

A KALMAN FILTER CLOCK ALGORITHM FOR USE IN THE PRESENCE OF FLICKER FREQUENCY MODULATION NOISE

J. A. Davis, C. A. Greenhall*, and P. W. Stacey

National Physical Laboratory
Queens Road, Teddington, Middlesex, TW11 0LW, UK

*Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Dr. MS 298-100, Pasadena, CA 91109, USA

Abstract

NPL has developed a Kalman-filter-based clock algorithm for combining measurements from its three active hydrogen masers. The algorithm is designed to produce a near optimal composite when the dominant noise process present is flicker frequency modulation (FFM). The FFM is modelled approximately by a linear combination of Markov noise processes. Each Markov process is included in the Kalman filter and contributes an additional component to its state vector. Both the validity of the model and the effectiveness of adding these additional components to the state vector are examined. The performance of the new algorithm is examined when applied to simulated measurements and also to measurements obtained from NPL's hydrogen masers.

1. INTRODUCTION

The National Physical Laboratory (NPL) maintains and develops the UK's national time scale UTC (NPL). NPL has a requirement to both improve the stability of UTC (NPL) and to steer UTC (NPL) as close to UTC as possible. The improved stability of UTC (NPL) is required for the evaluation of new caesium fountain clocks [1] and optical frequency standards at present under development.

Clock-combining algorithms are of two types: those based on the use of a Kalman filter, for example the GPS composite clock algorithm [2], and those based on a weighted average of the clock data, for example the NIST AT1 algorithm [3]. Most operational clock algorithms used in primary timing laboratories are of the latter type. There have been problems in the implementation of Kalman-filter-based clock algorithms. This is because the physical parameters estimated by the state vector are only partly observable from the measurements made between individual clock pairs. Greenhall [4] has recently overcome these problems through the development of a method described as covariance reduction. The resulting clock algorithm has the potential to produce a composite clock with a stability that is close to optimum for all averaging times.

Kalman-filter-based clock algorithms model the clock noise as a linear combination of White Frequency Modulation (WFM), Random Walk Frequency Modulation (RWFm), and Random Run Frequency Modulation (RRFM). By applying the "three-cornered-hat" technique to NPL's three active hydrogen masers, it is possible to obtain plots of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$ for each of NPL's masers. The results are shown in Figure 1. The plots reveal a complex noise structure, where some of

the noise present may be better described as Flicker Frequency Modulation (FFM), for which the curves have an underlying gradient of zero, in addition to WFM and RWFM, where the underlying gradients are $-1/2$ and $1/2$, respectively.

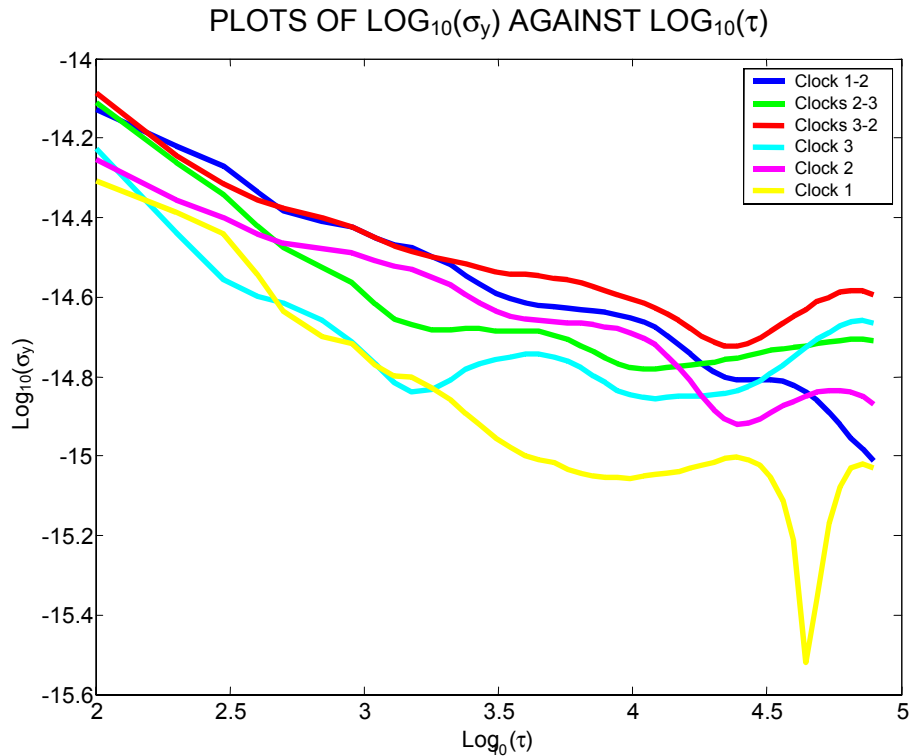


Figure 1. Plots of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$ obtained from measurements made between pairs of NPL's masers and for the individual clocks estimated using a three-cornered hat.

In this paper, we extend the work of Greenhall [4] to develop a clock algorithm that aims to provide close to optimal performance in the presence of FFM. The FFM is modelled as a linear combination of Markov noise processes. Each Markov process is included in the Kalman filter and contributes an additional component to the state vector. The properties of the resulting clock algorithm are examined using both simulated and real clock data. An iterative method has previously been developed to determine the magnitude of noise parameters from a measurement time series [5]. This method is extended and applied to a “three-cornered hat” of hydrogen maser measurements to estimate the required noise parameters.

2. MODELLING FLICKER FREQUENCY MODULATION AS A LINEAR COMBINATION OF MARKOV NOISE PROCESSES

Flicker Frequency Modulation (FFM) may be considered to be a power-law noise process for which the spectral density function of the phase measurements varies as $S_x(f) = \sigma^2 / |2\pi f|^3$, where σ^2 is the variance of an underlying white noise process and f is the Fourier frequency [6]. Recently, FFM has been described in terms of fractionally integrated noise processes [6, 7]. In this work, we adopt an older approach initially studied by Mandelbrot [8], where FFM is described as a linear combination of

Markov noise processes. Let $\{x_i; i = 1, \dots, m\}$ be a time series of measurements of x made at uniformly spaced times $\{t_i; i = 1, \dots, m\}$. A Markov noise process takes the following form:

$$x_{i+1} = kx_i + e_i, \quad (1)$$

where k is a constant for an individual Markov process, e_i, e_j are independent random errors, such that $\text{cov}(e_i, e_j) = 0, i \neq j$, with normal distribution, zero expectation $E(e) = 0$, and variance σ_M^2 . The variance of x_i as i approaches ∞ is given by U , where:

$$U = \frac{\sigma_M^2}{(1 - k^2)}. \quad (2)$$

Assuming $|k^2| < 1$ then the autocorrelation function $\rho(l)$ is given by:

$$\rho(l) = \exp(-Rl) \quad (3)$$

where l is the lag and $k = \exp(-R\tau_0)$, τ_0 is the time interval between successive points. In the work undertaken here, FFM is modelled as a linear combination of h Markov processes, where:

$$y_{f,i} = \sum_{j=1}^h y_{j,i}, \quad \text{and} \quad U_j = \text{var}(y_j) \quad (4)$$

$y_{j,i}$ is the i^{th} element of the j^{th} Markov process, and $y_{f,i}$ is the i^{th} element of the resulting approximation to FFM frequency measurements. We choose the magnitude of the Markov components so that the variances of the component $U_j = \text{var}(y_j)$ are equal. This will result in time series where plots of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$ will have a gradient of approximately zero.

The properties of FFM modelled in terms of a linear combination of Markov processes are shown in Figures 2, 3, and 4 below. Four Markov processes are used to model FFM with $R_{j+1} = R_j/8$ and $R_1 = 0.75$, $\tau_0 = 1$. Curves of a typical FFM simulation and their component Markov processes are shown in Figure 2. Plots of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$ are shown in Figure 3, based on theory (continuous lines) and simulated data (marks), and including results for the component Markov processes.

When simulated data are used, the initial values of the Markov time series $y_{j,1}$ are generated so that

$$\text{var}(y_{j,1}) = \frac{\sigma_M^2}{(1 - k_j^2)}. \quad (5)$$

This results in a stationary time series.

A plot of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$ for an individual Markov process possesses a single maximum at $\tau \approx 1.81R$. Plots of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$ obtained from the sum of several series of Markov processes summed as described in equation (4) results in a plot of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$ that exhibits a ‘‘ripple’’ effect; see Figure 4. The FFM may be constructed so that the underlying gradient approximates zero. It is relatively simple to construct a linear combination of Markov processes such that a plot of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$ has a gradient between $-1/2$ and $1/2$, and so models any power-law noise process intermediate in characteristics between WFM and RWFM.

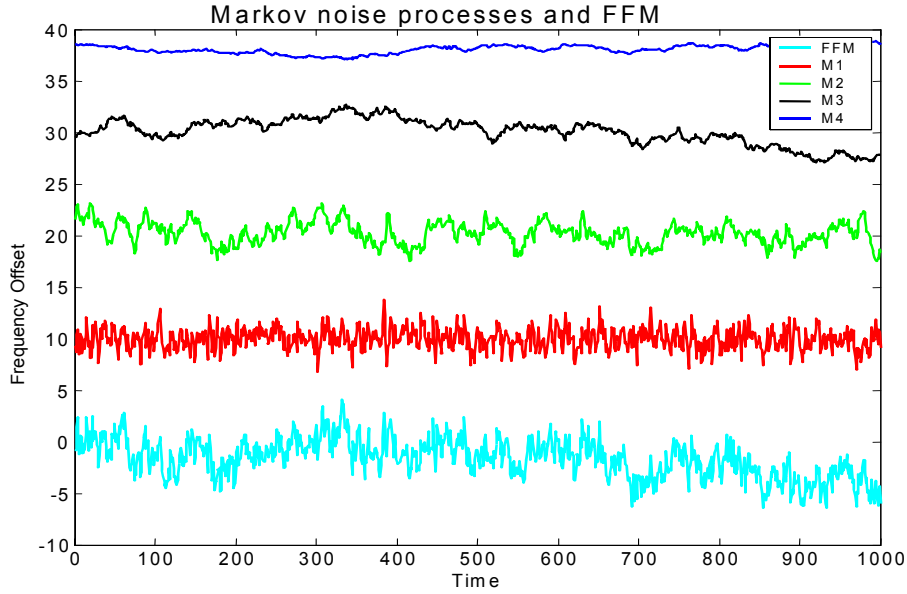


Figure 2. An example of a simulated FFM noise process and its Markov components (M1 to M4), offset by 10, 20 30 and 40 units, respectively, for clarity.

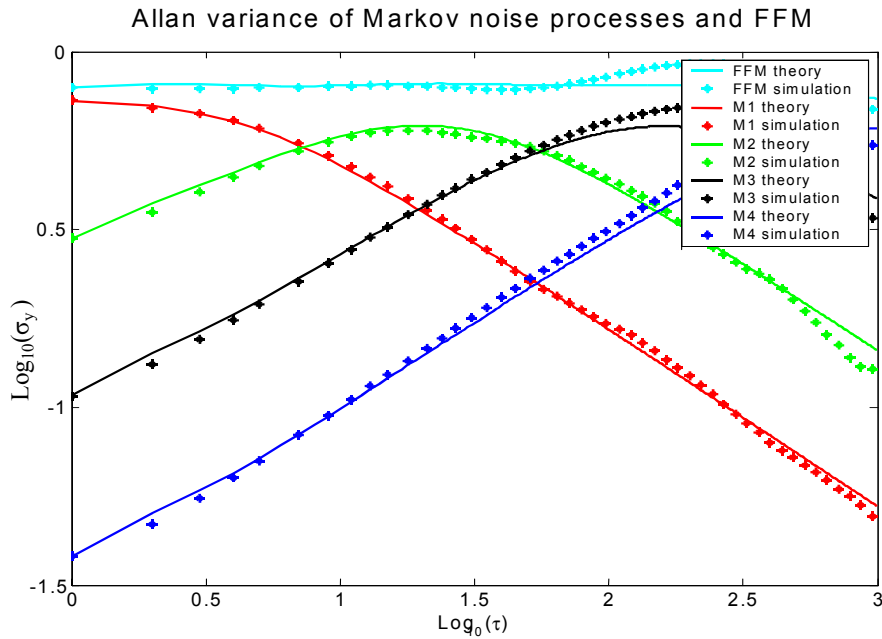


Figure 3. Plots of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$ obtained from both simulation and theory for a FFM noise process and its Markov components (M1 to M4).

The effect of increasing the “spacing ratio” $s = R_j / R_{j+1}$ between the Markov processes from $R_{j+1} = R_j / 8$ to $R_{j+1} = R_j / 16$ is shown in Figure 4. The magnitude and period of the “ripple” observed on the plots of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$ is significantly increased. There is, however, a price to be paid in using a small spacing ratio in terms of additional computational time. The “ripple”

is not clearly observed on the FFM curves of Figure 3 due to the y-axis scale interval being much larger.

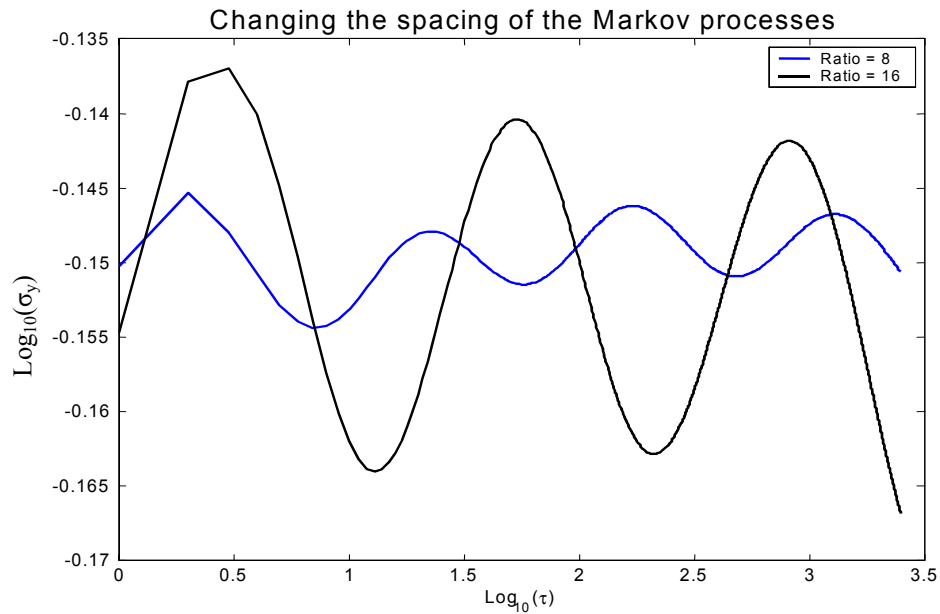


Figure 4. Plots of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$ obtained from theory when modelling FFM using Markov noise processes with spacing ratios of 8 and 16 respectively. A bias has been added to the (ratio = 16) curve for clarity.

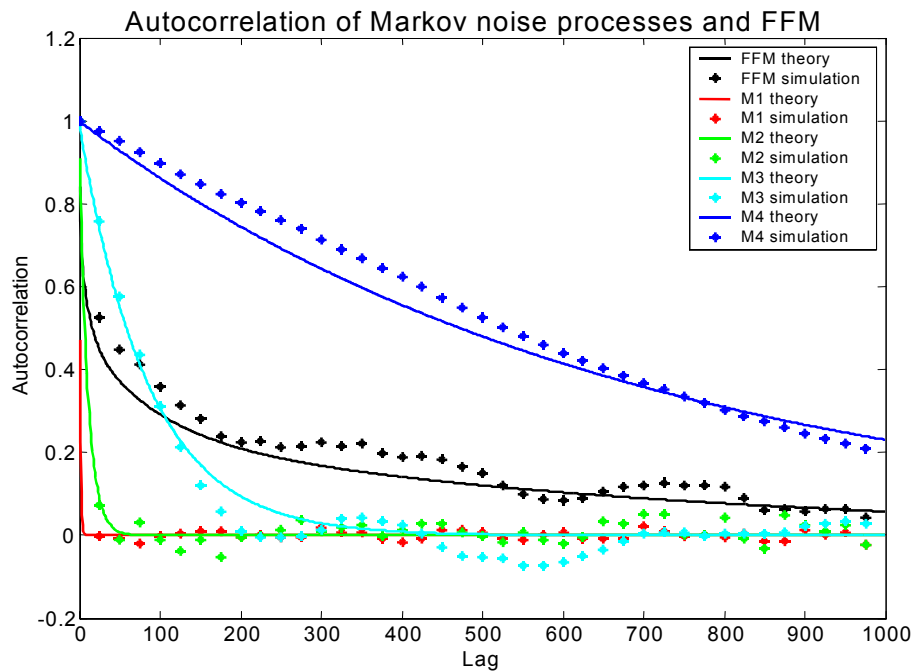


Figure 5. Plots of the autocorrelation function $\rho(l)$ obtained as a function of lag l derived from theory and by simulation for FFM noise processes and for the four Markov processes (M1 to M4) used in its construction. 20,000 points were used in the simulation.

Figure 5 shows plots of the autocorrelation function, obtained from both theory and simulation. The autocorrelation of the FFM decays much slower than would be expected from a single exponential decay; this “long-memory” characteristic of the FFM is well known [7]. From the above plots it is reasonable to assume that a good approximation to FFM has been achieved. It must also be remembered that it is only possible to generate an autocorrelation function because we are modelling FFM with a finite number of Markov processes; otherwise, the FFM noise process would not be stationary.

3. DESIGN OF THE KALMAN FILTER

The development of Kalman filter clock algorithms have been described previously [2,4]. In this work, each Markov process contributes an extra component to the state vector. When using three clocks where the FFM for each clock is modelled by four Markov processes, the state vector is given by:

$$\hat{x} = (x_1, y_1, z_1, m_{11} \cdots m_{14}, x_2, y_2, z_2, m_{21} \cdots m_{24}, x_3, y_3, z_3, m_{31} \cdots m_{34})^T \quad (6)$$

where x_i , y_i and z_i are the time offset, frequency offset, and linear frequency drift offset between the clock i and an ideal clock. m_{ij} is the frequency offset between the j^{th} Markov component of the FFM of clock i . The composite clock is obtained as the Kalman filter’s estimate of this “ideal” clock.

The component Φ_{ii} of the state propagation matrix for clock i is given by:

$$\Phi_{ii} = \begin{pmatrix} 1 & \tau_0 & \frac{\tau_0^2}{2} & \frac{(1 - \exp(-R_1 \tau_0))}{R_1} & \dots & \dots & \frac{(1 - \exp(-R_4 \tau_0))}{R_4} \\ & 1 & \tau_0 & & & & \\ & & 1 & & & & \\ & & & \exp(-R_1 \tau_0) & & & \\ & & & & \ddots & & \\ & & & & & \ddots & \\ & & & & & & \exp(-R_4 \tau_0) \end{pmatrix} \quad (7)$$

where $\Phi = \text{diag}(\Phi^1, \dots, \Phi^n)$, there being n clocks in the ensemble. The top left 3-by-3 sub-matrix is well-known and the additional terms represent the Markov components [9].

The component Q_{ii} of the process covariance matrix Q for clock i is given by:

$$Q_{ii} = \begin{pmatrix} Q_{ii1} & \sigma_{RWFEM}^2 \tau_0^2 / 2 & 0 & \tau_0^2 a_{12}(R_1 \tau_0) \sigma_{MC1}^2 & \dots & \dots & \tau_0^2 a_{12}(R_4 \tau_0) \sigma_{MC4}^2 \\ \sigma_{RWFEM}^2 \tau_0^2 / 2 & \sigma_{RWFEM}^2 \tau_0 & 0 & & & & \\ 0 & 0 & 0 & & & & \\ \tau_0^2 a_{12}(R_1 \tau_0) \sigma_{MC1}^2 & & & \tau_0 a_{22}(R_1 \tau_0) \sigma_{MC1}^2 & & & \\ & & & & \ddots & & \\ & & & & & \ddots & \\ \tau_0^2 a_{12}(R_4 \tau_0) \sigma_{MC4}^2 & & & & & & \tau_0 a_{22}(R_4 \tau_0) \sigma_{MC4}^2 \end{pmatrix} \quad (8a)$$

where $S = \text{diag}(S_{11}, \dots, S_{mm})$. S is formed from an identity matrix where the elements corresponding to the x (phase) component of the state vector are set to zero. The covariance reduction technique is applied to all work undertaken in this paper.

5. EVALUATING THE PERFORMANCE OF THE KALMAN FILTER USING SIMULATED DATA

The performance of the Kalman filter for simulated data is shown in Figures 6 to 9 below. Three clocks are used in each simulation. In Figures 6 and 7, clocks 1 and 3 exhibit only WFM noise and clock 2 FFM. In Figures 8 and 9, clocks 1 and 3 exhibit RWFM and clock 3 FFM, the FFM being modelled using four Markov processes. The clock phase outputs are shown in Figures 6 and 8. Figures 7 and 9 show plots of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$.

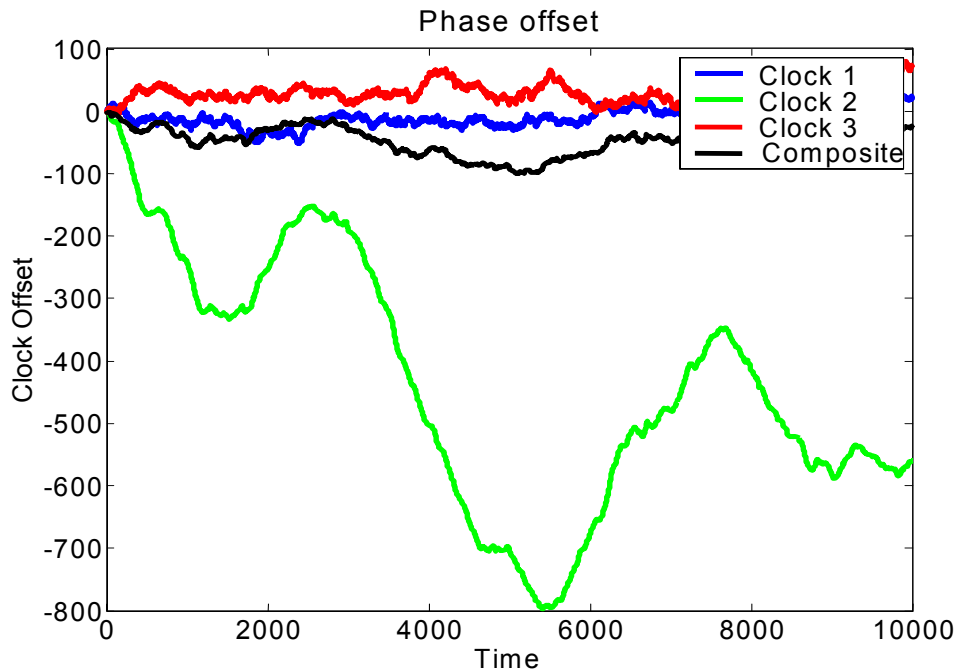


Figure 6. Phase offsets of the three clocks used in the simulation. Clocks 1 and 3 exhibit WFM and clock 2 FFM. The resulting composite is shown.

Figures 6 to 9 show the performance of the individual clocks and the resulting composite. At almost all averaging times, the stability of the composite clock is equal to or better than the stability of the most stable clock (Figures 7 and 9). Data from the first 5% of the time series are not used in the computation of the plots of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$ so as to give the Kalman Filter time to “settle down.” At any averaging time (τ), it is possible to form an optimally weighted mean of the clocks that will minimize the AVAR of the resulting composite; these weights are given by w_i , where:

$$w_i = \frac{w'_i}{\sum_{k=1}^n w'_k} \quad \text{and} \quad w'_i = 1/\sigma_y^2(\tau) \quad (12)$$

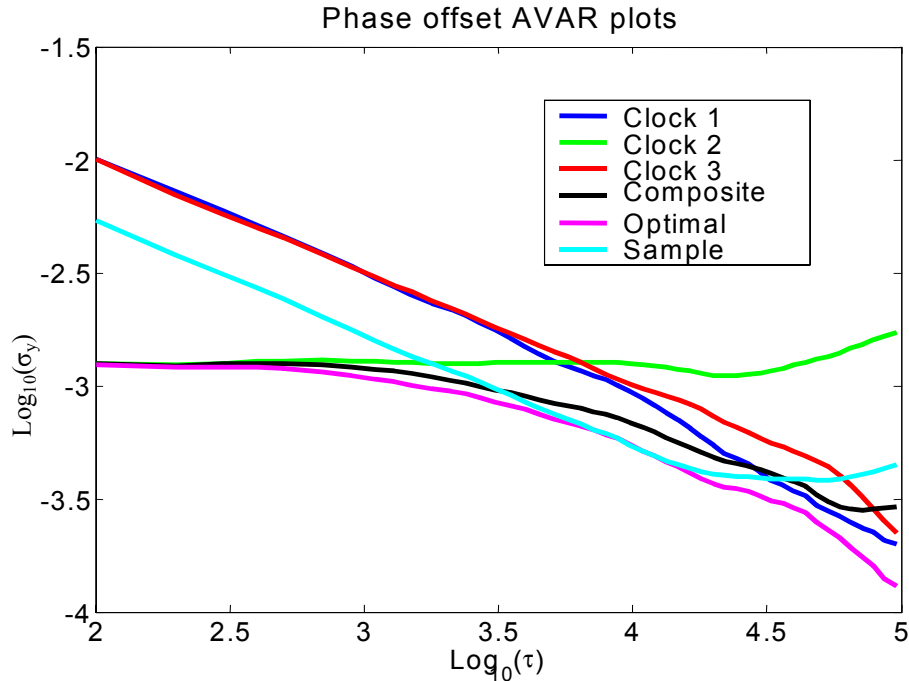


Figure 7. Plots of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$ obtained from the individual clocks and from the composite obtained from the Kalman filter. The final two curves show the σ_y obtainable from an optimal weighted mean of the clock measurements and those obtained from a fixed weighting.

The resulting plot of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$ is shown as the “optimal” curve in Figures 7 and 9. It is important to note that, for this curve, different clock weights have been used at each τ . By comparing the composite curve against the optimal curve, we conclude that the composite curve approaches the optimal curve at all averaging times, demonstrating that the addition of Markov components to the state vector may result in a clock algorithm that is close to optimal at all averaging times. Finally, we take a single clock weighting optimal at $\tau = 10,000\text{s}$. By comparing this “sample” curve with the composite curve, we are able to observe the difference in performance between the Kalman-filter-based algorithm and a simple weighted mean algorithm. The Kalman-filter-based clock algorithm performs close to optimal at all averaging times, while the simple weighted mean clock algorithm is optimal at a single averaging time, and performs sub-optimally at all other averaging times. It must be remembered, however, that the superior performance of the Kalman-filter-based clock algorithm may only be obtained if the noise processes do indeed approximate a linear combination of WFM, FFM, and RWFM noise processes and requires that good estimates for the corresponding noise parameters are available.

6. OBSERVING THE MARKOV COMPONENTS OF THE STATE VECTOR

The successful operation of this clock algorithm is due the Kalman filter’s ability to correctly estimate the Markov components of the state vector. Figure 10 shows the simulated data for two of the Markov processes M3 and M4 and the resulting state vector estimate. There is a high degree of correlation between the two M4 curves and, to a lesser extent, between the two M3 curves. These results suggest that the Kalman filter is correctly estimating the Markov components of the state vector.

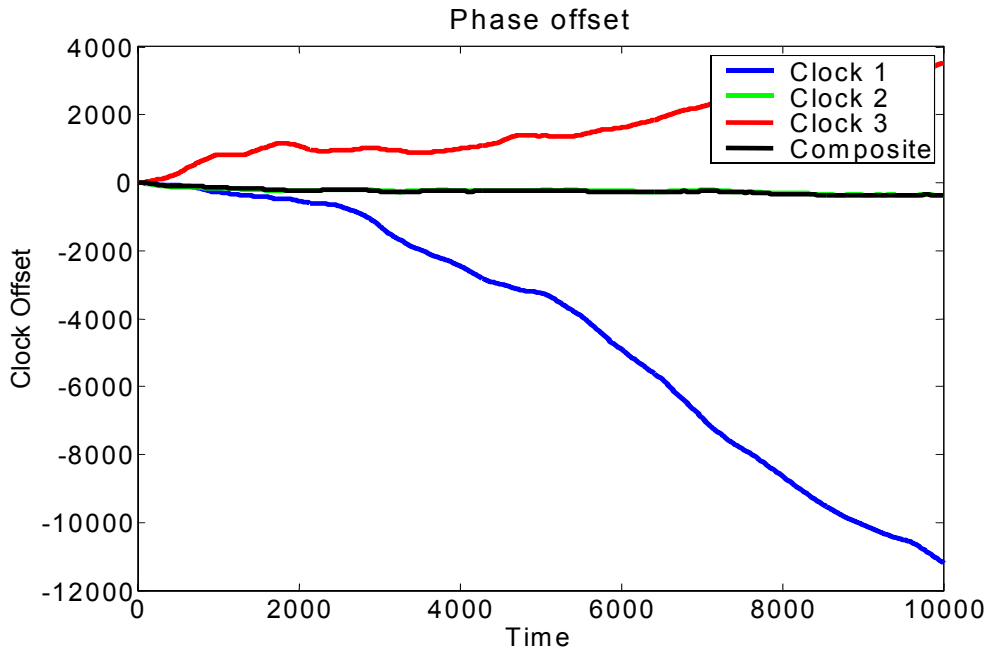


Figure 8. Phase offsets of the three clocks used in the simulation. Clocks 1 and 3 exhibit RWFM and clock 2 FFM. The resulting composite is shown.

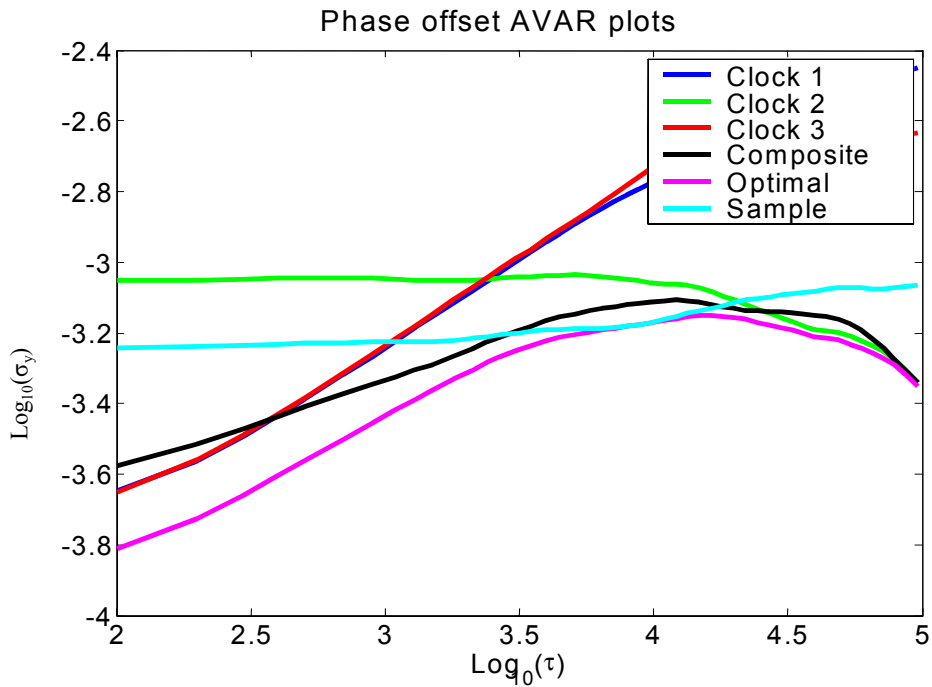


Figure 9. Plots of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$ obtained from the individual clocks and from the composite obtained from the Kalman filter. The final two curves show the σ_y obtainable from an optimal weighted mean of the clock measurements and those obtained from a fixed weighting.

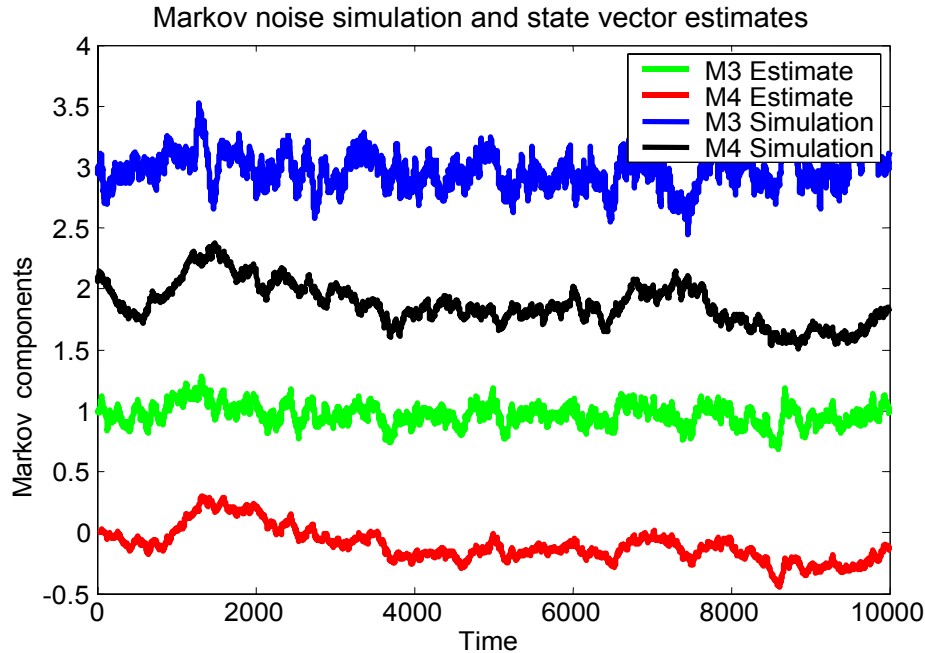


Figure 10 shows the Markov state vector estimates M3 and M4 and corresponding simulated Markov noise processes. Integer offsets of between zero and three units have been added for clarity.

7. ESTIMATING THE CLOCK NOISE PARAMETERS

To operate this Kalman filter clock algorithm using real clock data, it is necessary to first estimate the magnitude of the noise parameters σ_{WFM}^2 , σ_{RWFM}^2 , and σ_{MC}^2 . A method to estimate the magnitude of noise parameters is described in [5]. Estimates of a scaled AVAR are obtained from the measurement data set at averaging times $\tau = 2^n \tau_0$, where n is a positive integer and τ_0 is the minimum spacing of the data set. An iterative least squares process is then applied to produce a near optimal estimate of the noise parameters in the presence of highly correlated AVAR estimates.

In this paper, the use of a scaled AVAR has been replaced with a $D(4)$ Daubechies wavelet variance [11]. The $D(4)$ wavelet is itself a second-difference variance. In this work, the wavelet variances are applied to frequency measurements, and are used to resolve σ_{WFM}^2 , σ_{FFM}^2 , and σ_{RWFM}^2 noise parameters. σ_{FFM}^2 represents a FFM noise process that is approximated as the sum of n Markov noise processes, each with variance $\sigma_M^2 / (1 - k_n^2)$ as described in equation (2). The advantage of using

the wavelet variance is that, in the presence of power-law noise processes, the correlation between the estimates of individual variances is minimal and may be neglected. This significantly simplifies the analysis process, in particular where there are a relatively large number of variance estimates and measurements. It must be remembered that we are making measurements between pairs of clocks; the noise parameter resolution technique must be combined with a three-cornered hat. In the case of only three clocks being present, we construct the following least-squares problem:

$$\begin{pmatrix} d_{1AB} \\ \vdots \\ d_{nAB} \\ d_{1AC} \\ \vdots \\ d_{nAC} \\ d_{1BC} \\ \vdots \\ d_{nBC} \end{pmatrix} = \begin{pmatrix} h_{1AB,AWFM} & \cdots & h_{1AB,CRWFM} \\ \vdots & & \vdots \\ h_{nBC,AWFM} & \cdots & h_{nBC,RWFM} \end{pmatrix} \begin{pmatrix} \sigma_{AWFM}^2 \\ \sigma_{AFFM}^2 \\ \sigma_{ARWFM}^2 \\ \sigma_{BWFM}^2 \\ \sigma_{BFFM}^2 \\ \sigma_{BRWFM}^2 \\ \sigma_{CWFM}^2 \\ \sigma_{CFFM}^2 \\ \sigma_{CRWFM}^2 \end{pmatrix} \quad (13a)$$

$$D = H\sigma \quad (13b)$$

where d_{ijk} is the i^{th} D (4) wavelet variance estimated from measurements of the (j-k) clock pair. The elements of the design matrix H may be simply computed for wavelets of small width and obtained for wavelets of larger width by extrapolation. The variance of the wavelet variance $\text{var}(d_j)$ may be estimated as:

$$\text{var}(d_j) = 2d_j^2 / n \quad (14)$$

where n is the number of samples in the wavelet estimate; these are assumed to be uncorrelated.

An example of the estimate of noise parameters is shown in Figure 11 below. The first three curves show plots of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$ obtained for NPL's three active hydrogen masers using a three-cornered hat. The same data set is then used to estimate the σ_{WFM}^2 , σ_{FFM}^2 , and σ_{RWFM}^2 noise parameters for each clock. These noise parameters are themselves used to determine theoretical curves of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$; these are shown in the second three curves. The two curves obtained for each clock are in broad agreement.

8. FIRST RESULTS USING REAL CLOCK MEASUREMENTS

The Kalman filter clock algorithm has been applied to 10 days of data from NPL's three active hydrogen masers. The resulting plots of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$ obtained from (Composite – Individual Clock) estimates are shown in Figure 12. Because we are unable to compare against an “ideal” clock, the stability of (Composite – Individual Clock) estimates are examined. The results are encouraging; however, it must be remembered that there will be a high degree of correlation between the most stable clocks used in the algorithm and the composite. Future work will include comparing the stability of (Composite – Individual Clock) estimates, obtained from theory and those obtained from applying the clock algorithm to real data.

In the longer term, it is hoped to compare the output of the clock algorithm against NPL's cesium fountain, as well as against other UTC (k) time scales using two-way time transfer measurements, and also to develop methods to estimate the stability of the resulting composite.

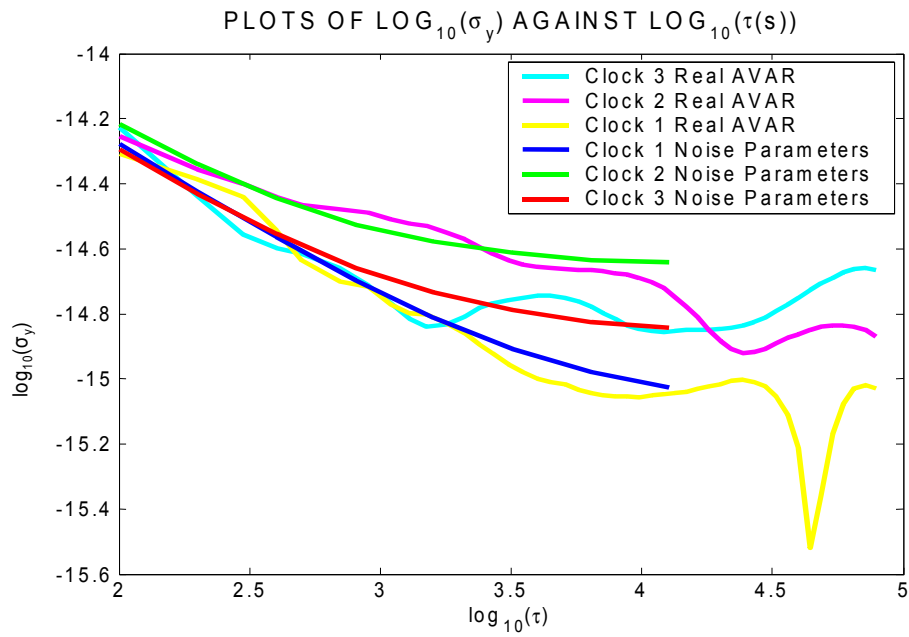


Figure 11. Plots of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$ obtained for NPL's three active hydrogen masers using a three-cornered hat, and obtained using the estimated noise parameters.

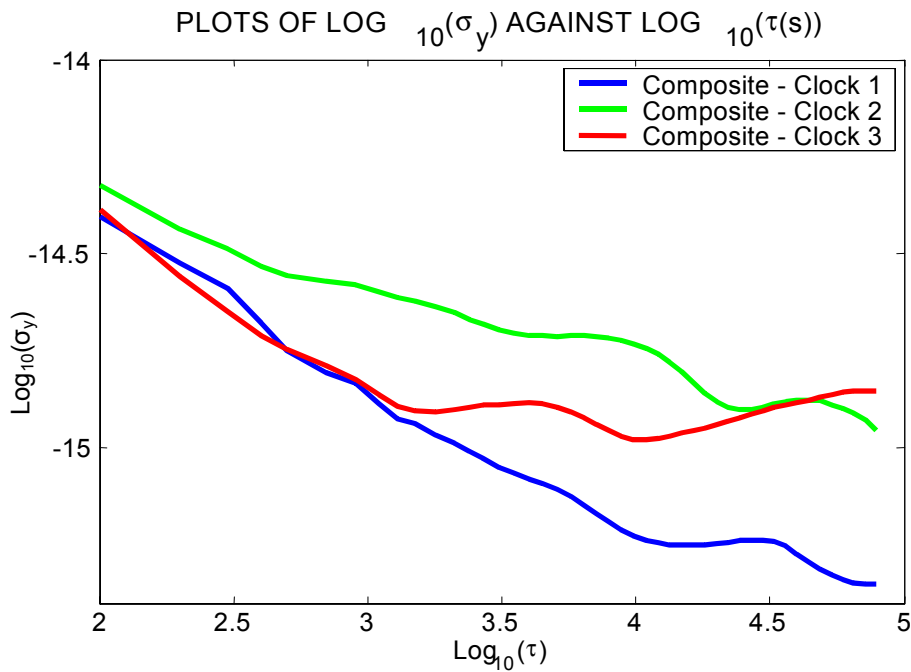


Figure 12. Plots of $\text{Log}_{10}(\sigma_y)$ against $\text{Log}_{10}(\tau)$ obtained from (Composite – Individual Clock) estimates obtained from applying the Kalman filter clock algorithm to NPL's three active hydrogen masers.

9 CONCLUSIONS

We have shown that Flicker Frequency Modulation noise may be described approximately as a linear combination of Markov noise processes. This description of FFM has been successfully included in a Kalman-filter-based clock algorithm. The result is a clock algorithm that will produce close to optimal performance at all averaging times, when operating in the presence of FFM. The algorithm has been tested using simulated data. First results have been obtained using real clock data.

REFERENCES

- [1] K. Szymaniec, W. Chalupczak, and D. Henderson, 2003, “*Initial evaluation of the NPL caesium fountain frequency standard*,” in Proceedings of the 2003 International Frequency Control Symposium & PDA Exhibition Jointly with the 17th European Frequency and Time Forum (EFTF), 5-8 May 2003, Tampa, Florida, USA (IEEE Publication 03CH37409C), pp. 112-114.
- [2] K. R. Brown, 1991, “*The theory of the GPS composite clock*,” in Proceedings of ION GPS-91 (Institute of Navigation, Alexandria, Virginia), pp. 223-241.
- [3] M. A. Weiss, D. W. Allan, and T. K. Pepler, 1989, “*A study of the NBS time scale algorithm*,” **IEEE Transactions on Instrumentation and Measurement**, **IM-38**, 631-643.
- [4] C. A. Greenhall, 2003, “*Forming stable timescales from the Jones-Tryon Kalman filter*,” **Metrologia**, **40**, S335-S341.
- [5] P. M. Harris, J. A. Davis, M. G. Cox, and S. L. Shemar, 2003, “*Least-squares analysis of time series data and its application to two-way satellite time and frequency transfer measurements*,” **Metrologia**, **40**, S342-S347.
- [6] D. B. Percival, 2003, “*Stochastic models and statistical analysis for clock noise*,” **Metrologia**, **40**, S289-S304.
- [7] L. S. Schmidt, 2003, “*Atomic clock models using fractionally integrated noise processes*,” **Metrologia**, **40**, S305-S311.
- [8] B. B. Mandelbrot, 1971, “*A fast fractional gaussian noise generator*,” **Water Resources Research**, **7**, no. 3, pp. 543-553.
- [9] R. G. Brown and P. Y. C. Hwang, 1997, **Introduction to Random Signals and Applied Kalman Filtering**, 3rd edition (John Wiley & Sons, New York), pp. 200-202.
- [10] S. T. Hutsell, 1996, “*Relating the Hadamard variance to MSC Kalman filter clock estimation*,” in Proceedings of the 27th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, 29 November-1 December 1995, San Diego, California, USA (NASA Conference Publication 3334), pp. 291-301.
- [11] D. B. Percival and A. T. Walden, 2000, **Wavelet Methods for Time Series Analysis** (Cambridge University Press), p. 59.

QUESTIONS AND ANSWERS

DON PERCIVAL (University of Washington): In statistical terms, what you are doing is looking at a linear combination of autoregressive processes of order one. Another approach which has been taken in the literature to try to approximate flicker noise processes is to look at higher-order autoregressive processes. Instead of using n th first-order processes, you could think about an n th fourth-order autoregressive process. That can be embedded in the Kalman filter very simply. Is there some advantage to using the linear combination of first-order processes, as opposed to the n th fourth-order autoregressive process?

JOHN DAVIS: To be quite honest, it is just the simplest way I could see how to do it, adding up the simple Markov processes. We have not looked at anything else. If we took a view after, I could certainly try to do that. That might be interesting.

