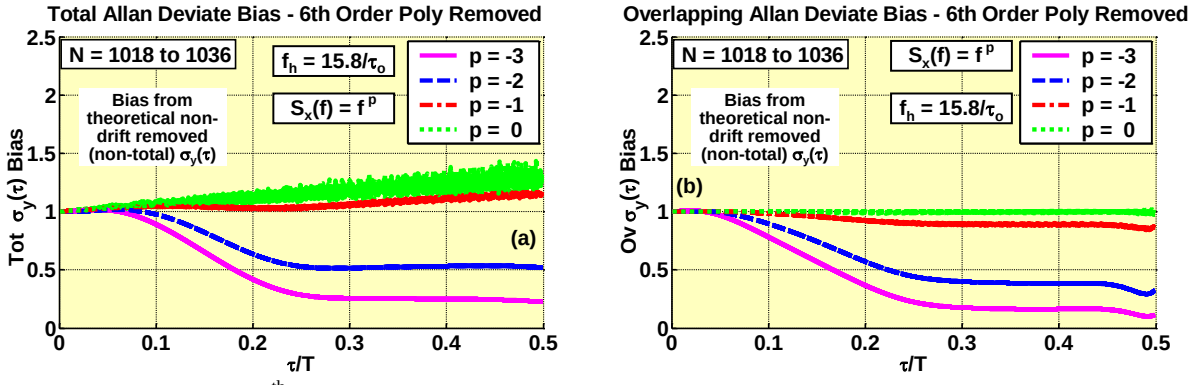

Figure 9. The 6th order polynomial drift-removed total and overlapping Allan bias vs τ/T.

Removing temporally complex environmental signatures from data can also cause $K_{stat}(f)$-bias. Figure 10 shows the $K_{stat}(f)$-bias in such an environmentally drift-removed overlapping $\sigma_x(\tau)^2$ $K_{stat}(f)$. Here, frequency drift due to the temperature profile shown was removed using a single parameter ULSF and the temperature signature. Note that the bias results here are similar to that in Figure 8 for 2nd order polynomial removal. The author has found that other environmental signatures can generate quite different bias results, varying from no bias to extremely high bias, depending on the specifics of the temperature signature. However, this investigation is not been completed to the point that its results can be published here.

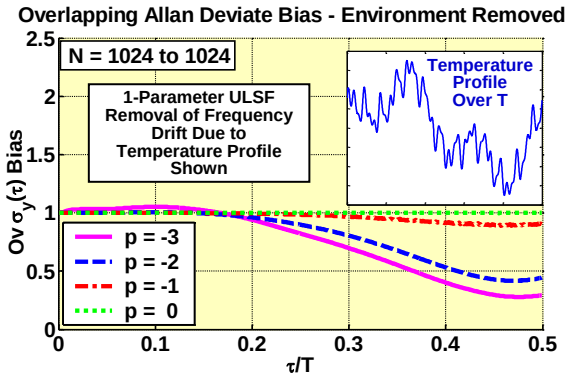

Figure 10. The effect of removing a temperature profile on the overlapping Allan deviate.

Graphical Examination of Drift-Removed Allan of x Kernels. Figures 11-13 show $\sigma_x(\tau)\, K_{stat}(f)$ for the Allan variance statistics and drift removal methods we having been discussing. These are shown for various τ/T values and N = 1024. Also shown are non-drift-removed $K_{stat}(f)$ computed both numerically and theoretically for comparison. Note that all the numerically computed and theoretical non-drift-removed $K_{stat}(f)$ agree, which is another verification of the validity of the numerical technique.

Figure 11 shows $K_{stat}(f)$ statistics for τ equal to one data sample τ_o (τ/T = 0.001 for this N). Observe here that all non-drift and drift-removed $K_{stat}(f)$ are equal for this τ, as is expected. Figure 12 shows these $K_{stat}(f)$ for τ/T = 0.1. Observe here that the 6th order polynomial drift-removed $K_{stat}(f)$ already have significant deviations at low f, while the other drift-removed $K_{stat}(f)$ show minimal deviations at low f from those of the non-drift-removed $K_{stat}(f)$. This shows why $K_{stat}(f)$-bias for p = -2 and -3 noise with 6th order polynomial drift removal starts to increase at lower τ/T values than that for the other drift-removal methods. These early deviations for 6th order polynomial drift-removal are due to its sharper data-residual highpass filtering and higher transition frequency as compared to lower-order polynomial drift removal (see Figure 2).

Figure 13 shows the $K_{stat}(f)$ for τ/T at their maximum possible values for each statistic. Observe here that all drift-removed $K_{stat}(f)$ show significant deviations at low f. This explains the significant bias at these τ/T for all drift-removal methods. Finally, note that low-f deviations in Greenhall and 2nd order polynomial drift-removed $K_{stat}(f)$ are significantly less for the total statistic than those for the other statistics. This explains the lower drift-removed bias for p = -2 and -3 noise in the total statistic.

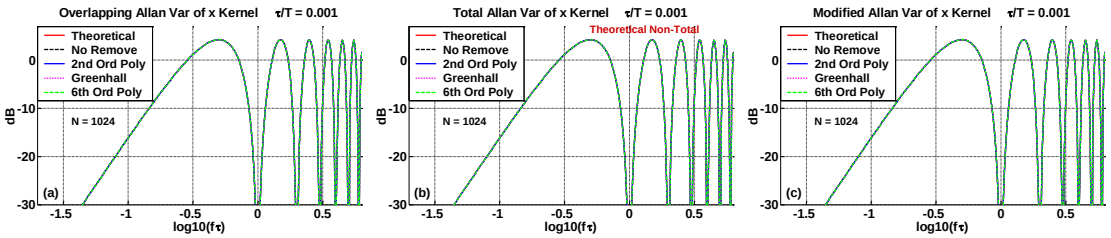

Figure 11. Drift-removed Mod, total, and overlapping Allan kernels for τ/T =0.001 (1-sample spacing).

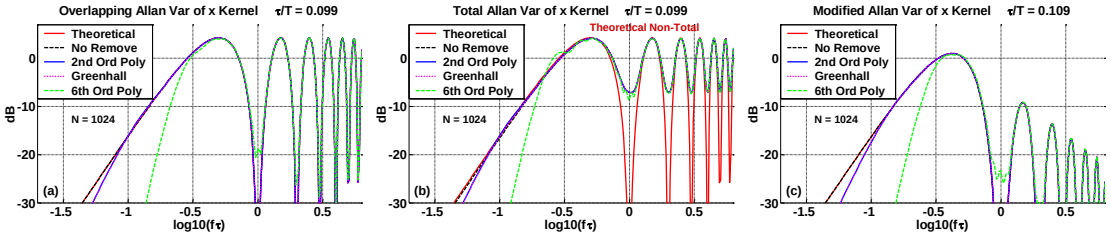

Figure 12. Drift-removed Mod, total, and overlapping Allan kernels for τ/T = 0.1.

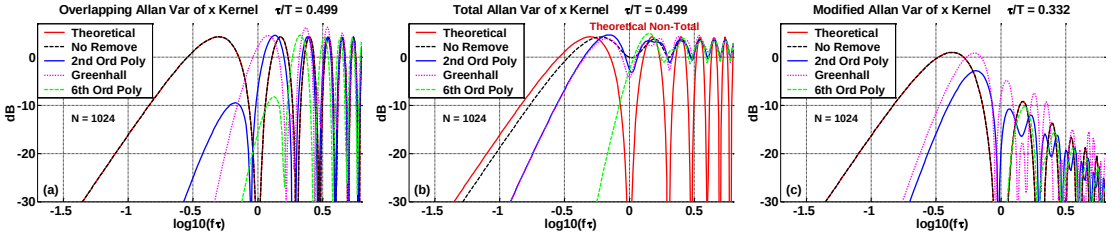

Figure 13. Drift-removed Mod, total, and overlapping Allan kernels for maximum τ/T.

HADAMARD DEVIATE BIAS

Figure 14(a) shows 6th order polynomial drift-removed $K_{stat}(f)$-bias for the overlapping ${}_H\sigma_y(\tau)$ (or ${}_H\sigma_x(\tau)$) versus τ/T, and Figure 14(b) shows $K_{stat}(f)$ for this statistic at its maximum τ/T for various drift-removal methods and no drift removal. Note in Figure 14(b) that 2nd order drift Greenhall and 2nd order polynomial ULSF drift-removed $K_{stat}(f)$ are identical to the non-drift-removed $K_{stat}(f)$ as expected, because the Hadamard deviate is frequency-drift insensitive **[5-6]**. Finally in Figure 14(a), note that there is significant 6th order polynomial drift-removal bias for p = -2 and -3 noise that starts at relatively low values of τ/T, similar to that in Allan statistics. This also implies that temporally complex drift removal or noise whitening can cause significant bias in Hadamard deviates for appreciable values of τ/T as well as Allan deviates.

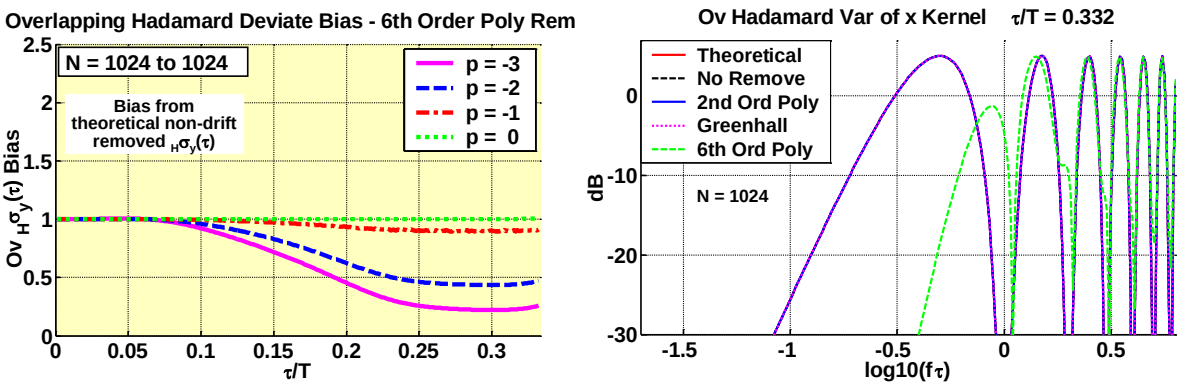

Figure 14. The 6th order poly drift-removed overlapping Hadamard bias vs τ/T and kernels for maximum τ/T.

CONCLUSIONS

In this paper, we have investigated the nature of drift-removal $K_{stat}(f)$-bias in various AHV statistics. In doing this, we have introduced a computationally efficient technique for generating drift-removed AHV $K_{stat}(f)$ for p = -2 and -3 $S_x(f)$. In the paper, we have shown that

- the Greenhall drift-removed total Allan deviate has significantly lower bias as a measure of M^{th} order random stability than that for the overlapping and modified Allan deviates, and
- the total Allan deviate has significantly less bias at appreciable τ/T than that of these other Allan statistics for all drift removal methods.

We have also shown that

- at appreciable τ/T, white (p = 0) noise bias for Mod $\sigma_y(\tau)$ can vary widely with detailed N values at large τ, and
- higher-order polynomial drift removal generates higher $K_{stat}(f)$-bias for p = -2 and -3 $S_x(f)$ than lower-order drift removal.

A consequence of the above bullet is that

- noise whitening typically will generate a large $K_{stat}(f)$-bias when p = -2 and -3 noise is present, and that
- noise whitening cannot be reliably used to remove unknown deterministic drift when such noise is present.

We have finally shown that

- environmental drift removal also can cause significant $K_{stat}(f)$-bias when p = -2 and -3 noise is present.

ACKNOWLEDGEMENT

The author would like to acknowledge the help of Leo A. Mallette in preparing this paper.

REFERENCES

[1] *"Standard Definitions of Physical Quantities for Fundamental Frequency and Time Metrology—Random Instabilities,"* 2008, IEEE Standard 1139-2008, IEEE.

[2] A. Papoulis and S. U. Pillai, 2002, **Probability, Random Variables, and Stochastic Processes**, 4th ed. (McGraw-Hill, Boston, Massachusetts).

[3] V. S. Reinhardt, 2010, "*The Profound Impact of Negative Power Law Noise on the Estimation of Causal Behavior,"* **IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control**, **Vol. 57**, No. 1, 74-81.

[4] V. S. Reinhardt, 2009, *"Zero Mean Noise Processes that Do Not Appear to be Zero Mean,"* in Proc. Institute of Navigation International Technical Meeting 2009, 26-28 January 2009¸ Anaheim, California, pp. 384-390.

[5] V. S. Reinhardt, 2011, *"The Properties of Mth Order Differences and Their Relationship to Mth Order Random Stability,"* in Joint Meeting of the EFTF and the IEEE IFCS, 2-5 May 2011, San Francisco, California, pp. 155-160.

[6] W. J. Riley, 2007,*"The Hadamard Variance," http://www.ieee-uffc.org/frequency_control /teaching/ pdf/Hadamard.pdf*, download November 2007.

[7] J. J. Gagnepain, 1998, *"La Variance de B. Picinbono,"* **Traitement du Signal**, **Vol. 15**, #6, Special, 477-482.

[8] R. A. Baugh, 1971, *"Frequency Modulation Analysis with the Hadamard Variance,"* in Proc. 25th Annual Frequency Control Symposium (IEEE), 26-28 April 1971, Atlantic City, New Jersey, pp. 222–225.

[9] *"IEEE Guide for Measurement of Environmental Sensitivities of Standard Frequency Generators,"* 2003, IEEE Std 1193.

[10] S. M. Kay, 1993, **Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory** (Prentice Hall, Saddle River, New Jersey).

[11] C. A. Greenhall, 1997, *"A Frequency-Drift Estimator and Its Removal from Modified Allan Variance,"* in Proc. 1997 IEEE IFCS, 28-30 May 1997, Orlando, Florida, pp. 428-432.

[12] H. W. Sorenson, 1966, *"Kalman Filtering Techniques,"* **Advances in Control Systems (Vol. 3)**, C. T. Leondes (Ed.) (Academic Press, New York, New York).

[13] MIL-PRF-55310D, 1998, *"Performance Specification, Oscillator, Crystal Controlled, General Specification for,"* U.S. DoD, 25 March 1998.

[14] J. R. Vig, and T. R. Meeker, 1991, *"The Aging of Bulk Acoustic Wave Resonators, Filters and Oscillators,"* in Proc. 45th Annual Symp. Freq. Control, 29-31 May 1991, Los Angeles, California, pp. 77 - 101.

[15] V. S. Reinhardt, 2007, *"How extracting information from data highpass filters its additive noise,"* in Proc. 39th PTTI Systems and Applications Meeting, 26-29 November 2007, Long Beach, California, pp. 559-580.