# Optimal oscillator modelling for GNSS-disciplined clocks on holdover

D.N. Matsakis and M. Slavney, *Masterclock, Inc.*

**Abstract**

An important part of GNNS Disciplined Clock (GDC) design is to optimally estimate the oscillator's current frequency and drift, as that will determine the performance of the GDC under holdover. The simplest method is make a parabolic fit to a batch of the most recent GNSS data used by the GDC, after removing the effects of all steering that had been applied. The steering would then be applied to the derived parabola (phase, frequency, and drift) so as to extrapolate into the future. The accuracy of this extrapolation determines the capability of the GDC under holdover. It is dependent on the noise characteristics of the oscillator.

An insightful paper by Vernotte et al. (2001) shows how the time interval error (TIE) of an oscillator depends on the amount of its white phase noise, white frequency noise (random walk phase noise, RW), flicker frequency noise, and the amount of its random walk frequency noise (integrated random walk phase noise, IRW, or RR). This paper uses those results to derive the optimal baseline for any linear combination of RR and IRW, mostly under the assumption that the phase noise is negligible. For brevity, flicker frequency noise is ignored; its characteristics are intermediate between the two noise types covered. The rule of thumb for the parabolic fit is that the optimal baseline is about ten times the prediction distance for pure RW, but only 1.06 times the prediction distance for IRW. The optimal baseline distance of a linear combination more closely approximates that of RWFM as the fourth power of the prediction distance distances. Some analysis for white phase noise is also included.

Finally, it is shown that a perfectly tuned Kalman filter outperforms even the optimal quadratic fits, and the relative accuracy of the two algorithms is assessed.

**Biographies**

Dr. Demetrios Matsakis has been Chief Scientist for Masterclock, Inc. since his 2019 retirement from the U.S. Naval Observatory. In his 40 years at the USNO he worked on most aspects of precise timekeeping, served 16 years a director the Time Service department, served three years as President of the International Astronomy Union's time commission, represented the United States in Geneva as Vice-president of the ITU WP7a delegation, and chaired or served on numerous working groups. He is an editor for GPS Solutions, has published approximately 170 scientific papers, secured one patent, and coauthored *An Introduction to Modern Timekeeping and Time Transfer*. His definition of time can be found on the Encyclopedia Britannica's web pages. He was won the Newcomb Award, which is the highest internal award for science at the USNO, and the Institute of Navigation's PTTI Distinguished Service Award.

Matthew Slavney has been an engineer at Masterclock Inc. for five years, where he has worked on time and frequency calibration as well as the development and programming of precision clocks and time references. He graduated from Miami University with a bachelors in computer engineering.

## I. Introduction

In Vernotte et al. (2001) it was shown how accurately predictions can be made $t_p$ points into the future by least-squares equal-weighting fitting over $t_m$ points (Figure 1.1) for time series characterized by White Phase, Random Walk PM (RW), Flicker FM, and Integrated Random walk PM (IRW, also termed Random Run, RR). This work will consider only White Phase, RW and IRW noise. Phase noise is considered negligible on the holdover times of interest, although its optimal $T_m$ is infinite; a simple exposition of this is provided in Appendix 1. In Appendix II a linear model will be briefly described. While flicker noises are not considered in this work, two references on how

 https://doi.org/10.33012/2025.19958

the Kalman formalism can incorporate them are Van Dierendonck at al. (1984) and Brown and Hwang (1997), which corrects a mistake in the former work.

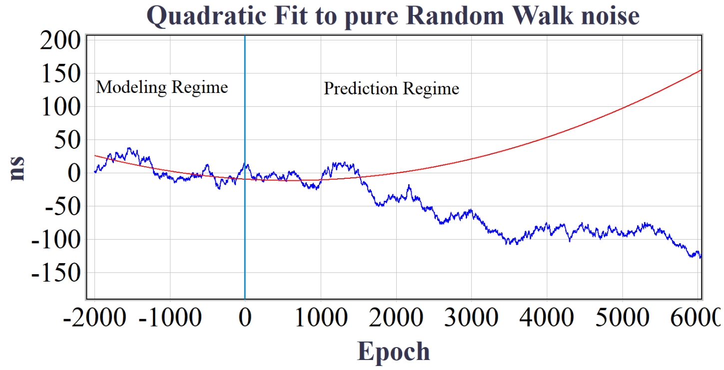



**Figure 1.1** A quadratic fit to the points in the modelling regime is extrapolated to predict the future

We shall consider the effects of the various types of noise separately, and in combination. Here are some of the standard equations and notations that will be employed for the Power Spectral Densities in terms of the fundamental noise types [1 and 2]:

$$S_x(f) = \sum_{a=-4}^{0} k_\alpha f^\alpha \tag{1.1}$$

$$S_y(f) = \sum_{a=-2}^{2} h_\alpha f^\alpha = \sum_{a=-4}^{0} 4\pi^2 k_\alpha f^{\alpha+2}, \text{ so } h_\alpha = 4\pi^2 k_{\alpha-2} \tag{1.2}$$

**II. White PM noise (WPM)**

The next equation is equation 52 in Vernotte et al. (2001),

$$TIE_W^2(T_m, T_p) = \frac{k_0 f_h}{N}\left(N + 180\frac{T_p^4}{T_m^4} + 360\frac{T_p^3}{T_m^3} + 252\frac{T_p^2}{T_m^2} + 72\frac{T_p}{T_m} + 9\right)$$ (2.1, equation 52 in Vernotte et al. (2001))

Here $TIE_W^2$ is the contribution to the TIE variance from white phase noise and N is related to the modeling time span $T_m$ by N=$T_m/\tau_0$. The high frequency cutoff $f_h$ (= $1/\tau_0$) is used in equation 45 of Vernotte et al. (2001) and 108 of Barnes et al. (1971) $\sigma_x^2 = k_0 f_h = \sigma_{y,White}^2 \frac{\tau^2}{3}$. Rewriting the previous equation,

$$TIE_W^2(T_m, T_p) = k_0\left(\frac{1}{\tau_0} + 180\frac{T_p^4}{T_m^5} + 360\frac{T_p^3}{T_m^4} + 252\frac{T_p^2}{T_m^3} + 72\frac{T_p}{T_m^2} + \frac{9}{T_m}\right) \tag{2.2}$$

And if the noise is pure white PN,

$$\sigma_{x,white}^2 = k_0 f_h = \sigma_{y,White}^2 \frac{\tau^2}{3} \tag{2.3}$$

And the contribution of white phase noise can be written

$$TIE_W^2(T_m, T_p) = \frac{\tau^2 \tau_0 \sigma_{y,White}^2}{3}\left(\frac{1}{\tau_0} + 180\frac{T_p^4}{T_m^5} + 360\frac{T_p^3}{T_m^4} + 252\frac{T_p^2}{T_m^3} + 72\frac{T_p}{T_m^2} + \frac{9}{T_m}\right) \tag{2.4}$$

Obviously, there is no finite value of $T_m$ to yield the lowest uncertainty. Instead TIE ($\sqrt{TIE_W^2}$) falls to $\sigma_{x,white} = \frac{\tau\sigma_{y,white}}{\sqrt{3}}$ as $T_m$ approaches infinity while the errors due to unknown quadratic fitting parameters become infinitesimal and the only uncertainty is due to the white phase noise at the next point.

**III. Random Walk (RW) phase noise**

RW would be the dominant contribution for a GDC's holdover at one day. As we shall see, the optimal computation period $t_m$ would be about ten days, so measurement noise should largely drop out and hopefully the variation of the frequency drift (IRW's contribution) would not yet be serious. To explore this, we need only consider equation 60 of Vernotte et al. (2001), which is reproduced here:

$$TIE_{RW}^2(T_m, T_p) = \frac{6\pi^2 k_{-2} T_m}{35}\left(50\frac{T_p^4}{T_m^4} + 100\frac{T_p^3}{T_m^3} + 69\frac{T_p^2}{T_m^2} + 19\frac{T_p}{T_m} + 1\right) \quad \text{(3.1 and 60 of Vernotte et al. (2001)}$$

In this equation $k_{-2}$ denotes the PSD coefficient of the RW noise component so that according to Vernotte et al. (2001) and equation 98 of Barnes et al. (1971):

$$\sigma_{y,RW}^2 = \frac{2\pi^2 k_{-2}}{\tau} = \frac{h_0}{2\tau} \tag{3.2}$$

Therefore the contribution of RW noise to TIE can be written:

$$TIE_{RW}^2(T_m, T_p) = \frac{3\tau\sigma_{y,RW}^2}{35}\left(50\frac{T_p^4}{T_m^3} + 100\frac{T_p^3}{T_m^2} + 69\frac{T_p^2}{T_m} + 19T_p + T_m\right) \tag{3.3}$$

In order to find the optimal $T_m$ we set the derivative with respect to it equal to 0. Writing $r = T_m/T_p$

$$0 = -150r^{-4} - 200r^{-3} - 69r^{-2} + 1 = -150 - 200r - 69r^2 + r^4 \tag{3.4}$$

This quartic equation can be solved analytically, and we find it has one positive root, at r=9.56774**.** So for optimal holdover at 1 day, we might want 9.5 days of data. But given that we may be solving over diurnal variations, 10.0 days would be safer. In a low-cost GDC the data would likely have to be decimated while an optimally tuned Kalman or exponential filter can discard old data while retaining their essence.

Fortunately the RW minimum turns out to be quite broad, so the price of using far less data is not that large. To find the width of the minimum in the case of RW, we can simplify equation (3.3) as follows

$$TIE_{RW}^2 \propto 50r^{-3} + 100r^{-2} + 69r^{-1} + 19 + r \tag{3.5}$$

Figure 3.1 shows how the minimum depends on r. The ordinate is the uncertainty at the given r divided by the uncertainty at the optimal r, where $r=T_m/T_p$. A baseline of 1 day increases the prediction rms by a factor of 2.5.

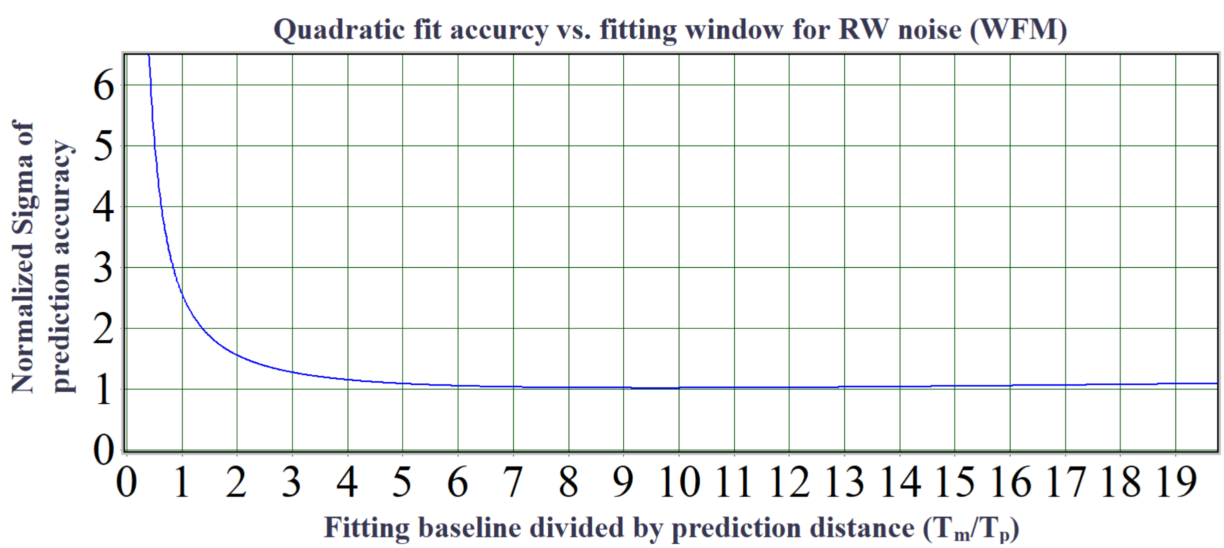


**FIGURE 3.1** Accuracy of quadratic fit's prediction at time $T_p$ in future with modelling baseline $T_m$

## IV. Integrated Random Walk (IRW, RR)

Equation (62) of Vernotte et al. (2001) gives the expected TIE for an IRW, which is merely reproduced here:

$TIE^2_{IRW}(T_m, T_p) = \frac{2\pi^4 k_{-4} T_m^3}{315}(450\frac{T_p^4}{T_m^4} + 690\frac{T_p^3}{T_m^3} + 303\frac{T_p^2}{T_m^2} + 42\frac{T_p}{T_m} + 2)$ (4.1 and 62 of Vernotte et al. (2001))

where $k_{-4}$ is the PSD of IRW.

$$TIE^2_{IRW}(T_m, T_p) = \frac{2\pi^4 k_{-4}}{315}(450\frac{T_p^4}{T_m} + 690T_p^3 + 303T_p^2 T_m + 42T_p T_m^2 + 2T_m^3) \tag{4.2}$$

Since for an IRW, $\sigma^2_{y,IRW} = \frac{2\pi^2\tau h_{-2}}{3} = \frac{8\pi^4\tau k_{-4}}{3}$ ; $k_{-4} = \frac{3\sigma^2_{y,IRW}}{8\pi^4\tau}$ [2, equation 87]). (4.3)

$$TIE^2_{IRW}(T_m, T_p) = \frac{\sigma^2_{y,IRW}}{420\tau}(450\frac{T_p^4}{T_m} + 690T_p^3 + 303T_p^2 T_m + 42T_p T_m^2 + 2T_m^3) \tag{4.4}$$

Setting zero equal to the derivative with respect to $T_m$, and with $r = T_m/T_p$

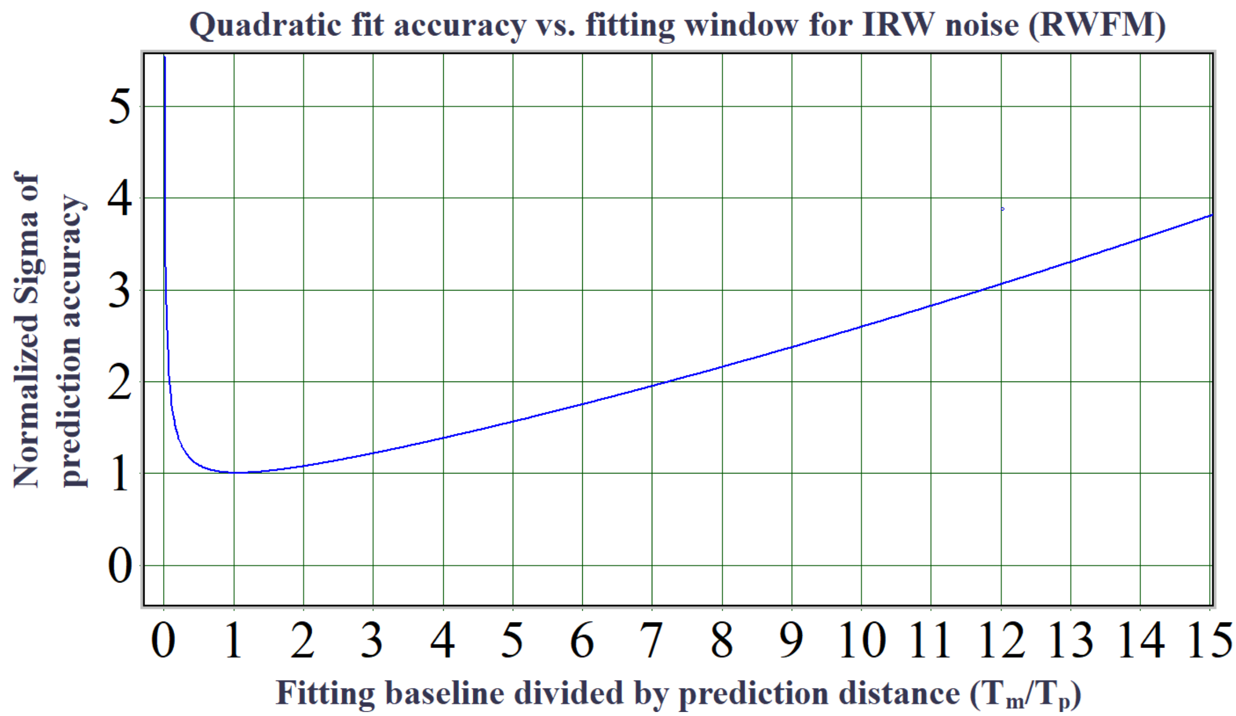
$$0 = -450\frac{T_p^4}{T_m^2} + 303T_p^2 + 84T_p T_m + 6T_m^2 = -150 + 101r^2 + 28r^3 + 3r^4 \tag{4.5}$$

This results in just one positive root, r=1.06 ($T_m = 1.06 * T_p$). One would expect a small optimum modeling window for IRW because RW benefits from averaging white frequency noise over a long window, while the larger frequency variations expected with IRW degrades the solution if the lookback time is too large.

To compute width of the minimum in the case of pure IRW noise $T_p$ can be taken to be a constant and:

$$TIE^2_{IRW} \propto 450r^{-1} + 690 + 303r + 42r^2 + 2r^3 \tag{4.6}$$

Figure 4.1 shows how the minimum depends on r. The ordinate is the uncertainty at the given r divided by the uncertainty at the optimal r, where $r = T_m/T_p$. A baseline of 1 day doubles the prediction rms.




**F4.1**. Width of minimum for IRW noise

**V. Optimal baseline for least squares fit to a linear combination of W, RW, and IRW noise.**

Ignoring white noise for now, the solution vector $X_{fit}$ (phase, frequency, and frequency drift) for a least-squares fit is given by

$$X_{fit} = (A^T * S^{-1} * A)^{-1} * A^T * S^{-1} * Y \tag{5.1}$$

Where $Y$ is the data vector, $S$ is the observation covariance matrix taken to be unity, and $A$ is the design matrix ($Y = A * X$), $X$ being the vector representing the true values of the parameters. The parameter estimation error, $E = (A^T * S^{-1} * A)^{-1}$, in the presence of linear combination RW noise and IRW noise is a linear combination of independent solutions to the noise types, and the squared error is the sum of the squared errors of those three independent noise types.

$$TIE^2_{W+RW+IRW} = k_0\left(\frac{1}{\tau_0} + 180\frac{T_p^4}{T_m^5} + 360\frac{T_p^3}{T_m^4} + 252\frac{T_p^2}{T_m^3} + 72\frac{T_p}{T_m^2} + \frac{9}{T_m}\right) + \frac{6\pi^2 k_{-2}}{35}\left(50\frac{T_p^4}{T_m^3} + 100\frac{T_p^3}{T_m^2} + 69\frac{T_p^2}{T_m} + 19T_p + T_m\right) + \frac{2\pi^4 k_{-4}}{315}\left(450\frac{T_p^4}{T_m} + 690T_p^3 + 303T_p^2 T_m + 42T_p T_m^2 + 2T_m^3\right) \tag{5.2}$$

For computational purposes later on the "k" terms can be replaced by their contribution to the Allan variance of the combination:

$$TIE^2_{W+RW+IRW} = \frac{\sigma^2_{y,White}\tau_0\tau^2}{3}\left(\frac{1}{\tau_0} + 180\frac{T_p^4}{T_m^5} + 360\frac{T_p^3}{T_m^4} + 252\frac{T_p^2}{T_m^3} + 72\frac{T_p}{T_m^2} + \frac{9}{T_m}\right) + (\frac{3\tau\sigma^2_{y,RW}}{35}\left(50\frac{T_p^4}{T_m^3} + 100\frac{T_p^3}{T_m^2} + 69\frac{T_p^2}{T_m} + 19T_p + T_m\right) + \frac{\sigma^2_{y,IRW}}{420\tau}\left(450\frac{T_p^4}{T_m} + 690T_p^3 + 303T_p^2 T_m + 42T_p T_m^2 + 2T_m^3\right) \tag{5.3}$$

Going back to equation (5.2) The derivative with respect to $T_m$ is

$$0 = -k_0\left(900\frac{T_p^4}{T_m^6} + 1440\frac{T_p^3}{T_m^5} + 756\frac{T_p^2}{T_m^4} + 144\frac{T_p}{T_m^3} + \frac{9}{T_m^2}\right) + \frac{6\pi^2 k_{-2}}{35}\left(-150\frac{T_p^4}{T_m^4} - 200\frac{T_p^3}{T_m^3} - 69\frac{T_p^2}{T_m^2} + 1\right) + \frac{2\pi^4 k_{-4}}{(35*9)}\left(-450\frac{T_p^4}{T_m^2} + 303T_p^2 + 84T_pT_m + 6T_m^2\right) \tag{5.4}$$

Multiplying by $T_m^6$

$$0 = -k_0\left(900T_p^4 + 1440T_mT_p^3 + 756T_m^2T_p^2 + 144T_m^3T_p + 9T_m^4\right) + \frac{6\pi^2 k_{-2}}{35}\left(-150T_m^2T_p^4 - 200T_m^3T_p^3 - 69T_m^4T_p^2 + T_m^6\right) + \frac{2\pi^4 k_{-4}}{(35*3)}\left(-150T_m^4T_p^4 + 101T_m^6T_p^2 + 28T_pT_m^7 + 2T_m^8\right) \tag{5.5}$$

$$0 = \frac{-k_0}{T_p^2}\left(900T_p^6 + 1440T_mT_p^5 + 756T_m^2T_p^4 + 144T_m^3T_p^3 + 9T_m^4T_p^2\right) + \frac{6\pi^2 k_{-2}}{35}\left(-150T_m^2T_p^4 - 200T_m^3T_p^3 - 69T_m^4T_p^2 + T_m^6\right) + \frac{2\pi^4 k_{-4}T_p^2}{(35*3)}\left(-150T_m^4T_p^2 + 101T_m^6 + 28T_m^7T_p^{-1} + 2T_m^8T_p^{-2}\right) \tag{5.6}$$

With R=$\frac{9k_{-2}\,T_p^{-2}}{\pi^2 k_{-4}}$ and $R' = \frac{105k_0}{2\pi^4 k_{-4}T_p^4}$ (from dividing the previous equation by $\frac{2\pi^4 k_{-4}T_p^2}{105}$)
Note that while R, R', and later f depend on both $T_p$ and the $k_n$, they can be rewritten in terms of only the relative contributions of the noise types at $\tau = T_p$.

$$0 = R'\left(-900T_p^6 - 1440T_mT_p^5 - 756T_m^2T_p^4 - 144T_m^3T_p^3 - 9T_m^4T_p\right) + R\left(-150T_m^2T_p^4 - 200T_m^3T_p^3 - 69T_m^4T_p^2 + T_m^6\right) - 150T_m^4T_p^2 + 101T_m^6 + \frac{28T_m^7}{T_p} + \frac{2T_m^8}{T_p^2} \tag{5.7}$$

$$0 = -R'\left(900T_p^6 - 1440rT_p^6 - 756r^2T_p^6 - 144r^3T_p^6 - 9r^4T_p^6\right) + R\left(-150r^2T_p^6 - 200r^3T_p^6 - 69r^4T_p^6 + r^6T_p^6\right) - 150r^4T_p^6 + 101r^6T_p^6 + 28r^7T_p^6 + 2r^8T_p^6 \tag{5.8}$$

$$0 = R'(-900 - 1440r - 756r^2 - 144r^3 - 9r^4) \quad + R(-150r^2 - 200r^3 - 69r^4 + r^6) - 150r^4 + 101r^6 + 28r^7 + 2r^8 \tag{5.9}$$

$$0 = -900R' - 1440R'r + r^2(-756R' - 150R) - r^3(200R + 144R') - r^4(69R + 150 - 9R') + r^6(R + 101) + 28r^7 + 2r^8 \tag{5.10}$$

with $f = \frac{R}{1+R} = \frac{1}{1+\frac{1}{R}} = \frac{1}{1+\frac{\pi^2 k_{-4}}{9k_{-2}}T_p^2} = \frac{9k_{-2}}{9k_{-2}+\pi^2 k_{-4}T_p^2} = \frac{9k_{RW}}{9k_{RW}+\pi^2 k_{IRW}T_p^2}$ , $R = \frac{f}{1-f}$ , $R'' = (1-f)R'$ equation (5.10) becomes

$$0 = -900R'' - 1440R''r + r^2(-756R'' - 150f) - r^3(200f - 144R'') - r^4(69f + 150(1-\mathrm{f}) + 9R'') + r^6(f + 101(1-f)) + 28(1-f)r^7 + 2(1-f)r^8 \tag{5.11}$$

This eight-order polynomial that reduces to sixth order if white noise is negligible, so that R'=R''=0.

$$0 = -150R - 200Rr - (69R + 150)r^2 + (R + 101)r^4 + 28r^5 + 2r^6) \tag{5.12}$$

And $0 = -150f - 200fr - (69f + 150(1-f))r^2 + (f + 101(1-f))r^4 + 28(1-f)r^5 + 2(1-f)r^6)$ (5.13)

This equation is easily solved in Matlab using the function "roots", with results as shown in the Figures 5.1-5.3.

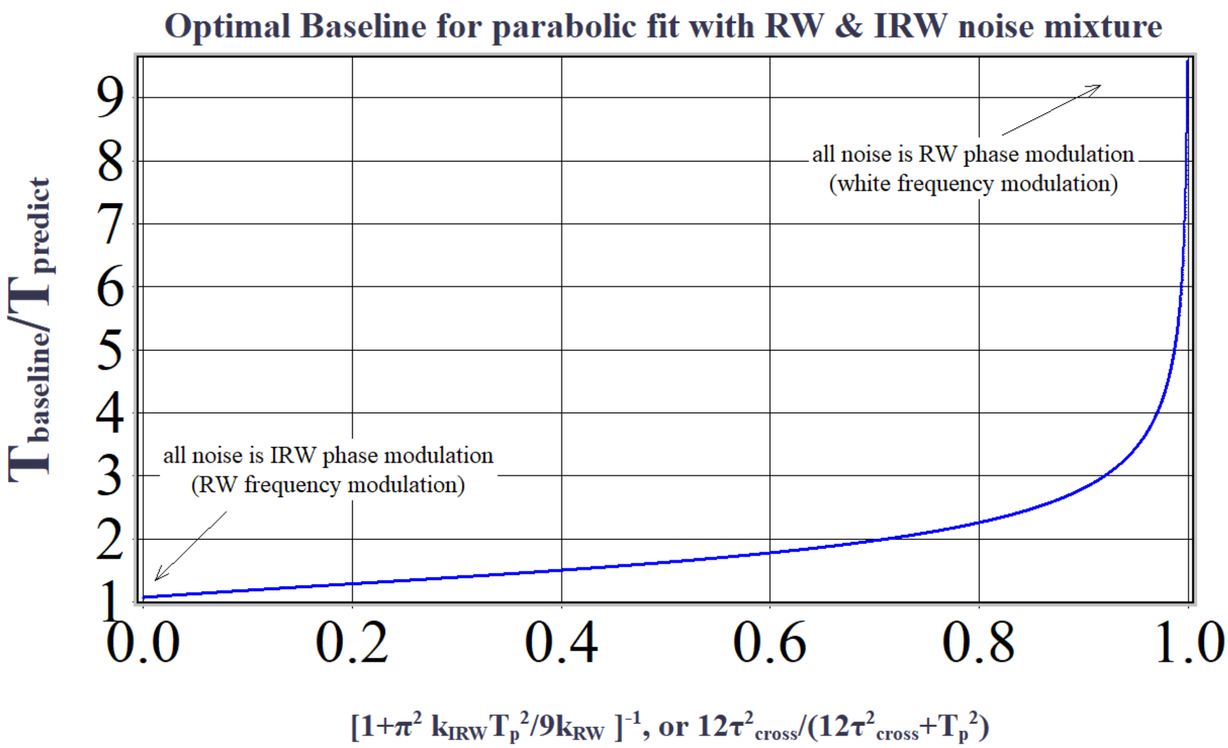



**FIGURE 5.1** Optimal baseline for prediction in presence of RW and IRW noise mixture

Figure 3 can also be drawn in a different but equivalent way. Adopting the notation $h_{RW} = h_0$, $k_{RW} = k_{-2}$ , $h_{IRW} = h_{-2}$ , and $k_{IRW} = k_{-4}$, and using the already presented relations $\sigma^2_{y,RW} = \frac{h_{RW}}{2\tau} = \frac{(2\pi)^2 k_{-2}}{2\tau}$ and $\sigma^2_{y,IRW} = \frac{h_{IRW}(2\pi)^2\tau}{6}$, we find that if the contributions of RW and IRW to the overall Allan variance are equal at $\tau = \tau_{cross}$,

$\frac{h_{RW}}{2\tau_{cross}} = \frac{h_{IRW}(2\pi)^2\tau_{cross}}{6}$ and $\frac{h_{RW}}{h_{IRW}} = \frac{k_{RW}}{k_{IRW}} = \frac{(2\pi)^2\tau^2_{cross}}{3} \sim 13\tau^2_{cross}$ . Therefore "f" in equation 5.11 can be expressed $f = \frac{1}{1+\frac{\pi^2 k_{-4}}{9k_{-2}}T_p^2} = \frac{1}{1+\frac{\pi^2 k_{IRW}}{9k_{RW}}T_p^2} = \frac{1}{1+\frac{T_p^2}{12\tau^2_{cross}}} = \frac{12\tau^2_{cross}}{12\tau^2_{cross}+T_p^2}$ (5.14)

Rearranging,

$12f\tau^2_{cross} + fT_p^2 = 12\tau^2_{cross}$ , $fT_p^2 = 12(1-f)\tau^2_{cross}$ , and $\frac{\tau^2_{cross}}{T_p^2} = \frac{f}{12(1-f)}$ (5.15)

Note that as $T_p$ increases (and $\tau_{eq}$ decreases) the importance of the IRW contribution goes up relative to the RW.

Figure 5.2 shows the same results are Figure 5.1, except using $\frac{\tau_{cross}}{T_P}$ as the abscissa.

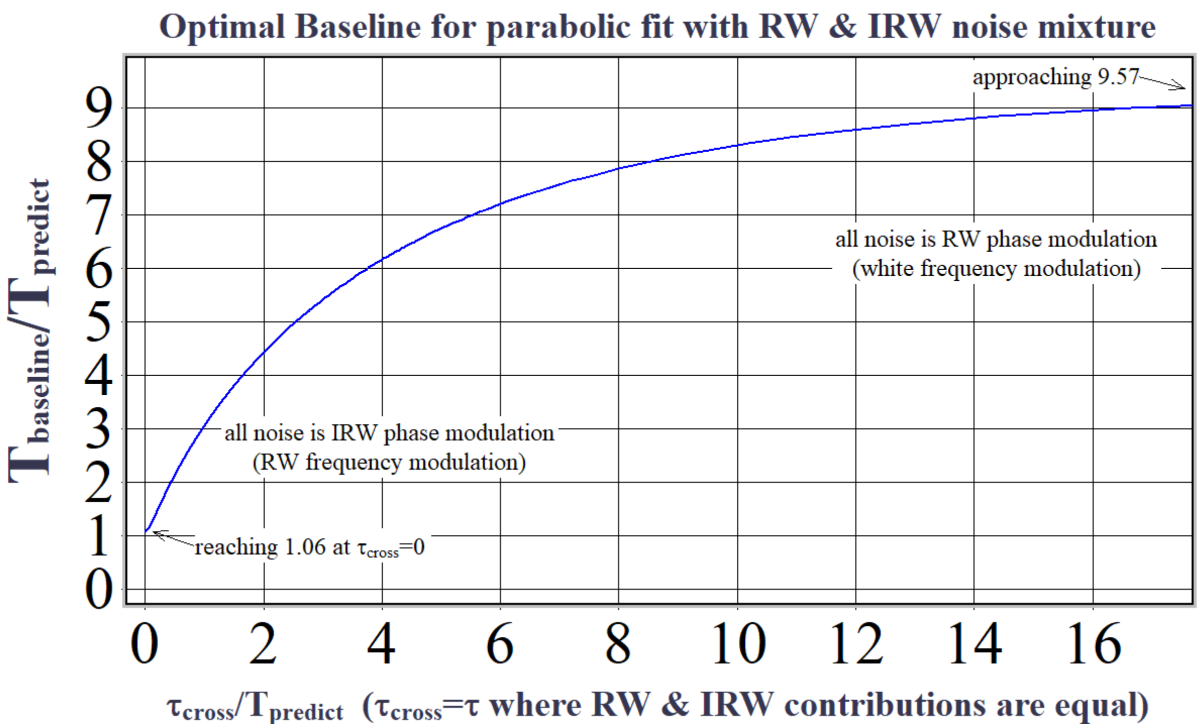


FIGURE 5.2 Optimal baseline to prediction ratio as function of $\frac{\tau_{cross}}{T_P}$

To explore the width of the minima in the absence of white phase noise, equation (5.1) can be expressed

$$TIE^2_{RW+IRW} = \frac{6\pi^2 k_{-2} T_p}{35}\left(\frac{50}{r^3} + \frac{100}{r^2} + \frac{69}{r} + 19 + r\right) + \frac{2\pi^4 k_{-4} T_p^3}{315}\left(\frac{450}{r} + 690 + 303r + 42r^2 + 2r^3\right) \qquad (5.16)$$

$$TIE^2_{RW+IRW} = \frac{6\pi^2 k_{-2} T_p}{35}\left(\frac{50}{r^3} + \frac{100}{r^2} + \frac{69}{r} + 19 + r\right) + \frac{2\pi^4 k_{-4} T_p^3}{(9*35)}\left(\frac{450}{r} + 690 + 303r + 42r^2 + 2r^3\right) \qquad (5.17)$$

If $C = \frac{6\pi^2 k_{-2} T_p}{35} + \frac{2\pi^4 k_{-4} T_p^3}{9*35} = \frac{2\pi^2 T_p}{35}\left(3k_{-2} + \frac{\pi^2 k_{-4} T_p^2}{9}\right)$ (5.18)

$$TIE^2_{RW+IRW} = C\{a\left(\frac{50}{r^3} + \frac{100}{r^2} + \frac{69}{r} + 19 + r\right) + (1-a)\left(\frac{450}{r} + 690 + 303r + 42r^2 + 2r^3\right)\} \qquad (5.19)$$

Where $a = \frac{6\pi^2 k_{-2} T_p}{2\pi^2 T_p\left(3k_{-2} + \frac{\pi^2 k_{-4} T_p^2}{9}\right)} = \frac{3k_{-2}}{3k_{-2} + \frac{\pi^2 k_{-4} T_p^2}{9}} = \frac{1}{1 + \frac{\pi^2 k_{-4} T_p^2}{9k_{-2}}} = \frac{1}{1 + \frac{3\pi^2 T_p^2}{18\pi^2 \tau_{cross}^2}} = \frac{6\tau_{cross}^2}{6\tau_{cross}^2 + T_p^2}$ (5.20)

$6a\tau_{cross}^2 + aT_p^2 6\tau_{cross}^2 + T_p^2$; therefore $\qquad \tau_{cross}^2 = \frac{1-6\alpha}{6\alpha} T_p^2$ (5.21)

Graphically equation (5.18) is shown in Figure 5.3, where 11 curves are plotted and normalized to their individual TIE minima. Here the n[th] curve from the bottom has $\alpha_n$=0.1*(n-1) pure IRW contribution, and the overall inaccuracy is scaled by sqrt(C'), and the absolute contribution of RW for the n[th] curve is $\sqrt{\frac{27k_{-2}}{\pi^2 k_{-4} T_p^2}} = \sqrt{\frac{27k_{RW}}{\pi^2 k_{IRW} T_p^2}} = \sqrt{\frac{36\tau_{cross}^2}{T_p^2}} = \frac{6\tau_{cross}}{T_p}$.

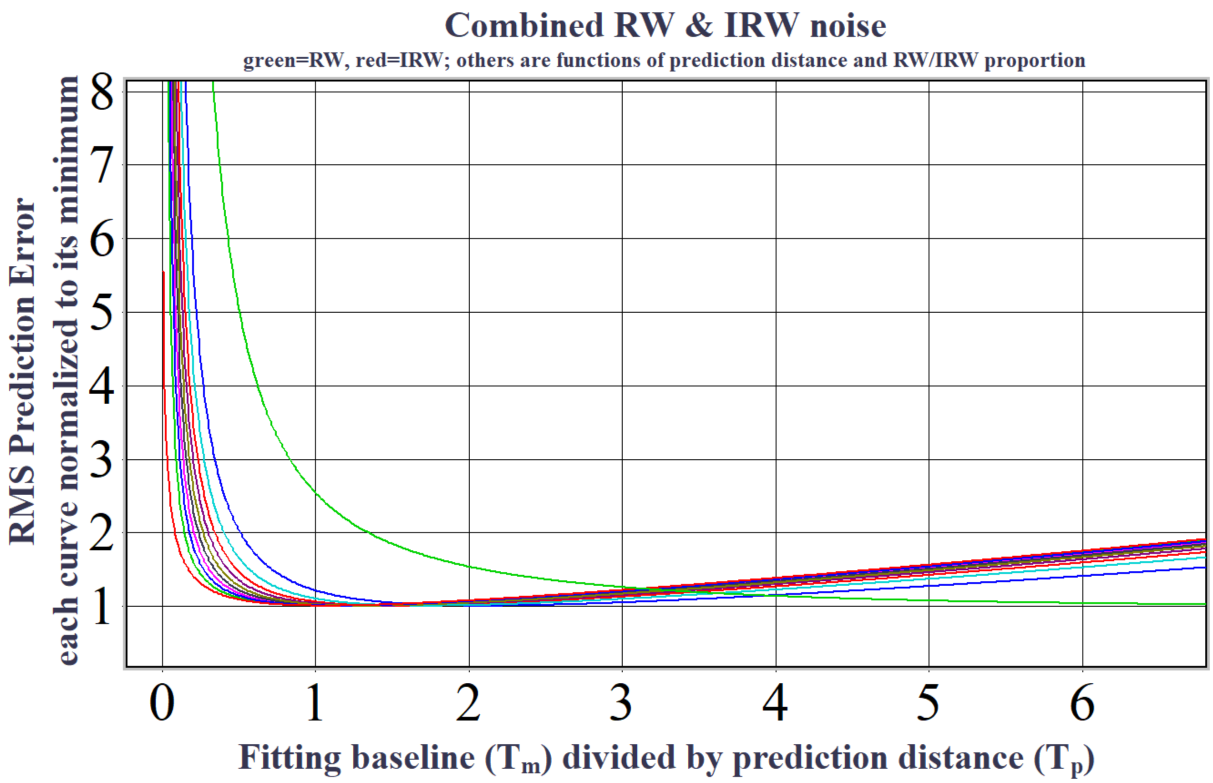



Figure 5.3 RMS Prediction error for combinations of RW and IRW noise. The curves in between the green (RW) and red (IRW) for functions of the noise ratios plus prediction distance $T_p$.

### VI. Computing the uncertainty of a Kalman fit

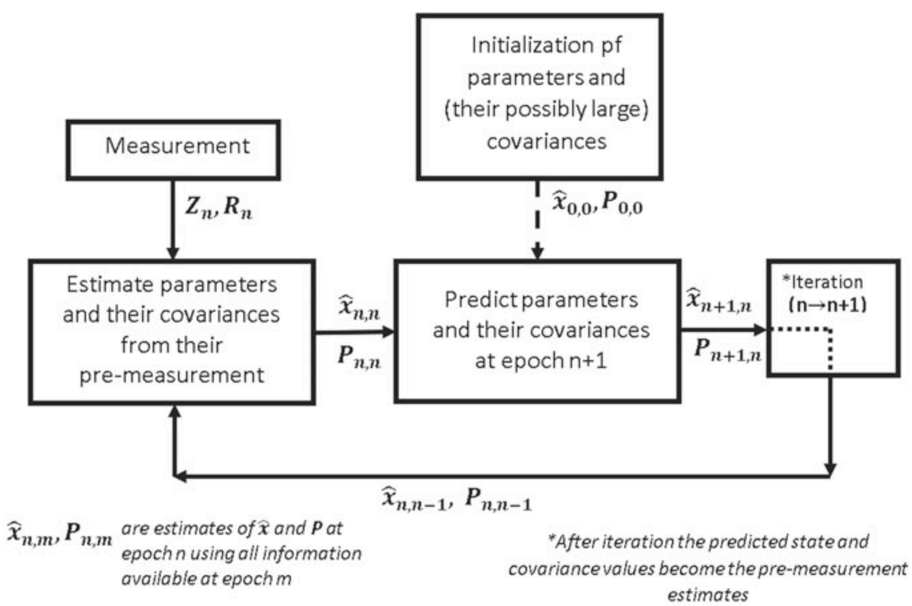



**FIGURE 6.1** Kalman Filter flow chart (Banerjee and Matsakis, 2023)

Figure 6.1 shows the Kalman formalism. If the process and measurement noises (Q and R) do not change, then the parameter uncertainty covariances (P) will reach a steady state in which the reduction when new measurements are taken (determined by R, and P) is cancelled by the increase between measurements (determined by Q). In terms of the equations, the steady-state parameters can be found by setting $P^m$ equal to what it is after lowering it by the measurement and raising it by the evolution amount:

$$P^m = \Phi(\mathrm{I} - \mathrm{KH})P^m\Phi' + Q \tag{6.1}$$
$$P^m - Q = \Phi(\mathrm{I} - KH)P^m\Phi' \tag{6.2}$$
$$P^m = Q + \Phi P^m\Phi' - \Phi KHP^m\Phi' \tag{6.3}$$

Where the Kalman Gain is $K = P^m H'(HP^m H' + R)^{-1}$ (6.4)

The equations are easy to solve in the case of one dimension, where the only variable is phase, because $\Phi = \mathrm{H} = 1$ and

$$P^m = P^m - P^m(P^m + R)^{-1}P^m + Q \tag{6.5}$$

$$P^m(P^m + R)^{-1}P^m = Q \tag{6.6}$$

$$(P^m)^2 = Q(P^m + R)\ ;\ (P^m)^2 - QP^m - QR = 0 \tag{6.7}$$

and

$$P^m = \frac{Q + \sqrt{Q^2 + 4RQ}}{2}\ ;\ P = \frac{-Q + \sqrt{Q^2 + 4RQ}}{2} \tag{6.8}$$

For the higher dimensional problems, the above Riccati equation (3.3) can be perfectly solved using Matlab's function dlqr. That actual expression, in Matlab notation, is:

[G,$P^m$,e]=dlqr(Phi',H',Q,R) (6.9)

Where Phi, H, Q, R, and $P^m$ have the same meaning as in equations (6.1) to (6.3), and the apostrophe indicates transpose. G and e are not relevant here.

In order to predict the holdover, one can solve for the steady-state $P^m$ and then propagate the solution forward as if there were no measurements, equivalent to setting the Kalman Gain to zero. Going $n$ steps into the future and realizing that $P^m = P$ in the absence of a measurement,

$$P^m = (\Phi P^m\Phi' + Q)^n \quad \text{and } P = (\Phi P\Phi' + Q)^n \tag{6.10}$$

This has been numerically confirmed by running solutions with a Kalman Filter and its library function dare.

## VII. Computing the Quadratic and Kalman fits

The Figure 7.1 shows the performance of a Kalman filter versus that of a quadratic least-squares fit for random run. In one curve for the quadratic least squares fit, the fitting baseline $T_m$ is 100, in the other it is whatever $T_m$ yields the lowest RMS for a prediction $T_p$ points into the future. The Kalman solution does not have fixed baselines, although the ratio of the process noise Q to measure noise variance R can be used to define effective look-back times for phase, frequency, and drift.

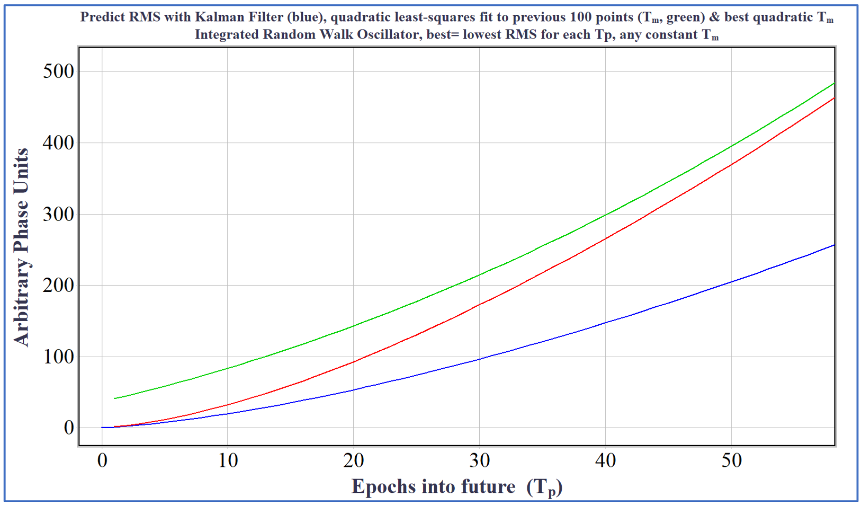



**FIGURE 7.1** Comparison of RMS for quadratic fit and Kalman filter in presence of IRW noise

The Kalman filter can be mistuned, and Figure 7.2 shows what can happen under extreme mistuning. In the figure, the usually highest curve is the predictive accuracy of a Kalman Filter that underestimates the process noise (Q) by a factor of 100, while the red curve, representing a quadratic fit, becomes worse for the most distant epochs. The green curve is a Kalman filter with too high a Q – raising the Q is often a good idea when mismodelling is suspected, because it leads to greater weight for newer data. The lowest curve is for a perfectly tuned Kalman filter.

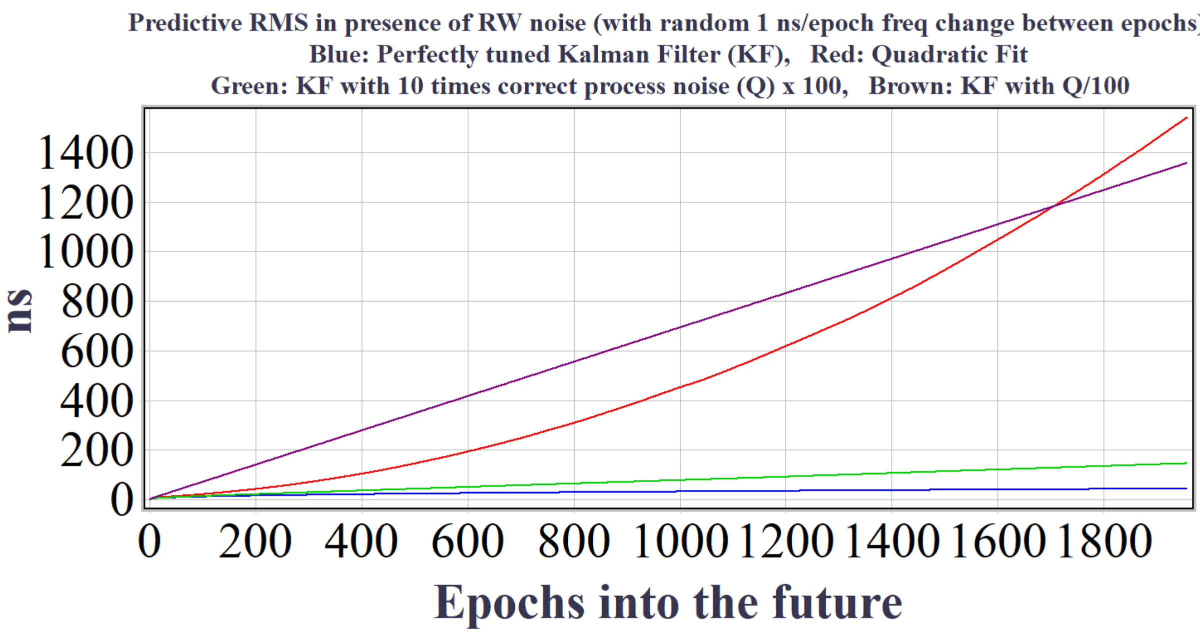



**FIGURE 7.2**. Prediction accuracy of correctly and incorrectly tuned Kalman filters for RW noise

**VIII. Conclusions**

The analysis of Vernotte et al. (2001) has been extended to find the optimal modelling baseline for an equally weighted quadratic fit. While the optimal modelling period can be impractically long, compromises can be made. However, a correctly tuned Kalman filter is better than the best quadratic fit.

**IX. Acknowledgements**

We thank the staff of Masterclock, Inc. whose constant efforts to build better products inspired us to look into these matters.

**REFERENCES**

Banerjee, P., and Matsakis, D. (2013). Introduction to Modern Timekeeping and Time Transfer, Springer Press.

Barnes, J., Chi, A., Cutler, L., Healey, D., Leeson, D, McGunigal, T., Mullen, J., Smith, W., Sydnor, R., Vessot, F. and Winkler, G. (1971). Characterization of Frequency Stability, IEEE-TM 20, pp 105-120.

Brown, R.G. and Hwang, P.Y.C. (1997). Introduction to Random Signals and Applied Kalman Filtering, with MATLAB Exercises and Solutions, Third Edition, J. Wiley & Sons, page 430.

Van Dierendonck, A.,J., McGrow, J.B., and R.G. Brown (1984), Proceedings Precise Time and Time Interval Meetings (PTTI), Note that the book Brown and Hwang corrects a mistake in this paper.
.
Vernotte, F., Delporte, J., Brunet, M. and Tournier, T., Metrologia, (2001). Uncertainties of drift coefficients and extrapolation errors: Application to clock error prediction, E1319.

**Appendix I. Uncertainty of estimation for equal-weight quadratic fit presence of white phase and measurement noise.**

In the notation of this paper, the standard least squares polynomial fit formalism is:

$$L = AX \tag{A1.1}$$

$$C = (A^T * S^{-1} * A)^{-1} \tag{A1.2}$$

Where *C* is the covariance matrix; *L* is the observation vector; *S* is the uncertainty matrix of those observations, which for uncorrelated observations characters by white noise is $k_0 I$; X is the 3-parameter vector of phase, frequency, and frequency drift. *A* is the observation matrix given by

$$\text{and} \quad A^T = \begin{pmatrix} 1 & 1 & 1 & \dots \\ t_1 - t_N & t_2 - t_N & t_3 - t_N & \dots \\ (t_1 - t_N)^2 & (t_2 - t_N)^2 & (t_3 - t_N)^2 & \dots \end{pmatrix} \text{ because } A = \begin{pmatrix} 1 & t_1 - t_N & (t_1 - t_N)^2 \\ 1 & t_2 - t_N & (t_2 - t_N)^2 \\ 1 & t_3 - N & (t_3 - t_N)^2 \\ \dots & \dots & \dots \end{pmatrix} \tag{A1.3}$$

and the modelling baseline $T_m$ corresponds to the N observations. For white phase noise of amplitude W, $S=W^2I$ and

$$C^{-1} = A^T * S^{-1} * A = W^{-2} \begin{pmatrix} N & \sum_{n=1}^{N}(t_n - t_N) & \sum_{n=1}^{N}(t_n - t_N)^2 \\ \sum_{n=1}^{N}(t_n - t_N) & \sum_{n=1}^{N}(t_n - t_N)^2 & \sum_{n=1}^{N}(t_n - t_N)^3 \\ \sum_{n=1}^{N}(t_n - t_N)^2 & \sum_{n=1}^{N}(t_n - t_N)^3 & \sum_{n=1}^{N}(t_n - t_N)^4 \end{pmatrix} \tag{A1.4}$$

The parameters are not orthonormal, because for holdover we are most interested in the instantaneous phase, frequency, and drift at the last data point of the fit. But the covariance matrix of the parameters is the inverse of (A1.4) can found by standard techniques. For example, chose units so that $t_n = n$. If you set $t_n = n\tau + t_0$, all it does is scale the frequency parameters and terms by $\tau$ and the frequency drift terms by $\tau^2$, both of those being squared for covariance matrices given their relation to variances.

$$\text{then } \sum_{n=1}^{N}(t_n - t_N) = \sum_{n=1}^{N-1} n = \frac{-N(N-1)}{2} \tag{A1.5}$$

$$\sum_{n=1}^{N}(t_n - t_N)^2 = \sum_{n=1}^{N-1} n^2 = \frac{N(N-1)(2N-1)}{6} \tag{A1.6}$$

$$\sum_{i=1}^{N}(t_i - t_N)^3 = -\sum_{n=1}^{N-1} n^3 = \frac{-N^2(N-1)^2}{4} \quad \text{(A1.7)}$$

and $\sum_{i=1}^{N}(t_i - t_N)^4 = \sum_{n=1}^{N-1} n^4 = \frac{N(N-1)(2N-1)(3N^2-3N-1)}{30}$ (A1.8)

Then

$$C^{-1} = A^T * S^{-1} * A = W^{-2}N \begin{pmatrix} 1 & \frac{-(N-1)}{2} & \frac{(N-1)(2N-1)}{6} \\ \frac{-(N-1)}{2} & \frac{(N-1)(2N-1)}{6} & \frac{-N(N-1)^2}{4} \\ \frac{(N-1)(2N-1)}{6} & \frac{-N(N-1)^2}{4} & \frac{(N-1)(2N-1)(3N^2-3N-1)}{30} \end{pmatrix} \quad \text{(A1.9)}$$

$$C = \frac{W^2}{N} Inv \begin{pmatrix} 1 & \frac{-(N-1)}{2} & \frac{(N-1)(2N-1)}{6} \\ \frac{-(N-1)}{2} & \frac{(N-1)(2N-1)}{6} & \frac{-N(N-1)^2}{4} \\ \frac{(N-1)(2N-1)}{6} & \frac{-N(N-1)^2}{4} & \frac{(N-1)(2N-1)(3N^2-3N-1)}{30} \end{pmatrix} = \frac{W^2}{N} Inv(M) \quad \text{(A.1.10)}$$

The determinant of the matrix "M" in the above equation is

$$detM = \frac{(N-1)(2N-1)}{6}\frac{(N-1)(2N-1)(3N^2-3N-1)}{30} - \frac{N^2(N-1)^4}{16} - \frac{(N-1)}{2}\frac{(N-1)(N-1)(2N-1)(3N^2-3N-1)}{60} + \frac{(N-1)}{2}\frac{N(N-1)^2}{4}\frac{(N-1)(2N-1)}{6} + \frac{(N-1)(2N-1)}{6}\frac{(N-1)}{2}\frac{N(N-1)^2}{8} - \frac{(N-1)(2N-1)}{6}\frac{(N-1)(2N-1)}{6}\frac{(N-1)(2N-1)}{6} \quad \text{(A1.11)}$$

$$= \frac{(N-1)^2(2N-1)^2(3N^2-3N-1)}{180} - \frac{N^2(N-1)^4}{16} - \frac{(N-1)^3(2N-1)(3N^2-3N-1)}{120} + \frac{N(N-1)^3(N-1)(2N-1)}{48} + \frac{N(N-1)(2N-1)(N-1)^3}{48} - \frac{(N-1)(2N-1)^3(N-1)^2}{216} \quad \text{(A1.12)}$$

$$= \frac{(N-1)^2(2N-1)^2(3N^2-3N-1)}{180} - \frac{N^2(N-1)^4}{16} - \frac{(N-1)^3(2N-1)(3N^2-3N-1)}{120} + \frac{N(N-1)^4(2N-1)}{24} - \frac{(2N-1)^3(N-1)^3}{216} \quad \text{(A1.13)}$$

$$= \frac{(N-1)^2}{4}\left[\frac{(2N-1)^2(3N^2-3N-1)}{45} - \frac{N^2(N-1)^2}{4} - \frac{(N-1)(2N-1)(3N^2-3N-1)}{30} + \frac{N(N-1)^2(2N-1)}{6} - \frac{(2N-1)^3(N-1)}{54}\right] \quad \text{( A1.14)}$$

The inverse of a matrix equals its adjugate (adjoint) divided by is determinant. The adjugate matrix of M is the transpose of the cofactor matrix, but since all these matrices are symmetric transposing isn't necessary.

$$C(1,1) = \frac{W^2}{Ndet(M)} adjM(1,1) = \frac{W^2}{Ndet(M)}\left[\frac{(N-1)^2(2N-1)^2(3N^2-3N-1)}{180} - \frac{N^2(N-1)^4}{16}\right] = \frac{W^2(N-1)^2}{4Ndet(M)}\left[\frac{(2N-1)^2(3N^2-3N-1)}{45} - \frac{N^2(N-1)^2}{4}\right] = \frac{W^2(N-1)^2}{4Ndet(M)}\left[\frac{(3N^2-3N-1)(4N^2-4N+1)}{45} - \frac{N^2(N^2-2N+1)}{4}\right] \quad \text{(A.1.15)}$$

The last term in brackets is

$$\frac{12N^4-12N^3+3N^2-12N^3+12N^2-3N-4N^2+4N-1}{45} - \frac{N^4-2N^3+N^2}{4} = \frac{12N^4-24N^3+11N^2+N-1}{45} + \frac{-N^4+2N^3-N^2}{4} \quad \text{(A1.16)}$$

$$= \frac{48N^4-96N^3+44N^2+4N-4-45N^4+90N^3-45N^2}{45*4} = \frac{3N^4-6N^3-N^2+4N-4}{180} \quad \text{(A1.17)}$$

So $C(1,1) = \frac{W^2(N-1)^2(3N^4-6N^3-N^2+4N-4)}{720Ndet(M)}$ (A1.18)

$$C(2,2) = \frac{W^2}{Ndet(M)} adjM(2,2) = \frac{W^2}{Ndet(M)}\left[\frac{(N-1)(2N-1)(3N^2-3N-1)}{30} - \frac{(N-1)^2(2N-1)^2}{36}\right] \quad \text{(A1.19)}$$

$$= \frac{W^2(2N-1)(N-1)}{6Ndet(M)}\left[\frac{(3N^2-3N-1)}{5} - \frac{(N-1)(2N-1)}{6}\right] \quad \text{(A1.20)}$$

$$= \frac{W^2(2N-1)(N-1)}{6Ndet(M)}\left[\frac{3N^2-3N-1}{5} - \frac{(2N^2-3N+1)}{6}\right] \tag{A1.21}$$

$$= \frac{W^2(2N-1)(N-1)}{6Ndet(M)}\left[\frac{18N^2-18N-6-10N^2+15N-5}{30}\right] = \frac{W^2(2N-1)(N-1)(8N^2-3N-11)}{180Ndet(M)} \tag{A1.22}$$

$$C(3,3) = \frac{W^2}{Ndet(M)} adjM(3,3) = \frac{W^2}{Ndet(M)}\left[\frac{(N-1)(2N-1)}{6} - \frac{(N-1)^2}{4}\right] = \frac{W^2(N-1)}{2Ndet(M)}\left[\frac{(2N-1)}{3} - \frac{N-1}{2}\right] = \frac{W^2(N-1)(N+1)}{12Ndet(M)} \tag{A1.23}$$

$$C(1,2) = C(2,1) = \frac{W^2}{Ndet(M)} adjM(1,2) = \frac{W^2}{Ndet(M)}\left[\frac{(N-1)^2(2N-1)(3N^2-3N-1)}{60} - \frac{N(N-1)^2(2N-1)(N-1)}{24}\right] \tag{A1.24}$$

$$= \frac{W^2(N-1)^2(2N-1)}{12Ndet(M)}\left[\frac{(3N^2-3N-1)}{5} - \frac{N^2-N}{2}\right] = \frac{W^2(N-1)^2(2N-1)}{12Ndet(M)}\left(\frac{3N^2-3N-1}{5} + \frac{-N^2+N}{2}\right) \tag{A1.25}$$

$$= \frac{W^2(N-1)^2(2N-1)}{12Ndet(M)}\left(\frac{6N^2-6N-2-5N^2+5N}{10}\right) = \frac{W^2(N-1)^2(2N-1)(N^2-N-2)}{120Ndet(M)} \tag{A1.26}$$

$$C(1,3) = C(3,1) = \frac{W^2}{Ndet(M)} adjM(1,3) = \frac{W^2}{Ndet(M)}\left[\frac{N(N-1)^3}{8} - \frac{(N-1)(N-1)(2N-1)^2}{36}\right] = \frac{W^2(N-1)^2}{4Ndet(M)}\left[\frac{N(N-1)}{2} - \frac{(2N-1)^2}{9}\right] \tag{A1.27}$$

$$= \frac{W^2(N-1)^2}{4Ndet(M)}\left[\frac{N^2-N}{2} - \frac{4N^2-4N+1}{9}\right] = \frac{W^2(N-1)^2(9N^2-9N-8N^2+8N-2)}{72Ndet(M)} = \frac{W^2(N-1)^2(N^2-N-2)}{72Ndet(M)} \tag{A1.28}$$

$$C(3,2) = C(2,3) = \frac{W^2}{Ndet(M)} adjM(3,2) = \frac{W^2}{Ndet(M)}\left[\frac{N(N-1)^2}{4} - \frac{(N-1)(N-1)(2N-1)}{12}\right] = \frac{W^2(N-1)^2}{4Ndet(M)}\left[N - \frac{(2N-1)}{3}\right] = \frac{W^2(N-1)^2(N+1)}{12Ndet(M)} \tag{A1.29}$$

Since the $Ndet(M)$ increases as $N^7$ while the terms in brackets increase by no more than $N^6$, we see that each term in the covariance matrix falls with at least the expected 1/N, and by inspection none of the diagonal terms are negative because a quadratic fit requires $N>2$.

If one is interested in the uncertainty of a prediction, it can be obtained by propagating the results forward using the full covariance matrix of the parameter estimates, in strict analogy to the Kalman filter's equation (6.10), except that the process noise Q is zero.

Of course the same results can be achieved using standard statistics packages. MATLAB for example will give you the fitted parameters as p with the command [p,S]=polyfit(x,y,order), apparently using the RMS of the fit residuals as a measure of the phase noise magnitude. The parameter covariances can be extracted with the command cov=inv(S.R)*(inv(S.R))'*(S.normr)^2/S.df. The MATLAB commands were in fact used to verify this appendix's equations. Figure 10 shows how the parameter uncertainties depend on the number of points (N) in the fit.

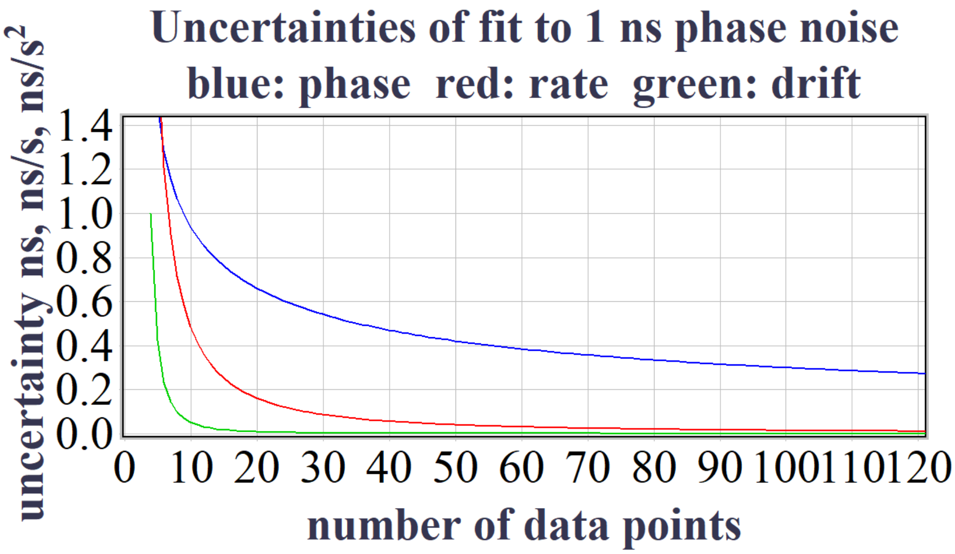



**FIGURE 10.** Expected standard deviation of phase, frequency, and frequency drift parameters in the presence of 1 ns white phase noise

**Appendix II. OPTIMAL MODELLING FOR LINEAR FITS**

For White FM (RW PM), using equation (3.2) $(\sigma^2_{y,RW} = \frac{2\pi^2 k_{-2}}{\tau})$, and equation (75) if Vernotte et al. (2001)

$$TIE^2_{RW}(T_m, T_p) = \frac{4\pi^2 k_{-2} T_m}{15}\left(\frac{9T_p^2}{T_m^2} + \frac{6T_p}{T_m} + 1\right) = \frac{2\,\tau\sigma^2_{y,RW}}{15}\left(\frac{9T_p^2}{T_m} + 6T_p + T_m\right) \quad \text{(A2.1)}$$

Zeroing the derivative with respect to $T_m$,

The maximum is at $0 = \frac{-18T_p^2}{T_m^2}+1$, or $\frac{T_m^2}{T_p^2} = 18$, yielding the ratio r$=\frac{T_m}{T_P} = 3\sqrt{2} = 4.2$ (A2.2)

For IRW noise (RW FM), using equation 76 of Vernotte et al. (2001) and equation 4.3 $(k_{-4} = \frac{3\sigma^2_{y,IRW}}{8\pi^4\tau})$

$$TIE^2_{IRW}(T_m, T_p) = \frac{8\pi^4 k_{-4} T_m^3}{105}\left(35\frac{T_p^3}{T_m^3} + 39\frac{T_p^2}{T_m^2} + 11\frac{T_p}{T_m} + 1\right) = \frac{\sigma^2_{y,IRW}}{35\tau}\left(35T_p^3 + 39T_p^2 T_m + 11T_p T_m^2 + T_m^3\right) \quad \text{(A2.3)}$$

This has no minimum, because without a quadratic term a fit cannot keep up with an IRW. In the absence of other forms of noise, the least bad solution would be to use the last two points. One can see this by inspection, but the standard technique for finding a minimum yields a negative (nonsense) value for $T_m$.

$$0 = 39T_p^2 + 22T_pT_m + 3T_m^2 = 39 + 22r + 3r^2; \quad r = \frac{-22+\sqrt{484-468}}{78} = \frac{-6}{78} = \frac{-1}{13} \quad \text{(A2.4)}$$

In the presence of phase noise, the equations of the first appendix are applicable if one ignores the third column and third row of the design matrix. Equation (A1.10) then becomes

$$C = \frac{W^2}{N} Inv\begin{pmatrix} 1 & \frac{-(N-1)}{2} \\ \frac{-(N-1)}{2} & \frac{(N-1)(2N-1)}{6} \end{pmatrix} = \frac{W^2}{N} Inv(M) \quad \text{(A2.5)}$$

The determinant of the matrix $M$ is $\frac{(N-1)(2N-1)}{6} - \frac{(N-1)^2}{4} = \frac{2N^2-3N+1}{6} - \frac{N^2-2N+1}{4} = \frac{4N^2-6N+2-3N^2+6N-3}{12} = \frac{N^2-1}{12}$ (A2.6)

And so its inverse is $\frac{12}{N^2-1}\begin{pmatrix}\frac{(N-1)(2N-1)}{6} & \frac{(N-1)}{2}\\ \frac{(N-1)}{2} & 1\end{pmatrix}$ and C=$\frac{12W^2}{N(N^2-1)}\begin{pmatrix}\frac{(N-1)(2N-1)}{6} & \frac{(N-1)}{2}\\ \frac{(N-1)}{2} & 1\end{pmatrix}$ (A2.7)

The propagation matrix $\phi = \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$ so the covariance 1 step forward is $\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix} C \begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}$ (A2.8)

Proceeding M steps forward, $\phi^M = \begin{pmatrix}1 & M\\ 0 & 1\end{pmatrix}$, so the covariance matrix of the prediction is given by

$$\phi^M C \phi'^M = \begin{pmatrix}1 & M\\ 0 & 1\end{pmatrix}\frac{12W^2}{N(N^2-1)}\begin{pmatrix}\frac{(N-1)(2N-1)}{6} & \frac{(N-1)}{2}\\ \frac{(N-1)}{2} & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ M & 1\end{pmatrix} \quad \text{(A2.9)}$$

$$= \frac{12W^2}{N(N^2-1)}\begin{pmatrix}1 & M\\ 0 & 1\end{pmatrix}\begin{pmatrix}\frac{(N-1)(2N-1)}{6} + \frac{M(N-1)}{2} & \frac{(N-1)}{2}\\ \frac{(N-1)}{2} + M & 1\end{pmatrix} \quad \text{(A2.10)}$$

$$= \frac{12W^2}{N(N^2-1)}\begin{pmatrix}\frac{(N-1)(2N-1)}{6} + \frac{M(N-1)}{2} + \frac{M(N-1)}{2} + M^2 & \frac{(N-1)}{2} + M\\ \frac{(N-1)}{2} + M & 1\end{pmatrix} \quad \text{(A2.11)}$$

Substituting $N = T_m$ and M$= T_p$, the covariance of the (1,1) component is

$$\frac{12W^2}{N(N^2-1)}\left[\frac{2N^2-3N+1}{6} + M(N-1) + M^2\right] = \frac{12W^2}{T_m(T_m^2-1)}\left[\frac{2T_m^2-3T_m+1}{6} + T_p(T_m-1) + T_p^2\right] \quad \text{(A2.12)}$$

As expected the variance at large $T_p$increases with the square of prediction distance, and falls with the modelling distance.

Figure 11 compares first order and second order Least Squares and Kalman filters, for a particular oscillator model that has RW noise of one unit jumps per epoch, IRW noise of integrated .01 unit jumps per epoch, an initial oscillator drift of $10^{-4}$ units per epoch squared, and a measurement noise of 10 units. A noise of 100 units gives essentially identical results. The results shown, and other simulations, are consistent with a linear fit being more accurate if the oscillator drift is negligible.

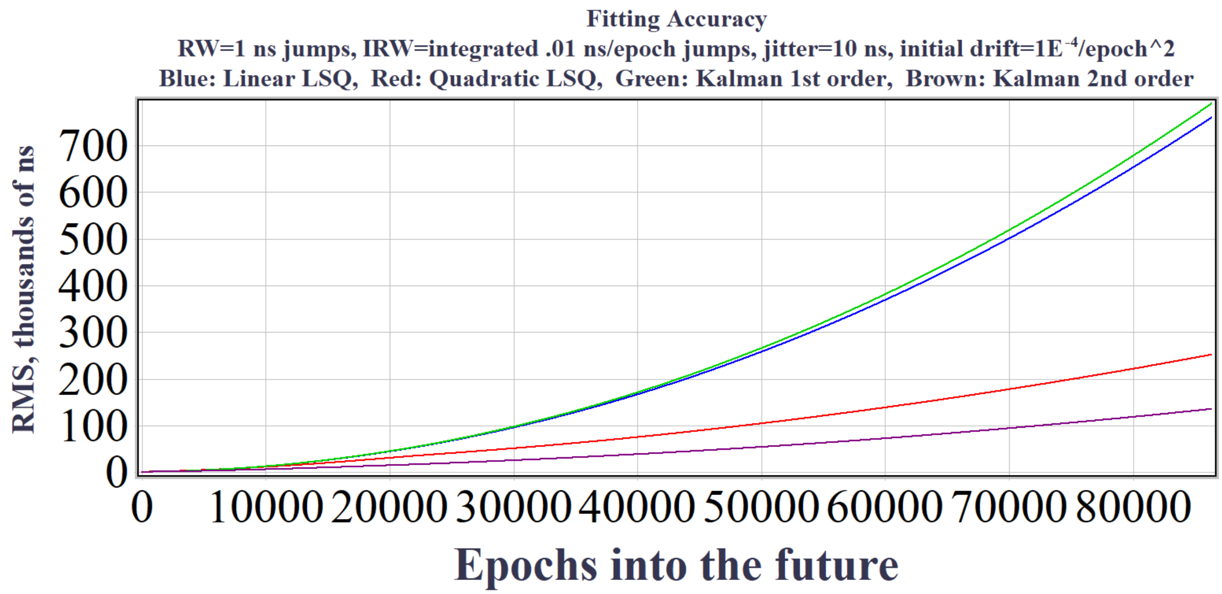



**Figure 11.** Time Errors of an oscillator (as specified immediately above the filter) fitted with linear and quadratic models The middle curve is a superposition of Kalman Filter 1st order and Least Squares 2nd order.