DIFFERENTIAL SAMPLING FOR FAST FREQUENCY ACQUISITION VIA ADAPTIVE EXTENDED LEAST SQUARES ALGORITHM

Rajendra Kumar
Jet Propulsion Laboratory
California Institute of Technology, Pasadena, CA 91109
Department of Electrical Engineering
California State University, Long Beach, CA 90804

ABSTRACT

This paper presents a differential signal model along with appropriate sampling techniques for least squares estimation of the frequency and frequency derivatives and possibly the phase and amplitude of a sinusoid received in the presence of noise. The proposed algorithm is recursive in measurements and thus the computational requirement increases only linearly with the number of measurements.

The dimension of the state vector in the proposed algorithm does not depend upon the number of measurements and is quite small, typically around four. This is an advantage when compared to previous algorithms wherein the dimension of the state vector increases monotonically with the product of the frequency uncertainty and the observation period. Such a computational simplification may possibly result in some loss of optimality. However, by applying the sampling techniques of the paper such a possible loss in optimality can be made small.

1. INTRODUCTION

The problem of estimating the phase and frequency of a received quasi-sinusoidal signal has been investigated by many researchers, see for example references [1-9] and their references. This has been obviously due to diverse and important engineering applications where such problems arise.

In purely theoretical terms the optimal estimates are simply the maximum-a-priori (MAP) estimates of the unknown parameters or signals. However, as the measurements are nonlinear functions of the unknown parameters, close-form solutions of the estimation problem are infinite-dimensional and too complex to implement [1]. Therefore, the researchers in the field have been motivated to apply either known or novel approximation
techniques so as to arrive at a suboptimal estimator which may be feasible to implement. The application of the extended Kalman filter (EKF) resulting in a digital phase locked loop (DPLL) [2,3] and the application of the Fast Fourier Transform (FFT) techniques [4-6] are two such approximations previously studied in the literature.

However, such approximations are based upon some specific assumptions about the signal model. For example, in the EKF approach, it is assumed that the one-step ahead prediction error is small and thus in the approximation of the nonlinear measurement function by its Taylor series expansion around the predicted estimates, all but the linear terms are negligible. In case of the frequency and phase estimation problem this in turn implies assumptions in terms of sampling rate, the uncertainty in terms of the frequency and its derivatives, signal-to-noise ratio etc. However, even with the above assumptions, the one-step-ahead prediction error (in the phase estimate) may be high in the initial estimation period (unless the frequency offset is negligible) and thus during this phase the performance can be drastically different from the intended optimal performance. In case of the FFT technique, a major assumption is that of constant frequency over the estimation period. Also computational limitations may result in a significant quantization error in the frequency estimate. In [4] a secondary algorithm has also been proposed for the interpolation of such quantized estimates. The FFT algorithm being nonrecursive in measurements, it is difficult to adapt it to time-varying situations and to obtain extensions so as to estimate the frequency derivatives, if these exist.

Our approach to the problem is to consider higher order approximations to the measurement function rather than just the linear one as in case of the extended Kalman filter. With sufficient number of terms retained in such an approximation, the resulting algorithm can be made as close to an optimum one as desired. Following such an approach, a least squares based algorithm has been proposed [7] which provides fast estimates of the frequency, phase and possibly the amplitude of the received signal. The complexity of such an algorithm increases with the product of frequency uncertainty and the observation period. Where this product is high, one could apply a suboptimum procedure [7] so as to reduce the computational requirements.

In this paper we show that with a suitable transformation of the signal model and with the use of a differential and cyclic sampling technique, a least squares algorithm can be derived which requires significantly fewer computations than the previously proposed algorithm [7]. In fact, for many situations, the computational requirements may be only marginally higher than for a DPLL. The proposed algorithm being recursive in the number of observations, it is capable of adapting to time-varying situations. Also, it is possible to explicitly estimate various frequency derivatives if these are significant, rather than indirectly inferring these from the slowly varying frequency estimates. From the
simulations, it is also shown that the algorithm exhibits an exponential convergence of the estimates to the neighborhood of the true parameters.

2. THE SIGNAL MODEL

Consider first a simpler problem of estimating an unknown frequency $\omega_d$ from the measurements $y(t)$, $z(t)$ in (1) made in the presence of zero mean additive white Gaussian noise $n_i(t), n_q(t)$.

\[
y(t) = A \sin(\omega_d t + \phi) + n_i(t) \quad ; 0 \leq t \leq T_0
\]

\[
z(t) = A \cos(\omega_d t + \phi) + n_q(t)
\]

Here $\{y(t), z(t)\}$ represents the in-phase and quadrature components of a received signal $s(t)$ and $\omega_d$ represents the frequency offset between the frequencies of the carrier reference signal $r(t)$ and the received signal $s(t)$. In the first instance we consider the uniformly sampled versions of the signals $y(t), z(t)$ given by

\[
y(k) = A \sin(\omega_d t_k + \phi) + n_i(k) \quad ; k = 0, 1, ..., N
\]

\[
z(k) = A \cos(\omega_d t_k + \phi) + n_q(k)
\]

with $n_i(k)$ and $n_q(k)$ denoting the sampled versions of the quadrature components of $n(t)$ with variance $\sigma_i^2$. In order to obtain a differential version of (2) we expand $\sin(\omega_d t + \phi)$ in a Taylor series around $t_{k-1}$, i.e.,

\[
\sin(\omega_d t_k + \phi) = \sin(\omega_d t_{k-1} + \phi) + \omega_d \Delta t_k \cos(\omega_d t_{k-1} + \phi) - \frac{1}{2!} \omega_d^2 (\Delta t_k)^2 \sin(\omega_d t_{k-1} + \phi) - ... \Delta t_k \triangleq t_k - t_{k-1}, k = 1, 2, ..., N
\]

and obtain a similar expression for $\cos(\omega_d t_k + \phi)$. With the substitution of (3) in (2), an equivalent signal model in the following differential form is easily obtained.

\[
y_d(k) \triangleq y(k) - y(k-1) = \omega_d \Delta t_k z(k-1) - \frac{1}{2!} \omega_d^2 (\Delta t_k)^2 y(k-1) + ... + \xi_i(k)
\]

\[
z_d(k) \triangleq z(k) - z(k-1) = -\omega_d \Delta t_k y(k-1) - \frac{1}{2!} \omega_d^2 (\Delta t_k)^2 z(k-1) + ... + \xi_q(k)
\]

where

\[
\xi_i(k) \triangleq n_i(k) - n_i(k-1) - \omega_d \Delta t_k n_q(k-1) - \frac{1}{2!} (\omega_d \Delta t_k)^2 n_i(k-1) + ...
\]

\[
\xi_q(k) \triangleq n_q(k) - n_q(k-1) + \omega_d \Delta t_k n_i(k-1) + \frac{1}{2!} (\omega_d \Delta t_k)^2 n_q(k-1) + ...
\]
The measurement model (4) can now be approximated by

\[ v(k) = \Theta' x(k) + \xi(k) \quad ; k = 1, 2, ..., 2N \]  

(6)

where

\[
\begin{align*}
\Theta' &= [\omega_d, \frac{1}{\Delta t_1^{2}}, ..., \frac{1}{\Delta t_1^n}] \\
z'(2k-1) &= [\Delta t_k x(k-1) - (\Delta t_k)^2 y(k-1) \ldots (\Delta t_k)^n z(k-1)] \\
z'(2k) &= [-\Delta t_k y(k-1) - (\Delta t_k)^2 x(k-1) \ldots (\Delta t_k)^n y(k-1)] \\
v(2k-1) &= y_d(k), \quad v(2k) = z_d(k), \quad n(2k-1) = n_i(k), \quad n(2k) = n_q(k) \\
\xi(2k-1) &= \xi_i(k), \quad \xi(2k) = \xi_q(k) \quad ; k = 1, 2, ..., N
\end{align*}
\]

(7)

In the above ‘ denotes transpose and the terms of order smaller than \((\omega_d \Delta t_k)^n/n!\) have been ignored. For example, if uniform sampling is used with \(\Delta t_k = T_s\) then so as to satisfy the Nyquist criteria \(\omega_d \Delta t_k < \pi\) and in this case it may be sufficient to have only 5 or 6 terms in the model. For faster sampling rate, the number of terms may be even smaller.

The signal model (6,7) above is very similar to the one explored in [7] with a few significant differences. First, whereas the state vector \(x(k)\) in (7) has elements which are functions of \(t\) in [7] the elements of \(x(k)\) are powers of \(t\). This implies that whereas the model dimension \(n\) in [7] increases with the number of observations \(N\) to keep the approximation error small, here \(n\) does not depend upon \(N\). Second, whereas the algorithm of [7] also estimates simultaneously the phase and possibly the amplitude of the received waveform, here we are mainly interested in the frequency estimation. However, as discussed later in the paper, it is possible to also estimate the amplitude and phase by a suitable augmentation of the algorithm. Another main difference between the two signal models is that the additive noise appearing in (7) is not white.

3. PARAMETER ESTIMATION

The parameter vector \(\Theta\) in (6) can be estimated by a standard least squares algorithm in a recursive or nonrecursive form. In the nonrecursive form, the estimate of \(\Theta\) on the basis of measurements \(v(k), k = 2N\) is given by

\[
\hat{\Theta}(N) = \left\{ \sum_{j=1}^{N} x(j)z'(j)\lambda^{N-j} \right\}^{-1} \left\{ \sum_{j=1}^{N} x(j)v(j)\lambda^{N-j} \right\}
\]

(8)

where \(0 < \lambda < 1\) is the exponential data weighting factor. We note that as in [7], the matrix to be inverted in (8) can be written in a block Hankel matrix form. To achieve such a structure, we modify the parameter vector to
and make the corresponding change in the state vector (this simply changes the order of elements in $\Theta$ and $x(k)$). With the modified state vector denoted by $\bar{x}(k)$, the matrix $\bar{x}(k)\bar{x}'(k)$ is given by

$$
\bar{x}(k)\bar{x}'(k) = \begin{cases} 
Az^2(k-1) & Cz(k-1)y(k-1) \\
C'y(k-1)y(k-1) & By^2(k-1) \\
Ay^2(k-1) & -Cz(k-1)y(k-1) \\
-C'y(k-1)y(k-1) & Bz^2(k-1)
\end{cases}, \text{k odd}
$$

$$
\begin{cases} 
-Az^2(k-1) & Cz(k-1)y(k-1) \\
C'y(k-1)y(k-1) & By^2(k-1) \\
Ay^2(k-1) & -Cz(k-1)y(k-1) \\
-C'y(k-1)y(k-1) & Bz^2(k-1)
\end{cases}, \text{k even}
$$

where each of the matrices $A$, $B$, $C$ is of Hankel form. For example, with $\tau_k \triangleq \Delta t_k$,

$$
A = \begin{bmatrix}
\tau_k^2 & -\tau_k^4 & \tau_k^6 & \cdots & \tau_k^n \\
-\tau_k^4 & \tau_k^6 & -\tau_k^8 & \cdots & \tau_k^{n+2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\tau_k^n & -\tau_k^{n+2} & \tau_k^{n+4} & \cdots & \tau_k^{2n-2}
\end{bmatrix}, \text{(10)}
$$

With a block Hankel form of the matrix, considerable computational savings can be made in the computation of $\hat{\Theta}(N)$ by applying the fast algorithm of (10).

If the noise in model (6) were white then the estimate of $\Theta$ obtained from (8) would be consistent and thus the estimation error $\Theta - \hat{\Theta}(N)$ would approach 0 as $N \to \infty$. However, as $\xi(k)$ is colored, there would be considerable bias in the parameter estimates under low to medium signal-to-noise ratio $A^2/\sigma^2$. To eliminate such a bias we propose the following extended least squares algorithm which essentially whitens the noise on an adaptive basis.

Extended Least Squares Algorithm - The additive colored noise $\xi_i(k), \xi_q(k)$ in general may be expressed as a moving average of the white noise $n_i(k), n_q(k)$ as follows

$$
\begin{bmatrix}
\xi_i(k) \\
\xi_q(k)
\end{bmatrix} = \begin{bmatrix} D_{11}(k, z^{-1}) & D_{12}(k, z^{-1}) \\
D_{21}(k, z^{-1}) & D_{22}(k, z^{-1})\end{bmatrix} \begin{bmatrix} n_i(k) \\
n_q(k)\end{bmatrix}, \text{(11)}
$$

where $D_j(k, z^{-1}); 1 \leq I, j \leq 2$ are some possibly time-varying polynomials in the delay variable $z^{-1}$. For signal model (4-6) we have that
In the proposed recursive extended least squares algorithm [11] we modify the measurement \( v(k) \) so as to effectively whiten the additive noise, thus arriving at the following algorithm

\[
\begin{align*}
D_{11}(k, z^{-1}) &= 1 - \cos(\omega_d \Delta t_k) z^{-1}; D_{12}(k, z^{-1}) = -\sin(\omega_d \Delta t_k) z^{-1} \\
D_{21}(k, z^{-1}) &= \sin(\omega_d \Delta t_k) z^{-1}; D_{22}(k, z^{-1}) = 1 - \cos(\omega_d \Delta t_k) z^{-1}
\end{align*}
\]

In the initialization of the above algorithm \( \eta(1) \) and \( \eta(2) \) may be simply set equal to zero. In fact, in the actual implementation of the algorithm it may be advantageous to first use the standard least squares algorithm for the first few observations to obtain an approximate estimate for \( \Theta_d \) and then use the extended least squares algorithm.

4. CONVERGENCE OF EXTENDED LEAST SQUARES ALGORITHM

The heuristic basis for the convergence of the extended least squares algorithm [11] is as follows. As \( \hat{\Theta}(k) - \Theta \), we have from (6,12) that the prediction error \( \epsilon(k) \) approaches \( \xi(k) - \eta(k) \). This latter term would then approach the white noise \( \eta(k) \) from (11,13). Thus essentially the additive noise \( \xi(k) \) in the measurements \( v(k) \) is whitened so that the parameter estimates obtained by the algorithm are consistent, i.e., \( \hat{\Theta}(k) - \Theta \) as \( k \to \infty \).

References [11,12] contain a rigorous proof of a similar algorithm applied in a somewhat different context. The algorithm of [11] differs from the one proposed here in that in [11] the coefficients of the matrix corresponding to \( D(z^{-1}) \) of (11) are not related to the elements of \( \Theta \) and these are estimated simultaneously by extending the dimension of \( \Theta \). However in the present problem of frequency estimation, the noise parameters (elements of the matrix \( D(z^{-1}) \)) can be obtained from \( \hat{\Theta} \) and thus need not be explicitly estimated. Such a difference also permits us to include the possibility that the model parameters are varying with \( k \) as would be the case if \( \Delta t_k \) is dependent on \( k \).

In the convergence analysis of [11,12], a passivity condition is imposed on the polynomial matrix \( D(z^{-1}) \). This condition can easily be satisfied by the artifice of adding an
independent noise sequence \( \{ \zeta(k) \} \) (also independent of the noise sequence \( \{ n(k) \} \)) to the observations for the purpose of estimating \( \hat{h}(k) \). In the estimation of \( \hat{\Theta}(k) \) however, the additive noise component \( \zeta(k) \) is subtracted from \( \varepsilon(k) \) so that the asymptotic statistical efficiency of the algorithm is not altered due to such a modification.

Another condition for the convergence of the algorithm (12) is that the matrix \( P(k) \to 0 \) as \( k \to \infty \). Equivalently it is required that the matrix \( P^{-1}(k) \) approaches \( \infty \) in all its eigenvalues as \( k \to \infty \). As \( P^{-1}(k) = \sum_{j=1}^{k} x(j)x'(j) \) (assuming that \( \lambda(j) = 1 \)) we observe that this condition will not be satisfied if the dimension \( n \) in the model (6) is greater than 2 and uniform sampling is used, i.e., \( \Delta t_k = T_s \) is a constant. In fact, as is evident from (9,10), under such a condition, the matrix \( P^{-1}(k) \) would have a rank of 2 and the algorithm would break down.

5. FASTER SAMPLING

It is noted that if the dimension of \( \Theta \) is restricted to 2 then it is required that \( \Omega_d T_s \ll 1 \) (\( \Omega_d \) denotes an upper bound on \( \omega_d \)) or the signal needs to be sampled at a much faster rate than the Nyquist rate which corresponds to selecting \( \Delta t_k = T_s \frac{\Delta \pi}{\Omega_d} \). Such a procedure however, would result in a severe degradation of the signal-to-noise (SNR) ratio. For example, if \( T_s \) is selected to be \( 0.1 T_{sN} \) with the filter noise bandwidth equal to \( 1/2 T_s \), then the signal-to-noise ratio in the model (6) would be reduced by a factor of \( 10^3 \) where the SNR is defined as \( E[||x(k)||^2]/E[\xi^2(k)] \). Taking into account the fact that the number of observations would also increase by a factor of 10, the available SNR is still degraded by a factor of 100.

In order to insure that the above convergence condition on \( P^{-1}(k) \) is satisfied without incurring a severe loss in SNR, we propose the following two sampling schemes.

Nonuniform Cyclic Sampling - In this procedure the sampling interval \( \Delta t_k \) is varied periodically to have at least \( n/2 \) distinct values where \( n \) is the dimension of the signal model (6). For example if \( n = 6 \), then \( T_s \) may take values \( T_{s1}, T_{s2}, T_{s3}, T_{s4} \) ...where \( T_{s1}, T_{s2}, \) and \( T_{s3} \) need to be just distinct. For example, one may choose \( T_{s1} = T_{s3N}, T_{s2} = 0.9 T_{s3N}, T_{s3} = 0.8 T_{s3N} \). This would insure that within the first \( n \) measurements the matrix \( (\sum_{j=1}^{n} x_j x'_j) \) would have full rank \( n \) and all its eigenvalues, would increase with \( k \) at the same rate. The actual values or \( T_{s1}, T_{s2}, \) and \( T_{s3} \) would simply influence the condition number (ratio of maximum to minimum eigenvalue) of the matrix \( P^{-1}(k) \), i.e., higher difference among these values would imply smaller condition number and thus a higher robustness against computational inaccuracies. In the limit, of course, when \( T_{s1} = T_{s2} = T_{s3} \), the condition number is \( \infty \) and the algorithm breaks down.
Over Sampling (Uniform and Nonuniform) - In this procedure, to keep the dimension \( n \) small, the continuous-time signal \( \{ y(t), z(t) \} \) is filtered with a low pass filter of bandwidth \( B_N = \Omega_d / 2\pi \) but the filter output is sampled at a much faster rate, i.e., \( \Delta t_k << \pi / \Omega_d \). Such a sampling scheme results in the variance of \( n_i(k), n_q(k) \) to be \( 2N_0B_N \) where \( 2N_0 \) is the one-sided noise spectral density for the continuous-time noise \( n_i(t), n_q(t) \). However, the sequence \( \{ n_i(k) \} \) would now be highly correlated resulting in considerably smaller variance for the differential noise \( n_i(k) - n_i(k-1) \) appealing in the signal model (5,6). Notice that the remaining noise terms in (5) are also small in variance as \( \omega_d \Delta t_k << 1 \).

For example, if the low pass filter has ideal characteristics with bandwidth \( B_N \), then the autocorrelation function of the filtered noise is \( \rho \) by

\[
R(\tau) = (2B_N N_0) \frac{\sin 2\pi B_N \tau}{2\pi B_N \tau}
\]

and for \( \tau = T_{SN} / 10 \) where \( T_{SN} \) is the Nyquist sampling interval as defined earlier we have the correlation

\[
R(\tau) = 2B_N N_0 (.9836).
\]

Selection of the above sampling rate then results in \( E[n_i(k) - n_i(k-1)]^2 = .033(2B_N N_0) = .033 \sigma_i^2 \). Also as \( \omega_d \Delta t_k = \pi / 10 \), the variance of the remaining noise terms in (5) is also of the same order of magnitude.

It should be emphasized that if proper consideration is not given to the differential signal model (5,6) and the filter bandwidth is simply selected to be equal to \( 1/2\Delta t_k \), with \( \Delta t_k = T_{SN} / M \), then the noise variance \( E[n_i(k) - n_i(k-1)]^2 \) would have a value of \( 4B_N M N_0 \) about 1000 times (with \( M = 10 \)) higher than obtained by the above procedure. This holds when the signal power \( E[|x_k|^2] \) is the same in both cases.

With \( \omega_d \Delta t_k \) equal to \( \pi / 10 \) as in this example, it is sufficient to have the dimension of the signal model \( n \) equal to 2. For higher values of \( \Delta t_k \) \( n \) may have a moderately higher value, say \( n = 4 \), in which case the sampling period \( \Delta t_k \) should be periodically varied with a minimum period of 2. The period for the variation of \( \Delta t_k \) must be \( \geq n/2 \) for any other \( n \).

Differential Noise Model - The sampled noise \( \{ n_i(k) \} \) obtained by the oversampling procedure above may be expressed as the moving average of a white noise sequence \( \{ \omega_i(k) \} \), i.e.,

\[
n_i(k) \cong \sum_j h(j) \omega_i(k - j) T_s
\]

(15)
where \( \omega_i(k) \) has a variance of \( N_0/T_{sn} \) and \( \{ h(j) \} \) represents the sampled version of the low pass filter impulse response. For the case of an ideal filter for example,

\[
h(j) = 2B_N \frac{\sin(2\pi B_s T_j)}{2\pi B_s T_s}
\]

Equivalently denoting by \( \bar{h}(t) \) the normalized impulse response \( h(t)12B_N \) and with \( M = T_{sn}/T_s \) as above, we have,

\[
n_i(k) \approx \frac{1}{M} \sum_j \bar{h}(j) w_i(k - j)
\] (16)

The sum in (16) has approximately 2M significant terms. From (16) one could derive a moving average model for the differential noise \( \Delta n_i \equiv n_i(k) - n_i(k-1) \). Clearly there are a large number of significant terms in the expression for the differential noise and one could not even approximately whiten it. (Theoretically such an operation corresponds to filtering the noise by an inverse filter). However, we observe that

\[
E[\Delta n_i(k)\Delta n_i(j)] = 2R_{n_i}(k - j) - R_{n_i}(k - j - 1) - R_{n_i}(k - j + 1)
\]

with \( R_n \), denoting the sampled autocorrelation function of \( n_i(k) \), and that \( E[\Delta n_i(k)\Delta n_i(j)] /\sigma^2_{\Delta n_i} \approx 0 \) for \( |i - j| = M \). Thus by using only a subsequence of the differential measurements placed \( T_{sn} \) sec. apart, in the update of parameter estimates, the \( [n_i(k) - n_i(k-1)] \) component in the expression for \( \xi_i(k) \) in (5) is white. The remaining measurements are only used for the estimate of \( n_i(k-1) \) term in (5), so as to eliminate it from the measurements. Similar considerations apply for the quadrature noise component \( n_q(k) \).

An Alternative Algorithm - Denoting by \( \hat{\Theta}_{\varphi}(k-1) \) the parameter vector estimate based on the differential measurements \( v(j), 1 \leq j \leq 2(k-1)M \), and letting \( \hat{\omega}_d,k(k-1) \) denote the corresponding frequency estimate, we use the following recursion (see (4-6) for the rationale),

\[
\begin{align*}
\hat{n}_i(j) &= y_d(j) - \hat{\Theta}'_M(k-1)x(2j-1) + \cos(\hat{\omega}_d,M(k-1)\Delta t_j)\hat{n}_i(j-1) + \sin(\hat{\omega}_d,M(k-1)\Delta t_j)\hat{n}_q(j-1) \\
\hat{n}_q(j) &= y_d(j) - \hat{\Theta}'_M(k-1)x(2j) + \cos(\hat{\omega}_d,M(k-1)\Delta t_j)\hat{n}_i(j-1) - \sin(\hat{\omega}_d,M(k-1)\Delta t_j)\hat{n}_q(j-1)
\end{align*}
\] (17)

At the end of recursion (17), one would have the estimates of \( n_i(kM-1), n_q(kM-1) \) required in the update of \( \hat{\Theta}_{\varphi}(k) \). The modified parameter estimation algorithm may be written as,
\[ \hat{\Theta}_M(k) = \hat{\Theta}_M(k-1) + K(k)\zeta(k) \]
\[ K(k) = \tilde{P}(k-1)\tilde{x}(k)[\lambda(k)I + \tilde{x}'(k)\tilde{P}(k-1)\tilde{x}(k)]^{-1} \]
\[ \tilde{P}(k) = \{ \tilde{P}(k-1) - \tilde{P}(k-1)\tilde{x}(k)[\lambda(k)I + \tilde{x}'(k)\tilde{P}(k-1)\tilde{x}(k)]^{-1}\tilde{x}'(k)\tilde{P}(k-1)\tilde{x}(k) \} / \lambda(k) \]
\[ \tilde{x}'(k) = \begin{bmatrix} x'(2kM-1) \\ x'(2kM) \end{bmatrix} \]
\[ \zeta(k) = \begin{bmatrix} y_d(kM) \\ z_d(kM) \end{bmatrix} - \tilde{x}'(k)\hat{\Theta}_M(k-1) - \tilde{\omega}_{d,M}(k-1)\Delta t_{kM} \begin{bmatrix} -n_c(kM-1) \\ n_c(kM-1) \end{bmatrix} ; \]
\[ k = 1, 2, \ldots \]

6. ESTIMATION ERROR ANALYSIS

As shown in [7], the parameter estimation error covariance matrix for the case of additive white noise in the signal model may be written as

\[ E \left[ \tilde{\Theta}(N)\tilde{\Theta}'(N) \right] = \left( \sum_{j=1}^{N} x_j x_j' \right)^{-1} \sigma^2 \]

where \( \sigma^2 \) is the variance of the additive white noise and \( \tilde{\Theta}(N) = \Theta - \hat{\Theta}(N) \) is the parameter estimation error. For the extended least squares algorithm considered in this paper, the additive noise is colored. However, asymptotically the noise is effectively whitened and the above expression is applicable, at least approximately.

For example, for the case of algorithm (16,17) and assuming that the signal dimension \( n \) is equal to 2, one can easily show that with \( M=10 \) as above,

\[ E \left[ \tilde{\omega}^2(N) \right] \approx \frac{3.3N_0}{NP_cT^2} \]

where \( \tilde{\omega}(N) \) denotes the frequency estimation error and \( P_c \) denotes the carrier power. Similar expression may be derived for the algorithm (12).

It may be noted, however, that the expression in (20) is mostly of academic interest as it only applies to the case of large number of observations without exponential data weighting. For small number of observations and with exponential data weighting, the algorithm actually exhibits a faster than exponential convergence phase followed by an exponential rate of convergence. Such a convergence is illustrated by computer simulations. However, if the maximum value of the data weighting coefficient \( \lambda_{\text{max}} \) is strictly less than one, then the parameter estimation error does not converge to zero with the number of observations. It is only for such an asymptotic convergence (perhaps not
of much interest in the present context) that \( \lambda_{\text{max}} \) is set equal to 1 and then the result of (20) is applicable asymptotically.

7. COMPARISON WITH PREVIOUS ALGORITHMS

Comparing the performance of the proposed algorithms with the one studied in [7], there is an obvious loss in terms of the statistical efficiency. However, one also achieves an order of magnitude reduction in the computational complexity. Thus, whereas the dimension \( n \) of the signal model in [7] depends directly upon \( \Omega_d T \), the product of the frequency uncertainty and the observation time, here the signal model dimension \( n \) is independent of \( \Omega_d T \) and can be made very small, in fact \( n \) can be selected to be 1 if \( M \) is made sufficiently large.

It should be emphasized however, that over the initial period of estimation, the algorithm is fast, exhibiting an exponentially fast convergence phase similar to that in [7]. It is also possible to devise algorithms which provide a tradeoff between the convergence rate and computational complexity by taking advantage of both of the algorithms.

For instance, in one possible approach, one may use the simpler algorithm proposed here, first to estimate \( \omega_r \) for an appropriate period so that the estimation error \( \tilde{\omega}_r \) is sufficiently small. At this point a correction is made in the reference oscillator frequency by \( \hat{\omega}_d \). The unknown frequency after the correction is simply equal to \( \hat{\omega}_d \). This is followed by the more precise algorithm of [7] where the signal dimension \( n \) is selected according to (18) of [7] with \( 3\sigma_{\omega_d} \) substituted for \( \Omega \), i.e.,

\[
 n^3 = 216 \frac{N_0}{P} 3\sigma_{\omega_d}
\]

For example if \( n=5 \), and \( P/N_0 = 25 \text{ dB-Hz} \) then one uses the simpler algorithm until \( \sigma_{\omega_d} \) - the standard deviation of the estimation error is less than 61 rad/sec. and then switch as to the more precise algorithm. In the latter phase the algorithm also provides the estimates of the phase \( \phi \) and amplitude \( A \) of the carrier.

In yet another approach, one could work with measurements

\[
\{y(1) - y(0), y(2) - y(0), \ldots, y(M) - y(0), y(M + 1) - y(M), \ldots, y(2M) - y(M), \ldots\}
\]

with the signal model dimension \( n \) higher than \( \Omega_d M T_{sn} \). In this case one could write a signal model similar to (4-6) and an algorithm similar to that of (10-12). The performance of such an algorithm is expected to be in between the algorithm of [7] and the one proposed earlier in this paper. The details of such a possible modification are omitted.
8. CARRIER PHASE AND AMPLITUDE ESTIMATION

When the frequency estimation error $\tilde{\omega}$ is small, it may then be desirable to also obtain an estimate of the carrier phase $\phi$ so that a phase correction may be applied to the reference oscillator in order to achieve a phase lock condition as is done in phase-locked loops. For this purpose, instead of estimating the initial phase at $t = 0$, it is more appropriate to estimate the phase $\phi(T)$ at the end of estimation period $T$. To achieve this we rewrite the observation model (2) in the following modified form.

$$
\bar{y}(k) \triangleq y(N - k) = A\sin(-\omega_d k T_s + \phi(T)) + n_i(k) \\
\bar{z}(k) \triangleq z(N - k) = A\cos(-\omega_d k T_s + \phi(T)) + n_q(k); \quad k = 0, 1, ..., \bar{N}
$$

(21)

where for simplicity of the expression an uniform sampling of the signal has been assumed. Equation (21) in turn may be written in the following equivalent form,

$$
\begin{bmatrix}
\bar{y}(k) \\
\bar{z}(k)
\end{bmatrix} = 
\begin{bmatrix}
\cos(\omega_d k T_s) & -\sin(\omega_d k T_s) \\
\sin(\omega_d k T_s) & \cos(\omega_d k T_s)
\end{bmatrix}
\begin{bmatrix}
A\sin(\phi(T)) \\
A\cos(\phi(T))
\end{bmatrix} + 
\begin{bmatrix}
n_i(k) \\
n_q(k)
\end{bmatrix}; \quad k = 0, 1, ..., \bar{N}
$$

(22)

An extended least squares (Kalman filter) algorithm can then be applied to (22) for estimating $A\sin(\phi(T))$ and $A\cos(\phi(T))$. In such an algorithm the unobserved states $\cos(\omega_d k T_s)$ and $\sin(\omega_d k T_s)$ can be replaced by their estimates respectively obtained by simply replacing $\omega_d$ by $\tilde{\omega}_d(N)$. The value of $\bar{N} < N$ is such that $\bar{N} \tilde{\omega}_d(N) T_s$ and thus the state estimation error is small. From the knowledge of $p(T) \triangleq A\sin(\phi(T))$ and $q(T) \triangleq A\cos(\phi(T))$, one may easily obtain the following estimates of $A$ and $\phi(T)$ (modulo $2\pi$),

$$
\dot{A} = \left\{ \dot{p}^2(T) + \dot{q}^2(T) \right\}^{1/2} \\
\dot{\phi}(T) = \tan^{-1} \left[ \dot{p}(T)/\dot{q}(T) \right], \quad \dot{q}(t) > 0 \\
= \pi + \tan^{-1} \left[ \dot{p}(T)/\dot{q}(T) \right], \quad \dot{q}(T) < 0
$$

(23)

Thus far in the development of the paper, we have assumed that the frequency $\omega_d$ is constant. If $\omega_d$ is not constant but is slowly varying with time, it may be tracked by exponential data weighting with $\lambda(k) \leq \lambda_{max} < 1$ with $\lambda_{max}$ selected in accordance with the rate of variation in $\omega_d$. However, if the variation in $\omega_d$ is fast then it is more appropriate to simultaneously estimate $\omega_d$ and possibly also the higher order derivatives of $\omega_d$. The proposed algorithm can be extended for such an extension as discussed below.

9. ESTIMATION OF THE FREQUENCY DERIVATIVES

Here we model the time-varying frequency $\omega_d$ as a polynomial in $t$ of sufficiently high degree. For example, if only the first frequency derivative has a significant contribution
over the estimation period, then \( \omega_d = \omega_0 + \lambda_0 t \) where \( \omega_0 \) and \( \lambda_0 \) are some constants or possibly slowly varying functions of time. In this case the expansion similar to (3) for \( \sin(\omega_d t_k + \phi) \) around \( t_{k-1} \) is given by,

\[
\sin(\omega_d t_k + \phi) = \sin(\omega_d t_{k-1} + \phi) + (\omega_0 + \lambda_0 t_{k-1}) \Delta t_k \cos(\omega_d t_{k-1} + \phi) \\
+ \lambda_0 \frac{\Delta t_k^2}{2!} \cos(\omega_d t_{k-1} + \phi) - (\lambda_0 + \omega_0 t_{k-1})^2 \frac{\Delta t_k^2}{2!} \sin(\omega_d t_{k-1} + \phi) + \ldots
\] (24)

In (24) above the dependence of \( \omega_d \) on \( t \) is not shown explicitly. It is however understood that \( \omega_d \) on the left hand side is evaluated at \( t = t_k \) whereas in the right hand side terms the value of \( \omega_d \) at \( t_{k-1} \) is substituted.

Following the procedure similar to that used in the derivation of (4-7), and assuming that \( (\omega_d \Delta t_k)^3 \) and higher order terms are negligible, one obtains the following differential model for \( y(k) \),

\[
y_d(k) = \Theta' z_i(k) + \xi_i(k)
\] (25)

where

\[
\Theta' = [\omega_0 \quad \lambda_0 \quad \omega_0 \lambda_0 \quad \lambda_0^2]
\]

\[
z_i(k) = \begin{bmatrix}
\Delta t_k z(k-1) & (t_{k-1} + \frac{1}{2} \Delta t_k) \Delta t_k z(k-1) & -\frac{1}{2} \Delta t_k^2 y(k-1) & -t_{k-1} \Delta t_k^2 y(k-1) & \frac{1}{2!} \omega_0^2 \Delta t_k^2 y(k-1)
\end{bmatrix}
\]

\[
\xi_i(k) = n_i(k) - n_i(k-1) - \omega_d \Delta t_k n_i(k-1) - \frac{1}{2} \lambda_0 \Delta t_k^2 n_i(k-1) - \frac{1}{2!} \omega_d^2 \Delta t_k^2 n_i(k-1)
\]

A similar expression can easily be written for the differential \( z_i(k) \). Extended least squares algorithm can now be applied to model (25) to estimate the parameter vector \( \Theta \) and thus the unknown frequency \( \omega_0 \) and it derivative \( \lambda_0 \).

### 10. SIMULATIONS

Figures 1 through 3 present the performance of the least squares algorithm presented in the paper. For the simulations, the unknown frequency \( \omega_d \) is taken to be 1 rad/s. For frequencies much higher than one, an appropriate scaling of the frequency and time can be made as in [7]. Such a transformation of the frequency and time as in [7], leaves the error analysis of the algorithm invariant. However, this imparts an additional robustness to the algorithm against finite-dimensional truncation error.

In view of the above remarks, the plots in Figures 1 through 3 are equally applicable to any arbitrary frequency (say 1000 rad/s), if the estimation error in the figures is considered to be also normalized, i.e., in such a case the ordinate in Figures 1-2 represent \( \tilde{\omega}_d / \Omega_d \) and the number of samples/ cycles are with respect to the known frequency bound \( \Omega_d \).
As is evident from Figure 1, in the absence of noise the frequency estimate comes close to the true frequency (within a few percent) in only a fraction of one period of the unknown frequency. It is also apparent that the frequency estimation error exhibits an initial, faster than exponential convergence phase, followed by an exponential convergence phase (linear on logarithm scale). When the estimation error becomes very small, say about .001 of the true frequency, the truncation error may become relatively significant (the contribution due to higher than \((\omega_d \Delta t_k)^4\) terms may not be negligible) and to reduce the effect of such unmodelled error, a simple smoother is used which exponentially weights the filtered frequency estimates. For the estimation algorithm, the exponential data weighting of 0.995 is used.

In Figure 2 is plotted the frequency estimation error vs the estimation period when noise is present in the measurements. In the simulation example, an additive noise of variance .01 is added to the differential signal model (3-6). This would thus be the variance of the whitened version of the additive noise \(\xi(k)\) in (6). As discussed in the section on the sampling techniques, the white noise variance in the differential signal model is only a small fraction of the additive noise variance in the original measurements if proper sampling procedure is followed. Here for example the oversampling factor \(M=5\) and for this case \(E[\Delta n_i(k)]^2 = .065\sigma_{ni}^2\) (see the discussion following equation (14)). Thus the effective noise variance considered is \(\sigma_{ni}^2 = .15\) with \(10\log_{10}(A^2/2\sigma_{ni}^2) = 5.2\) dB. The figure does not clearly expose the initial convergence phase. However comparison with Figure 1 shows that over the initial 10 samples of estimation period, the estimation error for this case is approximately the same as for the noise free case. It is easily verified that in this initial period, the performance of the algorithm is very close to the optimum possible performance [7].

Figure 3 illustrates the capability of the algorithm to track time-varying frequency. In this example the unknown frequency is varying sinusoidally according to \(\omega_d(t) = \omega_0[1 + 0.5\sin(0.01\omega_0 t)]\) representing a maximum frequency derivative equal to .005\(\omega_0\). The result is given normalized to \(\omega_0 = 1\) rad/s. For the case of \(\omega_0 = 1000\) rad/s this represents, for example, a maximum magnitude of the derivative equal to 5000 rad/s\(^2\). It is noted that the estimate follows closely the true frequency. We also remark here that in this example, there is no explicit estimation of the frequency derivative. Thus with the application of the extended algorithm (24,25) much more accurate estimates can be obtained in this case as the frequency variation may hardly be considered to be slow in this example.

It should be remarked that by reducing the effect of model truncation errors (most dominant over intermediate estimation periods), the performance of the algorithm can be kept close to the optimum over the complete estimation period. This is achieved simply by a periodic adjustment of the reference oscillator frequency (by the estimate of \(\omega_d\)).
Thus as the offset (equal to $\tilde{\omega}_d$) tends to zero, the effect of the truncation error becomes negligible and $\tilde{\omega}_d$ tends to zero.

11. CONCLUSIONS

We have presented a differential signal model along with appropriate sampling techniques for least squares estimation of the frequency and frequency derivatives and also the phase and amplitude of a sinusoid received in the presence of noise. The proposed algorithm is recursive in measurements and thus the computational requirements are of order $N$ where $N$ is the number of measurements. Depending upon the sampling rate, the dimension of the state vector in the least squares algorithm may be anywhere from 1 to 5. With exponential data weighting, the proposed algorithm exhibits a faster than exponential convergence phase followed by an exponential phase as is shown in [14] for the recursive least squares algorithm applied to a different problem.

When compared to the FFT algorithm the algorithm of the paper shares the advantage of the previously proposed least squares algorithm of [7]. However, the computational requirements here are considerably lower than for the previous algorithm. Moreover, whereas the dimension of the state vector in [7] increases with the product of frequency uncertainty $\Omega_d$ and the observation period $T$, here the state vector dimension does not depend upon $\Omega_d T$ and is small. The disadvantage of the algorithm may be some loss of asymptotic optimality in terms of frequency estimation errors. However, using appropriate sampling procedures proposed in the paper, such a difference can be made small especially over initial estimation periods, where the estimation errors are more significant. Moreover, with an appropriate synthesis of the two algorithms as suggested in the paper, one may obtain near optimum performance with reduced computational requirements.

ACKNOWLEDGEMENT

The research described in this paper was carried out in part by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration, in part sponsored by the United States Air Force Systems Command through an agreement with the National Aeronautics and Space Administration, under Task Order RE-182, Amendment 452. A significant part of the work was carried out at the California State University, Long Beach.

REFERENCES


Figure 1. Convergence of Least Squares Algorithm (Noise Free Case)
Figure 2. Convergence of Least Squares Algorithm in the Presence of Noise
Figure 3. Acquisition of Rapidly Varying Frequency Without Explicit Estimation of the Frequency Derivative