Victor S. Reinhardt, "On Difference Variances as Residual Error Measures in Geolocation," Institute of Navigation 2008 National Technical Meeting, January 28-30, 2008, San Diego, CA, USA.

On Difference Variances as Residual Error Measures in Geolocation

Victor S. Reinhardt, *Raytheon Space and Airborne Systems*

Post publication corrections: Equation (A.11)

BIOGRAPHY

Dr. Reinhardt received his Ph. D. in Physics from Harvard University, where he worked under Professor Norman Ramsey. He is currently employed at Raytheon Space and Airborne Systems as an Engineering Fellow. Dr. Reinhardt has 23 patents and has authored numerous papers in a large variety of technical areas. Dr. Reinhardt's professional activities include being a member of the IEEE FCS Technical Program Committee, the PTTI Advisory Board, the IEEE Standards Committee for the Definitions of Physical Quantities for Fundamental Frequency and Time Metrology, and the ITU Study Group on Satellite Time and Frequency Transfer and Dissemination.

ABSTRACT

Two types of random error measures are utilized by the navigation community. Observable residual error (R) variances are used in geolocation to characterize the mean square (MS) error of data from a trajectory estimate generated from the data itself. On the other hand, M^{th} order difference (Δ) variances, such as Allan and Hadamard variances, are used in time and frequency (T&F) source characterization to determine the MS of the M^{th} order difference of data over a separation interval τ. Δ-variances are used in this capacity because they have excellent convergence properties for the negative power law (neg-p) noise present in T&F sources, while R-variances are not known for such properties. When neg-p noise is present in geolocation problems, it would be desirable to use Δ-variances as measures of residual error. This paper shows that Δ-variances can indeed be used in this capacity under certain conditions, principally when the model function used estimate the trajectory is an $(M-1)^{th}$ order polynomial. It is also shown that R-variances share the good convergence properties of Δ-variances when neg-p noise is present, even when the Δ-variances can't be used to represent residual error. It is further shown that the convergence properties of R-variances are a function of the complexity of the model function used. Therefore, R-variances are guaranteed to converge in the presence of neg-p noise if one is free to choose the model function.

INTRODUCTION

Two types of random error measures are utilized in the navigation community: Observable residual error (R) variances and M^{th} order difference (Δ) variances. R-variances are used in geolocation, and also other areas, to statistically characterize the residual error of measured data from a trajectory, or any other causal function, imbedded in the data $x(t_n)$, as is shown in Figure 1 [Wolberg, 1967]. Observable R-variances, which are the types we will discuss, are given by various mean square (MS) statistics of the residual error after removing an *estimate* of the causal trajectory from position, range, or time data $x(t_n)$ [Reinhardt, 2007-2]. These R-variances are often called standard variances or MS (RMS) errors [Wolberg, 1967], but we will use the neutral term R-variances to avoid confusion with conflicting definitions or connotations.

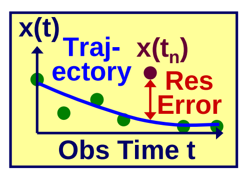

Figure 1. Residual error from an estimated trajectory.

Δ-variances [Reinhardt, 2007-1], such as Allan [Allan, 1966; Std 1139, 1999] and Hadamard [Baugh, 1971] or Picinbono [Gagnepain, 1998] variances, on the other hand, are used to characterize time and frequency (T&F) source error, because of their good convergence properties in the presence of the negative power law (neg-p) noise that these sources generate. Δ-variances are given by various MS statistics of $\Delta(\tau)^M x(t_n)$ the M^{th} power of the forward difference operator $\Delta(\tau)$ over an interval τ operating on the data $x(t_n)$. $\Delta(\tau)$ is defined by [Reinhardt, 2007-1; Reinhardt, 2007-2]

$$\Delta(\tau)x(t) \equiv x(t+\tau) - x(t) \quad (1)$$

Note that $\Delta(\tau)x(t)$ is just the time interval error (ITE) over τ when $x(t)$ is the time error [Std 1139, 1999]. The first two orders of $\Delta(\tau)^M x(t_n)$ are pictured in Figure 2.

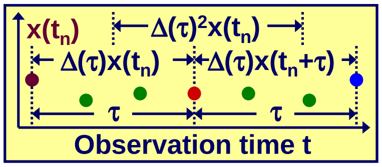

Figure 2. 1st and 2nd order differences over the interval τ.

By neg-p noise, we mean wide-sense stationary but divergent x(t) noise with a single sideband (SSB) power spectral density (PSD) given by [Std 1139, 1999; Reinhardt, 2007-1]

$$L_x(f) \propto f^p \quad (p < 0) \tag{2}$$

We note that $L_x(f)$ is defined for this paper as

$$L_x(f) = \Im_{f,\tau}\{R_x(\tau)\} \equiv \int_{-\infty}^{+\infty} d\tau e^{-j\omega\tau} R_x(\tau) \tag{3}$$

where $R_x(\tau)$ the autocorrelation function for a wide-sense stationary real process x(t) is given by [Scharf, 1998]

$$R_x(\tau) = E\{x(t+\tau/2)x(t-\tau/2)\} \tag{4}$$

$E\{..\}$ is the ensemble average, and $\omega = 2\pi f$. We also note that the double-sideband PSD $S_x(f)$ used in many T&F works [Std 1139, 1999], not the SSB PSD $L_x(f)$ used in this paper.

One can show that M^{th} order Δ-variances highpass (HP) filter $L_x(f)$ with a $2M^{th}$ order zero at f=0 [Reinhardt, 2007-1]. This guarantees that an M^{th} order Δ-variance will converge for poles in $L_x(f)$ up to order $-p = 2M$. Thus the Allan variance, which is proportional to the 2^{nd} order Δ-variance of x(t), is guaranteed to converge for $p \geq -4$ and the Hadamard variance, which is proportional to the 3^{rd} order Δ-variance of x(t), is guaranteed to converge for $p \geq -6$ [Reinhardt, 2007-1].

The common wisdom is that R-variances do not have good convergence properties in the presence of neg-p noise. However, such variances are the proper statistical measures of geolocation error regardless of any convergence difficulties, because they properly address the statistical questions being posed. Thus, it would be desirable if Δ-variances could be also used as measures of MS residual error in geolocation problems when neg-p noise is present. In this paper, we will demonstrate that this indeed can be done under certain conditions, where the most important condition is the use of an $(M-1)^{th}$ order polynomial as the model function to estimate the trajectory.

Because of this conditional equivalence between R and Δ variances, one would expect residual error variances to have convergence properties similar to those of their Δ variance counterparts. In the second part of this paper, we will demonstrate that this is indeed true and is true under conditions much more general than those required for the equivalence between R and Δ variances. We will demonstrate that the very process of fitting a model function to data necessarily HP filters the noise in the observable residual error under very general conditions. We will also demonstrate that the order of this HP filtering is a function of the complexity of the model function used to estimate the trajectory. Therefore, R-variances are guaranteed to converge in the presence of any order of neg-p noise, if one is free to choose the model function used for the estimation process.

RESIDUAL ERRORS IN SIMPLIFIED GEOLOCATION

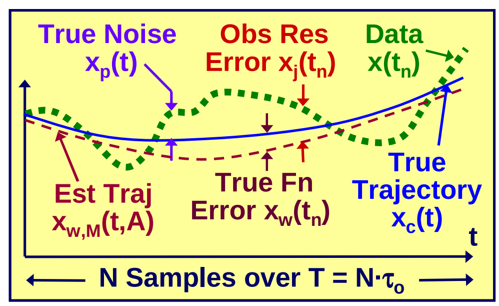

Figure 3. Simplified geolocation problem.

Figure 3 shows the simplified (post hoc) geolocation problem we will use for our discussion. Here we start with N data samples $x(t_n)$ distributed uniformly over a total observation time

$$T = N\tau_o \tag{5}$$

where τ_o is the sampling interval. We will assume the data is given by

$$x(t_n) = x_c(t_n) + x_p(t_n) \tag{6}$$

where $x_c(t_n)$ is the true trajectory that we wish to estimate and $x_p(t_n)$ is the true noise.

We will further assume that the continuous process $x(t)$ corresponding to the data samples $x(t_n)$ (and thus $x_c(t)$

and $x_p(t)$) has been pre-filtered by a system response function $H_s(f)$ [Reinhardt, 2006]. $H_s(f)$ is given by $\Im_{f,t}\{h_c(t)\}$ such that

$$x(t) = \int_{-\infty}^{+\infty} dt\, h_s(t-t')x_{in}(t') \tag{7}$$

where $x_{in}(t)$ is the variable prior to system filtering. More will be said about $H_s(f)$ later.

To solve our simplified geolocation problem, we want to estimate $x_c(t_n)$ by fitting a model function $x_{w,M}(t,\mathbf{A})$ containing M adjustable parameters $\mathbf{A} = (a_0, a_1, ...a_{M-1})$ to the N data samples $x(t_n)$. For this paper, the fitting process, in general, is not specified, Thus, it can be a uniformly of non-uniformly weighted least squares fit (LSQF) [Wolberg, 1967], a Kalman filter [Sorenson, 1966; Brown, 1983], etc. Also, some of the important results in this paper will be independent of the specific fitting process chosen.

The observable residual error from such a fit is thus

$$x_j(t_n) = x(t_n) - x_{w,M}(t_n,\mathbf{A}) \tag{8}$$

Two statistical error measures that can be formed from $x_j(t_n)$ are as follows. The most detailed measure is $E\{x_j(t_n)^2\}$ the ensemble MS residual error at the point $x(t_n)$, which we will call its point variance. This is strictly not an observable error measure, since one can only determine a single $x_j(t_n)^2$ from a single data set, but one can determine the properties of $E\{x_j(t_n)^2\}$ from knowledge of $L_x(f)$ and statistically apply it to direct observations. A less detailed measure is average variance of $x_j(t_n)$ over the data given by

$$\sigma_{x-j}^2 = \sum_{n=0}^{N-1} \xi_n E\{x_j(t_n)^2\} \tag{9}$$

where the weights ξ_n are generally determined by the weighting used in the fitting process. Here again, the ensemble average E{…} is used for the purposes of later analysis, but the observable statistic would be the above without the E{…}.

Another type of error that is not observable, but is the ultimate error one wishes to determine, is the true function error is given by

$$x_w(t_n) = x_{w,M}(t_n,\mathbf{A}) - x_c(t_n) \tag{10}$$

The statistical measures formed from $x_w(t_n)$ that we will discuss are its point variance $E\{x_w(t_n)^2\}$ and its average variance

$$\sigma_{x-w}^2 = E\{\sum_{n=0}^{N-1} \xi_n x_w(t_n)^2\} \tag{11}$$

We again note that $E\{x_w(t_n)^2\}$ and σ_{x-w}^2 are the true but non-observable measures of the accuracy of the fit.

Δ-VARIANCES AS MEASURES OF RESIDUAL ERROR

For this paper, we define the Δ variance of $x(t_n)$ as [Reinhardt, 2007-1]

$$\sigma_{x,M}^2(\tau) = \lambda_M^{-1}(N-M)^{-1} \sum_{n=0}^{N-M-1} E\{[\Delta(\tau)^M x(t_n)]^2\} \tag{12}$$

where the normalization

$$\lambda_M = \sum_{m=0}^{M} \left(\frac{M!}{m!(M-m)!} \right)^2 \tag{13}$$

is designed to make all M-orders of $\sigma_{x,M}^2(\tau)$ equal for uncorrelated white ($p = 0$) noise. The arithmetic average in (12) is called an overlapping average [Std 1139, 1999] and is only one way to form statistics of $[\Delta(\tau)^M x(t_n)]^2$. Other arithmetical averaging techniques, which will not be addressed here, are total or modified averages [Std 1139, 1999].

For $N = M+1$ data points, one can prove that [Reinhardt, 2007-2]

$$\sigma_{x-j}^2(N = M+1, M) = \sigma_{x,M}^2(T/M)\ [\xi_n^{-1} = 1] \tag{14}$$

under the conditions listed in Table 1.

Table 1. Conditions for equivalence of $\sigma^2_{x,M}(T/M)$ and σ^2_{x-j}.

(a) A uniformly weighted least squares fit (LSQF) over T is used to determine $\mathbf{A}$ in $x_{w,M}(t,\mathbf{A})$.

(b) $x_{w,M}(t,\mathbf{A})$ is an $(M-1)^{th}$ order polynomial with M coefficients $\mathbf{A}$.

(c) σ^2_{x-j} uses the "unbiased" uniform weighting coefficients $\xi_n^{-1} = N - M$. Note that $\xi_n = 1$ for $N = M+1$ and that "unbiased" is in parentheses because this statistic is only unbiased for uncorrelated white noise (p=0).

For M=2, (14) is just the well-known fact that the Allan variance of $x(t)$ is proportional to the uniformly weighted 3-sample R-variance of $x(t_n)$ when LSQF estimates of the time and frequency offsets are removed from the data [Allan, 1966]. Similarly, for M=3, (14) states that the Hadamard variance of x(t) is proportional to the 4-sample uniformly weighted residual variance of $x(t_n)$ when LSQF estimates of the time and frequency offsets plus the frequency drift are removed from the data.

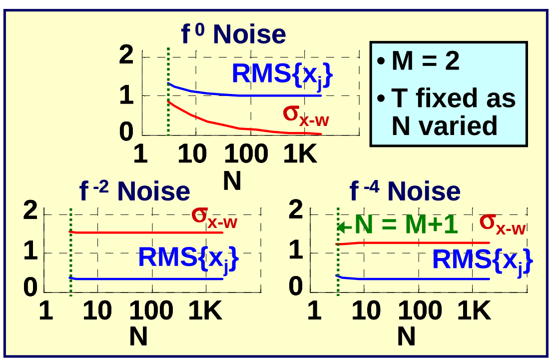

Figure 4. Errors in LSQF residuals as N is varied.

As shown in Figure 4, one can extend the relationship between σ^2_{x-j} and $\sigma^2_{x,M}(T/M)$ to any N by noting that the biased uniformly weighted form of σ_{x-j} ($\xi_n^{-1} = N$), which we'll designate $RMS\{x_j\}$ to avoid confusion with the "unbiased" ($\xi_n^{-1} = N - M$) σ^2_{x-j}, doesn't vary much with N. One can see from the figure that this is especially true for neg-p noise. We note here that N is being varied, but T is not. Setting this variation to zero, one can write the following approximate equality for any N

$$\sigma^2_{x-j}(N) = \frac{N}{N-M}\sigma^2_{x,M}(T/M)\,[\xi_n^{-1} = N-M] \qquad (15)$$

We note that one can obtain exact expressions similar to (15) for any specific p by deriving M^{th} order "bias" function relationships in a fashion similar to those derived by Allan [Allan, 1966] and Barnes [Barnes, 1971] for the Allan variance. One should note, however, that our bias functions have inherently different behavior from those of the Allan and Barnes bias functions because we fix both T and τ ($T = M\tau$) as N is varied, while Allan and Barnes fix τ as N is varied and let the total observation interval $N\tau$ change with N. (Allan and Barnes define T as the time between successive samples, not the total observation interval.) Thus, our bias functions are very close to $N/(N-M)$ for all p, unlike the Allan-Barnes bias functions, which have very different forms for different p.

From the above, one can see that $\sigma^2_{x,M}(T/M)$ can be interpreted as a measure of the MS residual error $\sigma^2_{x-j}(N,M)$ for any N when the conditions of Table 1 are met. This has several important consequences as follows:

(a) $\sigma^2_{x,M}(T/M)$ and $\sigma^2_{x-j}(N,M)$ are equivalent error measures when Table 1 conditions are met, and $\sigma^2_{x,M}(T/M)$ can be used for $\sigma^2_{x-j}(N,M)$ in geolocation problems under these conditions.

(b) When a non-uniformly weighted fit over T is used to generate $x_{w,M}(t,\mathbf{A})$, $\sigma^2_{x,M}(T_{eff}/M)$ should be a good approximate measure of $\sigma^2_{x-j}(N,M)$. (This has not been verified quantitatively.) Here, T_{eff} is a suitably defined correlation length for the non-uniform fit weighting over T. This approximation should also hold for Kalman filters, since a Kalman filter is equivalent to a LSQF under certain circumstances [Sorenson, 1966] and is also a least squares estimator [Brown, 1983].

(c) The advantage in using $\sigma^2_{x,M}(T_{eff}/M)$ to estimate $\sigma^2_{x-j}(N,M)$ in geolocation problems is that one need not perform the fit or remove $x_{w,M}(t,\mathbf{A})$ from the raw data for $\sigma^2_{x,M}(T_{eff}/M)$. This is due to the well-known insensitivity of $\sigma^2_{x,M}(\tau)$ to $(M-1)^{th}$ or lower order polynomial behavior [Reinhardt, 2007-1]. This holds whether $x_{w,M}(t,\mathbf{A})$ is a good model function for $x_c(t)$ or not, because the residual error due to poor modeling of $x_c(t)$ by $x_{w,M}(t,\mathbf{A})$ will effect both $\sigma^2_{x,M}(T_{eff}/M)$ and $\sigma^2_{x,M}(\tau)$ equally.

(d) The equivalence of $\sigma^2_{x,M}(T/M)$ and $\sigma^2_{x-j}(N,M)$ explains the well-known insensitivity of the Hadamard variance to frequency drift (2nd order polynomial coefficient of the time error) and the sensitivity of the Allan variance to such drift (because the equivalent $\sigma^2_{x-j}(N,2)$ doesn't model frequency drift).

THE HIGHPASS FILTERING OF NOISE IN THE OBSERVABLE RESIDUAL ERROR

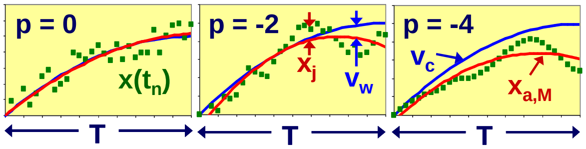

Figure 5. Simulated fit solutions for various p values.

In this section, we will demonstrate that the estimation of $x_c(t)$ from data using the model function $x_{w,M}(t,\mathbf{A})$ HP filters the noise in the residual error under very general conditions. We will first demonstrate this graphically and then formally prove the assertion. For the graphical discussion, consider Figure 5, where we show fit solutions for various power law noise indices p (and a large N). For $p = 0$ (white noise), one can see that the solution behaves in the classical manner [Wolberg, 1967], with $x_{w,M}(t,\mathbf{A})$ closely tracking $x_c(t)$. Using fitting theory for uncorrelated noise [Wolberg, 1967], one can show that $x_{w,M}(t,\mathbf{A}) \to x_c(t)$ ($\sigma_{x-w} \to 0$) as $N \to \infty$ so long as the bandwidth of the data is such that the data samples remain uncorrelated for all values of N.

For $p = -2$ and -4 (neg-p noise), however, one can see from the figure that there are significant systematic long-term deviations in $x_{w,M}(t,\mathbf{A})$ from $x_c(t)$. These are due to the highly correlated nature of neg-p noise. In fact, using fitting theory for correlated noise, one can show that these deviations will remain non-zero as $N \to \infty$, because the fit cannot distinguish highly correlated low frequency (LF) noise components from the causal behavior $x_c(t)$. We will later show that the transition Fourier frequency f_T at which this tracking occurs is approximately $1/T$ for a uniformly weighted fit over T and approximately $1/T_{eff}$ for a non-uniformly weighted fit over T, where T_{eff} is a suitably defined correlation length. Again, we note that T is fixed as N is varied for this discussion.

Thus, $x_{w,M}(t,\mathbf{A})$ tracks the LF components in the noise with $f \le f_T$, and this causes $L_x(f)$ to be HP filtered in $E\{x_j(t)^2\}$ and σ^2_{x-j}, with a highpass (HP) cut-off knee at f_T. One should point out that this LF tracking also occurs for white noise but is only apparent for neg-p noise. The effect stands out for neg-p noise because virtually all the power for neg-p noise is in Fourier components less than f_T (for any value of T), while for uncorrelated white noise, the power in the noise components below f_T is small for large T.

When the fitting process is linear in the data, one can write spectral representations for $E\{x_j(t_n)^2\}$ and σ^2_{x-j} as (see Appendix A)

$$E\{|x_j(t)|^2\} = \int_{-\infty}^{+\infty} df\, |G_j(t,f) H_s(f)|^2 L_x(f) + E\{|x_{j,c}(t)|^2\} \tag{16}$$

$$\sigma^2_{x\text{-}j} = 2\int_0^\infty df\, |H_s(f)|^2 K_{x\text{-}j}(f) L_x(f) + \sigma^2_{x\text{-}c} \tag{17}$$

First, we note the presence of $|H_s(f)|^2$ in (16) and (17). $H_s(f)$ represents the effects of filtering by a system on x(t) as given by (7) [Reinhardt, 2006]. $H_s(f)$ is generated both by actual filters in the system and by topological structures such as phase lock loops [Reinhardt, 2006]. $|H_s(f)|^2$ here replaces the simple LP cut-off f_h used in previous formulations of the spectral integral [Std 1139, 1999] and is a more accurate representation of a system's specific filtering properties than f_h.

The importance of using $H_s(f)$ in (16) and (17) is that it can be shown to have HP as well as LP behavior for many types of systems [Reinhardt, 2006]. This HP filtering helps both $E\{x_j(t)^2\}$ and σ^2_{x-j} to converge in the presence of neg-p noise and by itself can ensure the convergence of these error measures when certain values of neg-p noise are present [Reinhardt, 2006].

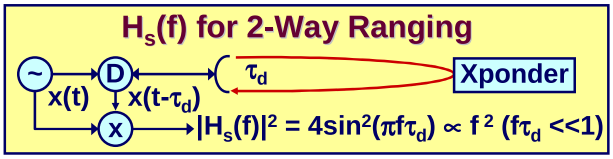

Figure 6. Delay system response function.

A classic example of such an $H_s(f)$ is that generate by the 2-way ranging topology shown in Figure 6 [Reinhardt, 2006]. This well-known response function arises because the ranging signal is mixed with a delayed version of itself. One can show that $|H_s(f)|^2 = 4\sin^2(\pi f \tau_d)$ for such a system [Reinhardt, 2006], which has a 2nd order zero at f=0. Thus this $H_s(f)$ suppresses poles in $L_x(f)$ at f=0 up to order $-p = 2$.

$G_j(t,f)$ in (16) comes from the fitting process used to determine $x_{w,M}(t,\mathbf{A})$. To explain how this $G_j(t,f)$ is generated, consider the following. (For more detail, see Appendix A and [Reinhardt, 2007-2].) When the trajectory fitting process is linear in the data, the solution can be represented in terms of a Green's function $g_w(t,t')$ as

$$x_{w,M}(t,\mathbf{A}) = \int_{-\infty}^{+\infty} dt' g_w(t,t') x(t') \qquad (18)$$

where $x(t)$ is the continuous process that equals $x(t_n)$ at $t = t_n$. Here, we've written $g_w(t,t')$ as a continuous function of t' so that we can write (16) in terms of the continuous PSD $L_x(f)$. We note that this $g_w(t,t')$ can be written in terms of the discrete sample solution Green's function $g_{w,N}(t,t_n)$ as

$$g_w(t,t') = \sum_{n=0}^{N-1} g_{w,N}(t,t_n)\delta(t-t_n) \qquad (19)$$

Using the properties of the Fourier transform, one can show that (18) can also be written as

$$x_{w,M}(t,\mathbf{A}) = \int_{-\infty}^{+\infty} df\, G_w(t,f) X(f) H_s(f) \qquad (20)$$

where $X(f) = \Im_{f,t}\{x_{in}(t)\}$, $x_{in}(t)$ is the data process before filtering by the system response function $H_s(f)$ and where $G_w(t,f)$ is the Wigner-Ville spectrum (for –f) [Scharf, 1998] of $g_w(t,t')$ given by $G_w(t,f) = \Im_{-f,t'}\{g_w(t,t')\}$.

$G_j(t,f)$ in (16) is then given in terms of $G_w(t,f)$ by (See Appendix A)

$$G_j(t,f) = e^{j\omega t} - G_w(t,f) \qquad (21)$$

From $|G_j(t,f)|^2$, one can also form $K_{x-j}(f)$ in (17) by averaging $|G_j(t,f)|^2$ over t (see Appendix A).

Let us now prove that $|G_j(t,f)|^2$ (and thus $K_{x-j}(f)$) has HP filtering properties for the general conditions given in Table 2.

Table 2. Conditions for HP filtering behavior in observable residual errors

(a) The fit solution is linear in the data $x(t_n)$.

(b) $x_j(t_n) = 0$ for the fit solution when there is no noise and $x_c(t) = x_{w,M}(t,\mathbf{A})$. (We note that the fit can be non-uniformly or uniformly weighted over T and can be a Kalman filter, LSQF, etc.)

To demonstrate the HP filtering of $|G_j(t,f)|^2$ under these conditions, let us first write $x(t_n)$ as

$$x(t_n) = \sum_f x_f(t_n) + x_c(t_n) \qquad (22)$$

where we have decomposed $x_p(t_n)$ into single-frequency noise components $x_f(t) = X_f e^{j2\pi ft}$. Because of the linearity of the fit solution, we can write the total solution $x_{w,M}(t,\mathbf{A})$ as

$$x_{w,M}(t,\mathbf{A}) = \sum_f x_{w,M}(t,\mathbf{A}^{(f)}) + x_{w,M}(t,\mathbf{A}^{(c)}) \qquad (23)$$

where $x_{w,M}(t,\mathbf{A}^{(f)})$ is the fit solution when $x_f(t_n)$ is the sole input data and $x_{w,M}(t,\mathbf{A}^{(c)})$ is the fit solution when $x_c(t_n)$ is the sole input data. Because the spectral noise components $x_f(t)$ for a wide-sense stationary noise process are uncorrelated with each other [Davenport, 1987] and with $x_c(t_n)$, from (23), we can write

$$E\{|x_j(t_n)|^2\} = \sum_f E\{|x_{j-f}(t_n)|^2\} + E\{|x_{j-c}(t_n)|^2\} \qquad (24)$$

where

$$\begin{aligned} x_{j-f}(t_n) &= x_f(t_n) - x_{w,M}(t_n,\mathbf{A}^{(f)}) \\ x_{j-c}(t_n) &= x_c(t_n) - x_{w,M}(t_n,\mathbf{A}^{(c)}) \end{aligned} \qquad (25)$$

Thus, since (16) is just the infinitesimal limit of (24), the noise filtering properties of $|G_j(t,f)|^2$ in (16) can be determined by demonstrating the noise filtering properties of each $E\{x_{j-f}(t_n)^2\}$ term separately.

To do so, let us expand $X_f e^{j2\pi ft}$ using the well-known Taylor Theorem as

$$x_f(t) = X_f e^{j2\pi f t_0} \sum_{k=0}^{M-1} \frac{(j2\pi f (t-t_0))^k}{k!} + X_f \frac{(j2\pi f (t'-t_0))^M}{M!} \tag{26}$$

where t' is somewhere in $[t_0, t]$ and t_0 is the beginning of the observation interval. If $x_f(t)$ were given only by the right hand finite sum term in (26), the residual error $x_{j-f}(t_n)$, and hence $E\{x_j(t_n)^2\}$, would be zero when $x_{a,M}(t,\mathbf{A})$ is an (M-1)th order polynomial because of Table 2 property (b). Thus, when the Taylor series converges, the value of $E\{x_j(t_n)^2\}$ must be proportional to the square of the right most term in (26), which is proportional to f^{2M}. Therefore, we must have

$$|G_j(t,f)|^2 \propto f^{2M} \text{ (for } f \ll 1) \quad [x_{a,M}(t,\mathbf{A}) = (M-1)^{th} \text{ order polynomial}] \tag{27}$$

assuming the Taylor series converges for small enough f. Similarly, for arbitrary $x_{a,M}(t,\mathbf{A})$ with a DC term a_o, one must have

$$|G_j(t,f)|^2 \propto f^{2m} \;\; m \ge 1 \text{ (for } f \ll 1) \quad [x_{a,M}(t,\mathbf{A}) \text{ has DC Term}] \tag{28}$$

because this is covered by the M=1 case for (27). Finally, (27) and (28) also apply to $K_{x-j}(f)$ because it is formed by averaging $|G_j(t,f)|^2$ over t (see Appendix A).

Figure 7 shows a Matlab simulation of $K_{x-j}(f)$ for a polynomial $x_{w,M}(t,\mathbf{A})$ with $M = 1$ to 5, $N = 1000$, and a uniformly weighted LSQF. This verifies that $K_{x-j}(f)$ has the HP filtering properties given by (27) and demonstrates that the HP knee f_T is approximately at 1/T for a uniformly weighted fit. Figure 8 shows a similar simulation for a non-uniformly weighed LSQF. This also verifies that $K_{x-j}(f)$ has the HP filtering properties given by (27) and demonstrates that the HP knee f_T is approximately at $1/T_{eff}$ for a non-uniformly weighted fit. We finally note that (27) proves that $E\{x_j(t_n)^2\}$ and σ^2_{x-j} are guaranteed to converge for any neg-p value, if one is free to chose the form of $x_{w,M}(t,\mathbf{A})$, because one can choose a polynomial of any order.

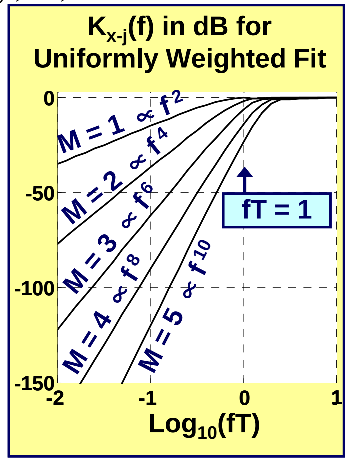

Figure 7. $K_{x-j}(f)$ in dB for (M-1)th order polynomial $x_{w,M}(t,\mathbf{A})$ and a uniformly weighted LSQF.

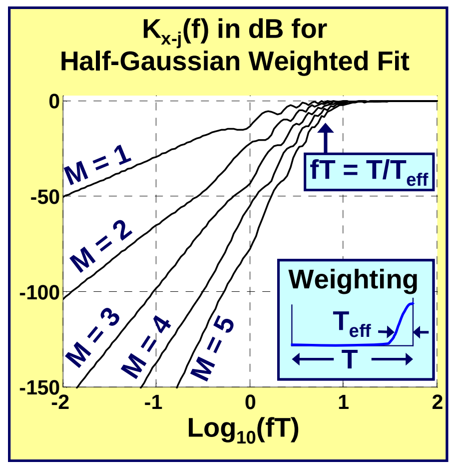

Figure 7. $K_{x-j}(f)$ in dB for (M-1)th order polynomial $x_{w,M}(t,\mathbf{A})$ and a non-uniformly weighted LSQF.

As a final topic, we note from (8) and (10) that we can write

$$x_p(t_n) = x_j(t_n) + x_w(t_n) \tag{29}$$

Thus, $E\{x_w(t_n)^2\}$ and σ^2_{x-w} must lowpass (LP) filter $L_x(f)$, because $x_w(t_n)$ is the complement of $x_j(t_n)$ with respect to the total noise $x_p(t_n)$. In Appendix A, spectral integrals for $E\{x_w(t_n)^2\}$, σ^2_{x-w}, $E\{x_{j-c}(t_n)^2\}$, and σ^2_{x-c} are also derived.

CONCLUSIONS

We have shown that $\sigma^2_{x,M}(T/M)$ is equivalent to σ^2_{x-j} when the conditions of Table 1 are met. Furthermore, if we relax the uniform weighting condition (a),

$\sigma^2_{x,M}(T_{eff}/M)$ should still be a good approximate measure of σ^2_{x-j}. Thus, $\sigma^2_{x,M}$ can be used as an exact or approximate measure of residual geolocation problems when the fit function is a polynomial.

Under the very general conditions of Table 2, we have also shown that the estimation of the trajectory in a geolocation problem causes $E\{x_j(t_n)^2\}$ and σ^2_{x-j} to have noise HP filtering properties that are a function of the complexity of the model function $x_{w,M}(t,\mathbf{A})$. This means that $E\{x_j(t_n)^2\}$ and σ^2_{x-j} are guaranteed to converge if one is free to choose the form of $x_{w,M}(t,\mathbf{A})$.

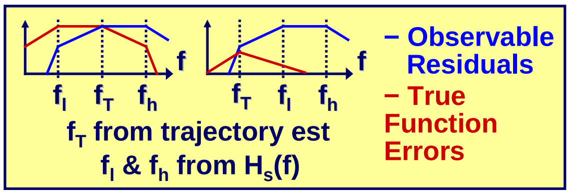

Figure 8. Noise filtering properties of the observable residual error and the true function error.

Figure 8 summarizes the noise filtering properties of the observable residual error measures $E\{x_j(t_n)^2\}$ and σ^2_{x-j} and the true function error measures $E\{x_w(t_n)^2\}$ and σ^2_{x-w}. As shown in the figure, the trajectory fitting process causes $E\{x_j(t_n)^2\}$ and σ^2_{x-j} to be HP filtered and $E\{x_w(t_n)^2\}$ and σ^2_{x-w} to be LP filtered with a knee frequency f_T. As is also shown in the figure, $H_s(f)$ filters all error measures equally, since $H_s(f)$ has the same effect on all variables. This $H_s(f)$ filtering is shown parametrically using an HP knee f_l and a LP knee f_h.

The right side of Figure 8 shows that $E\{x_w(t_n)^2\}$ and σ^2_{x-w} will disappear as $T \to \infty$ ($f_l >> f_T$), if the HP order of $H_s(f)$ is sufficient by itself to overcome the pole in $L_x(f)$ at f=0; that is, if $H_s(f)$ guarantees the convergence of $E\{x_w(t_n)^2\}$ and σ^2_{x-w}. This case is the transition to stationary but correlated statistics. Under these conditions $E\{x_w(t_n)^2\}$ and σ^2_{x-w} become measures of the true data errors when $f_l << f_T$. However, when $H_s(f)$ is not sufficient to make $E\{x_w(t_n)^2\}$ and σ^2_{x-w} converge, $E\{x_j(t_n)^2\}$ and σ^2_{x-j} will never be measures of the true error even though they themselves converge.

Applying the above to the GPS system, one confirms well-known wisdom about the behavior of error measures and gains new insights. First, the above confirms that $E\{x_j(t_n)^2\}$ and σ^2_{x-j} obtained from local receiver (Rx) data over small T are only measures of the consistency of the data, not its accuracy. Second, the convergence of these local error measures in the presence of neg-p noise is a function of the complexity of the local trajectory models used. Third, one can view the operation of the control segment as a phase lock loop $H_s(f)$ with an HP cut-off $f_l = 1/T_{GPS}$ that is sufficient to make $E\{x_w(t_n)^2\}$ and σ^2_{x-w} converge in the presence of neg-p noise. (That is, the system variables are internally convergent because they are tied to fixed ground sites, regardless of any divergences in the GPS time scale as a whole.) Because of this, $E\{x_j(t_n)^2\}$ and σ^2_{x-j} will be measures of the true errors for $T >> T_{GPS}$.

APPENDIX A. MODELING THE SPECTRAL RESPONSE OF TRAJECTORY ESTIMATION FROM DATA

When the trajectory fitting process is linear in the data $x(t_n)$, it was pointed out that, the solution $x_{w,M}(t,\mathbf{A})$ can be represented in terms of a Green's function $g_w(t,t')$ as given by (18). $x_{w,M}(t,\mathbf{A})$ was also written in terms of $G_w(t,f)$, $H_s(f)$, and $X(f) = \Im_{f,t}\{x_{in}(t)\}$, where $x_{in}(t)$ is given by (7), and $x_{w,M}(t,\mathbf{A})$ is given by (20). Because of the linearity of the solution, we can decompose $x_j(t_n)$ given by (8) and $x_w(t_n)$ given by (10) into $x_{j,p}(t_n)$, $x_{j,p}(t_n)$, and $x_{j,c}(t)$ as

$$\begin{aligned} x_j(t) &= x_{j,p}(t) - x_{j,c}(t) \\ x_w(t) &= x_{w,p}(t) + x_{j,c}(t) \end{aligned} \quad (A.1)$$

where (a)

$$\begin{aligned} x_{w,p}(t) &= x_{w,M}(t,\mathbf{A}^{(p)}) = \int_{-\infty}^{+\infty} df\, G_w(t,f) X_p(f) H_s(f) \\ x_{j,p}(t) &= x_p(t) - x_{w,M}(t,\mathbf{A}^{(p)}) \\ &= \int_{-\infty}^{+\infty} df\, G_j(t,f) X_p(f) H_s(f) \\ x_{j,c}(t) &= x_{w,M}(t,\mathbf{A}^{(c)}) - x_c(t) \\ &= -\int_{-\infty}^{+\infty} df\, G_j(t,f) X_c(f) H_s(f) \end{aligned} \quad (A.2)$$

(b) $x_{w,M}(t,\mathbf{A}^{(c)})$ is the solution when only the causal trajectory $x_c(t)$ is present,

(c) $x_{w,M}(t,\mathbf{A}^{(p)})$ is the solution when only the true noise $x_p(t)$ is present,

(d)

$$g_j(t,t') = \delta(t-t') - g_w(t,t')$$
$$G_j(t,f) = \Im_{-f,t'}\{g_j(t,t')\} = e^{j\omega t} - G_w(t,f) \qquad \text{(A.3)}$$

and (e)

$$x_{w,p}(t) + x_{j,p}(t) = x_w(t) + x_j(t) = x_P(t) \qquad \text{(A.4)}$$

By taking the ensemble average $E\{\ldots\}$ of the square of $x_j(t)$ and $x_w(t)$ and assuming $x_p(t)$ is uncorrelated with $x_c(t)$, one can write

$$E\{x_j(t)^2\} = E\{x_{j,p}(t)^2\} + E\{x_{j,c}(t)^2\}$$
$$E\{x_w(t)^2\} = E\{x_{w,p}(t)^2\} + E\{x_{j,c}(t)^2\} \qquad \text{(A.5)}$$

where

$$E\{x_{j,p}(t)^2\} = \int_{-\infty}^{+\infty} df\, |G_j(t,f)H_s(f)|^2\, L_x(f)$$
$$E\{x_{w,p}(t)^2\} = \int_{-\infty}^{+\infty} df\, |G_w(t,f)H_s(f)|^2\, L_x(f)$$
$$E\{x_{j,c}(t)^2 = \int_{-\infty}^{+\infty} df_g \int_{-\infty}^{+\infty} df\, L_c(f_g,f) \qquad \text{(A.6)}$$
$$\cdot G_j(t,f+0.5f_g)G_j^*(t,f-0.5f_g)$$
$$\cdot H_s(f+0.5f_g)H_s^*(f-0.5f_g)$$

with

$$L_x(f) = \Im_{f,\tau}\{R_p(\tau)\}$$
$$L_c(f_g,f) = \Im_{f_g,t_g}\{\Im_{f,\tau}\{R_c(t_g,\tau)\}\} \qquad \text{(A.7)}$$

$$R_p(\tau) = E\{x_p(t_g+\tau/2)x_p(t_g-\tau/2)\}$$
$$R_p(t_g,\tau) = E\{x_c(t_g+\tau/2)x_c(t_g-\tau/2)\} \qquad \text{(A.8)}$$

In the above, is $L_c(f_g,f)$ is the “rotated” Loève spectrum of $x_c(t)$ given by the double Fourier transform of the double time autocorrelation function $R_c(t_g,\tau)$, written in terms of the “rotationally” transformed global time $t_g = 0.5(t_1+t_2)$ and local or differential time $\tau = (t_1 - t_1)$ [Scharf, 1998; Cohen, 1995].

If we now integrate (A.6) over t, we obtain

$$\sigma^2_{x-j} = \sigma^2_{x-j,p} + \sigma^2_{x-c}$$
$$\sigma^2_{x-w} = \sigma^2_{x-w,p} + \sigma^2_{x-c} \qquad \text{(A.9)}$$

with

$$\sigma^2_{x-j,p} = \int_{-\infty}^{+\infty} df\, K_{x-j}(f)\,|H_s(f)|^2\, L_x(f)$$
$$\sigma^2_{x-w,p} = \int_{-\infty}^{+\infty} df\, K_{x-w}(f)|H_s(f)|^2\, L_x(f)$$
$$\sigma^2_{x-c} = \int_{-\infty}^{+\infty} df_g \int_{-\infty}^{+\infty} df\, K_{x-c}(f,f_g)L_c(f_g,f) \qquad \text{(A.10)}$$
$$H_s(f+0.5f_g)H_s^*(f-0.5f_g)$$

$$K_{x-j}(f) = \sum_{n=0}^{N-1} \xi_n\, |G_j(t_n,f)|^2$$
$$K_{x-w}(f) = \sum_{n=0}^{N-1} \xi_n\, |G_w(t_n,f)|^2 \qquad \text{(A.11)}$$
$$K_{x-c}(f,f_g) = \sum_{n=0}^{N-1} \xi_n G_j(t_n,f+0.5f_g)G_j^*(t_n,f-0.5f_g)$$

REFERENCES

[Allan, 1966] D. W. Allan, “Statistics of Atomic Frequency Standards,” Proceedings of the IEEE, v54, #2, Feb. 1966, pp 221- 230.

[Barnes, 1971] J. A. Barnes, 1971, “Characterization of Frequency Stability,” IEEE Trans. IM-20, May, 1971, pp 105-120.

[Baugh, 1971] Baugh, R.A., “Frequency Modulation Analysis with the Hadamard Variance,” 25th Annual Frequency Control Symposium, 1971, pp. 222–225.

[Brown, 1983] R. G. Brown, **Introduction to Random Signal analysis and Kalman Filtering**, John Wiley and ssons, 1983.

[Cohen, 1995] L. Cohen, **Time-Frequency Analysis**, Prentice-Hall, 1995.

[Davenport, 1987] Wilbur B. Jr. Davenport and William L. Root, **An Introduction to the Theory of Random Signals and Noise**, IEEE Press, 1987.

[Gagnepain, 1998] J. J. Gagnepain, "La Variance de B. Picinbono," Traitement du Signal, v15, #6, Special, pp. 477-482, 1998.

[Reinhardt, 2006] V. S. Reinhardt, “The Properties of Time and Phase Variances in the Presence of Power Law Noise for Various Systems,” 2006 IEEE International Frequency Control Symposium, June, 2006, pp745-749.

[Reinhardt, 2007-1] V. S. Reinhardt, "A Physical Interpretation of Difference Variances," 2007 Joint Meeting of the European Time and Frequency Forum (EFTF) and the IEEE International Frequency Control Symposium (IEEE-FCS), Geneva, Switzerland, May 29 - June 1, 2007.

[Reinhardt, 2007-2] V, S. Reinhardt, "How Extracting Information From Data Highpass Filters Its Additive Noise," 39th PTTI Systems and Applications Meeting, Nov 26 - 29, 2007.

[Scharf, 1998] L. L. Scharf, B. Firedlander, and D. J. Thomson, 1998, "Covavariant Estimators of Time-Frequency Descriptors for Nonstationary Random Processes," 32nd Asilomar Conference on Signals, Systems, and Computers, v1, Pacific Grove, CA, 1998, pp 808-811.

[Sorenson, 1966] H. W. Sorenson, "Kalman Filtering Techniques," in **Advances in Control systems** (Vol. 3) C. T. Leondes (Ed.), Acedemic Press, 1966.

[Std 1139, 1999] IEEE Standard 1139, "Standard Definitions of Physical Quantities for Fundamental Frequency and Time Metrology—Random Instabilities," IEEE, 1999.

[Wolberg, 1967] Wolberg, John R., **Prediction Analysis**, D. Van Nostrand and Co, 1967.