Performance of the Maximum Likelihood Constant Frequency Estimator for Frequency Tracking

Mehmet Karan, Robert C. Williamson, and Brian D. O. Anderson, *Fellow, IEEE*

***Abstract*— In this paper, the performance of maximum likelihood (ML) estimators for an important frequency estimation problem is considered when the signal model assumptions are not valid. The motivation for this problem is to understand the robustness of the hidden Markov model-maximum likelihood (HMM-ML) tandem frequency estimator [1], where the signal is divided into time blocks, and the frequency in each time block is estimated using the ML approach under the assumption that the signal has a constant frequency in each time block. In order to analyze the sensitivity of ML estimators to the model assumptions, the mean frequency of a discrete complex tone that has a time-varying (ramp) frequency is estimated under the incorrect assumption that it has a constant frequency. In particular, the behavior of the threshold region with respect to different chirp rates is analyzed, and a simple rule is given. The mean squared error above the threshold region is shown to be constant even at very high SNR levels. These results are supported by simulations.**

I. Introduction

ESTIMATING the parameters of a sinusoidal signal from a given discrete set of observation data is an important problem arising in different branches of science and engineering. In the time-series analysis literature, recent results for this problem are due to Walker [2], Hannan [3], and Hasan [4]. In those papers, the signal is assumed to be in the form

$$z[n] = A\cos[\Omega_0 n] + B\sin[\Omega_0 n] + \epsilon[n], \quad n = 0, \ldots, N-1 \tag{1}$$

where $\epsilon[n]$ is i.i.d. noise. It has been shown that the least-squares estimate of Ω_0, which is asymptotically equivalent to a maximum-likelihood estimate when $\epsilon[n]$ is a Gaussian noise, is the maximizer of the periodogram

$$I_N(\Omega) = \frac{1}{N}\left|\sum_{n=0}^{N-1} z[n]e^{-j\Omega n}\right|^2. \tag{2}$$

It has also been shown that the asymptotic variance of the frequency estimation is of order N^{-3}. However, in [5], it is concluded that the product of the amplitude of the signal and data size must be quite large for the reliability of the asymptotic analysis results.

Manuscript received February 10, 1993; revised December 10, 1993. This work was supported by the Cooperative Research Centre for Robust and Adaptive Systems by the Australian Commonwealth Government under the Cooperative Research Centres Program. The associate editor coordinating the review of this paper and approving it for publication was Prof. John Goutsias.

The authors are with the Department of Systems Engineering, Research School of Information Sciences and Engineering, Australian National University, Canberra, Australia.

IEEE Log Number 9403760.

In [6], the signal is assumed to be in the form of a single-frequency complex tone, and its parameters are estimated from a finite number of noisy discrete-time observations. In [6], Rife and Boorstyn showed that the maximum-likelihood estimate attains the Cramer-Rao bounds at high signal to noise ratio (SNR). In addition, at low SNR, the existence of a range of SNR's is observed where the mean-squared error (MSE) increases very rapidly with reduction in SNR. This threshold effect can be understood in terms of the "outliers" occurring in the maximization of (2) (see [6] for details). In [7], it has been shown that the approximate phase error variance is a good indicator for the threshold. The threshold effects are also discussed in [8] from an information theoretic point of view.

In the previous estimation approaches, the signal is assumed to have a constant frequency during the measurement interval. When the signal has a time-varying frequency over the measurement interval, the frequency can, in principle, be tracked by using either extended Kalman filters (EKF's) [9] (or phase-locked loops (PLL) if the amplitude is known) or a hidden Markov model-maximum likelihood (HMM-ML) tandem estimator [1]. In the EKF approach, the parameters of the signal are estimated by using a Kalman filter where the output matrix of the associated signal model is obtained by linearizing the measurement around the one-step prediction of the parameters. Here, there is an assumption on the random variation of the frequency but no assumption on the deterministic variation of the frequency. In contrast, Streit and Barrett [1] assumed that the frequency variation is piecewise constant. Accordingly, they divided the signal time function into fixed finite-sized blocks. Then, the frequency in each block was estimated using a ML approach, and frequency transitions between successive blocks were modeled using hidden Markov models (HMM's). Their algorithm works well even at very low SNR's.

Since fixed block sizes are used in [1], the assumption that the signal has constant frequency over these blocks may not always be valid. In order to assess the practical utility of the tandem estimator, it is important to understand the degradation in the performance of the HMM-ML tandem estimator when the signal has a time-varying frequency. The crucial question is as follows: What happens if the frequency is estimated using a ML estimator that is based on the assumption that the signal has a constant, unknown single frequency when, in fact, the signal has a slowly time-varying single frequency? In this paper, we will analyze this problem. The aim is to be able to give rules for the choice of the block length, i.e. design parameters. The techniques in Rife and Boorstyn [6] will be the basis for our analysis and identify some of the fundamental

mechanisms causing departure from predicted performance. The model that is selected for our purposes is a complex signal that has the simplest form of time-varying frequency, viz. a ramp. Although selection of the model for the time variation of the frequency as a ramp is restrictive, it is useful to understand the nature of the problem, and it clearly gives some insight for the general problem of estimation of signals that have a more general time-varying frequency.

A. Assumptions and Signal Model

Specifically, the signal that we will analyze in this paper is a linear FM signal that has the form

$$s(t) = b_0 \exp[j(\omega_0 + \Delta\omega(t))t] \quad t \in [0, T_1] \tag{3}$$

where $\Delta\omega(t) = \omega_1 t$ and T_1 is the length of the signal; the sampled version of the signal $s(t)$ is

$$s[n] = b_0 \exp[j(\Omega_0 + \Omega_1 n)n] \tag{4}$$

where T is the sampling period, $\Omega_0 = \omega_0 T$, and $\Omega_1 = \omega_1 T^2$. The sampling frequency f_s is defined by $f_s = 1/T$ and $\omega_s = 2\pi f_s$.

The measurement data $z[n]$, for $n = 0, \ldots, N-1$, is assumed to be

$$z[n] = s[n] + w[n] \tag{5}$$

where $w[n] = w_R[n] + jw_I[n]$, and both $w_R[\cdot]$ and $w_I[\cdot]$ are Gaussian noise sequences with mean zero and variance σ^2. They are statistically independent of each other. The SNR is defined as the the ratio between the average signal power and the average noise power, which is $b_0^2/2\sigma^2$ for the noisy complex linear FM signals.

The signal model is selected as a complex sinusoid in order to compare our results with the results of Rife and Boorstyn [6], and in practice, this kind of signal model is used to avoid leakage problems (see [10]). In addition, note that one can create a complex signal from its real part by using the Hilbert transform. Care must be taken since, in this case, $w_R[\cdot]$ and $w_I[\cdot]$ will no longer be independent noises; however, the cross correlation between $w_R[\cdot]$ and $w_I[\cdot]$ can be removed by down sampling the complex measurement signal.

In the next section, we will summarize the outlier analysis of Rife and Boorstyn for the constant frequency case. In the third section, we will derive the statistics of the periodogram of the complex measurement signal $z[n]$. The fourth section is devoted to generalization of the outlier analysis for the estimation of the linear FM signal under the incorrect assumption that its frequency is constant. In the last section, mean squared error of the frequency estimates analyzed at different SNR levels and a simple rule for the threshold SNR is given in terms of the chirp rate. Simulation results and theoretical results are compared.

II. Outlier Analysis of Rife and Boorstyn for Constant Frequency

In [6], the parameters of a complex, single-frequency tone are estimated from a finite number of noisy discrete-time observations by using the maximum-likelihood (ML) estimation technique. In that paper, the parameters of $s(t)$ in (3) (b_0 and ω_0) are estimated when $\Delta\omega(t) = 0$ from the noisy measurement signal $z[n]$ for $n = 0, \ldots, N-1$. Then, the ML estimate of the frequency $\Omega_0 = \omega_0/T$ is given by

$$\hat{\Omega}_0 := \arg\max_{\Omega} \left\{ \frac{1}{N} \mid Z(\Omega) \mid \right\} \tag{6}$$

where $Z(\Omega)$ is the discrete-time Fourier transform of $z[\cdot]$, which is defined by

$$Z(\Omega) := \sum_{n=0}^{N-1} z[n] \exp(-j\Omega n). \tag{7}$$

Like most nonlinear estimators, the ML frequency estimator exhibits threshold effects. At high SNR's, the frequency estimate occurs near the true frequency. As the SNR decreases, the probability that the global maximum of $\mid Z(\Omega) \mid /N$ lies far from the true frequency increases. The occurrence of these outliers (i.e., the frequency estimates that are far from the true frequency) causes a sudden decrease in the performance of the ML estimator as the SNR is reduced. The mean-squared error can be written as

$$\begin{aligned} E\{(\hat{\omega} - \omega_0)^2\} &= (1-q)E\{(\hat{\omega} - \omega_0)^2 \mid \text{No outlier}\} \\ &\quad + qE\{(\hat{\omega} - \omega_0)^2 \mid \text{outlier}\} \end{aligned} \tag{8}$$

where q denotes the outlier probability, when $f_0 = f_s/2$. If $f_0 \neq f_s/2$, (8) is still valid, but the terms $E\{(\hat{\omega} - \omega_0)^2 \mid \text{outlier}\}$ and $E\{(\hat{\omega} - \omega_0)^2 \mid \text{No outlier}\}$ will differ. Since the frequency estimate is close to the true frequency when there is no outlier, $E\{(\hat{\omega} - \omega_0)^2 \mid \text{No outlier}\}$ is approximately equal to the Cramer-Rao bound, which is derived in [6]. In addition, when the DFT size is equal to the data size, the DFT of the noise is i.i.d; therefore, when an outlier occurs, the probability density function of the DFT of the noise is approximately uniform; thus, $E\{(\hat{\omega} - \omega_0)^2 \mid \text{outlier}\}$ can be calculated easily. Finally, when $\omega_0 = \omega_s/2$, i.e., when there is no bias, the outlier probability q is calculated as

$$1 - q = \Pr\{\text{All } \mathcal{Z}_k < \mathcal{Z}_{N/2}\} \tag{9}$$

$$= \int_0^\infty f_{\mathcal{Z}_{N/2}}(x) \left[\int_0^x f_{\mathcal{Z}_k}(y)dy \right]^{N-1} dx \tag{10}$$

where $\mathcal{Z}_k := \left\{ \frac{1}{N} \mid Z(\Omega) \mid \right\} \mid_{\Omega = 2\pi k/N}$ ($k = 0, \ldots, N-1$ and $k \neq N/2$), $f_{\mathcal{Z}_k}(x)$ is the Rayleigh probability density function defined by

$$f_{\mathcal{Z}_k}(x) = \frac{Nx}{\sigma^2} \exp\left(-\frac{Nx^2}{2\sigma^2} \right) \tag{11}$$

and $f_{\mathcal{Z}_{N/2}}(x)$ is the Rician probability density function defined by

$$f_{\mathcal{Z}_{N/2}}(x) = \frac{Nx}{\sigma^2} \exp\left[-\frac{N(x^2 + b_0^2)}{2\sigma^2} \right] I_0\left(\frac{Nb_0 x}{\sigma^2} \right). \tag{12}$$

Here, $I_0(x)$ is the modified Bessel function of the first kind.

III. Statistics of the Periodogram of the Linear FM Signal

In order to analyze the threshold effects for the ML estimator analytically under the assumptions of the first section, we need to derive the outlier probability q, which depends on the statistics of absolute value of the DFT of the measurement signal. Since $\arg\max|Z(\Omega)|$ is equal to $\arg\max\{Z_R(\Omega)^2 + Z_I(\Omega)^2\}$, where $Z_R(\cdot)$ and $Z_I(\cdot)$ are the real and imaginary parts of the DFT of the measurement signal $z[n]$, the outlier probability can be derived from the statistics of $Z_R(\cdot)$ and $Z_I(\cdot)$. However, the real and imaginary parts of the DFT of a complex signal are the DFT's of the complex conjugate (even) and anticonjugate (odd) parts of the signal. Let $z_e[n]$ and $z_o[n]$ denote the even and odd parts of the signal $z[n]$ defined by

$$z_e[n] = \frac{z[n] + z^*[M-n]}{2} \tag{13}$$

$$z_o[n] = \frac{z[n] - z^*[M-n]}{2} \tag{14}$$

where M is the DFT size. Then, $Z_R(\Omega) := \mathrm{DFT}\{z_e[n]\}$ and $Z_I(\Omega) := \mathrm{DFT}\{z_o[n]\}$. Furthermore,

$$Z_R(\Omega) = S_R(\Omega) + W_R(\Omega) \tag{15}$$

and

$$Z_I(\Omega) = S_I(\Omega) + W_I(\Omega), \tag{16}$$

where $S_R(\cdot)$ and $S_I(\cdot)$ are the real and imaginary parts of the discrete-time Fourier transform of $s[\cdot]$, and $W_R(\cdot)$ and $W_I(\cdot)$ are the real and imaginary parts of the discrete-time Fourier transform of $w[\cdot]$.

It can be shown [11] that the continuous-time Fourier transform of $s(t)$ defined by (3) is

$$S(\omega) = b_0\sqrt{\frac{\pi}{2\omega_1}}\exp\left(-\frac{j(\omega-\omega_0)^2}{4\omega_1}\right) \times\left\{K\left[\sqrt{\frac{2\omega_1}{\pi}}\left(T_1 - xs\frac{\omega-\omega_0}{2\omega_1}\right)\right] -K\left[\sqrt{\frac{2\omega_1}{\pi}}\left(\frac{\omega_0-\omega}{2\omega_1}\right)\right]\right\} \tag{17}$$

where $K(x)$ is the Fresnel integral defined by

$$K(x) = \mathcal{C}(x) + j\mathcal{S}(x) = \int_0^x e^{j\pi\tau^2/2}d\tau. \tag{18}$$

It can be seen from the graphs of the functions $\mathcal{C}(x)$ and $\mathcal{S}(x)$ in Fig. 1 and (17) that when $\Delta\omega$ increases while the measurement signal length T_1 stays fixed, the peak amplitude of the Fourier transform of the signal decreases since $\Delta\omega = \omega_1 T_1$. This can be seen also in Fig. 2, which shows both noiseless quantities, viz. $|S(f)|$, and noisy quantities, viz. $|Z(f)|$. Since the instantaneous frequency of the signal changes between ω_0 and $\omega_0 + 2\omega_1 T_1$, it is not surprising that the spectrum of the signal is approximately nonzero in the interval $[\omega_0, \omega_0 + 2\Delta\omega]$. As can be seen from Fig. 2, the magnitude of the Fourier transform of the signal can be approximated by a rectangular window that has a width

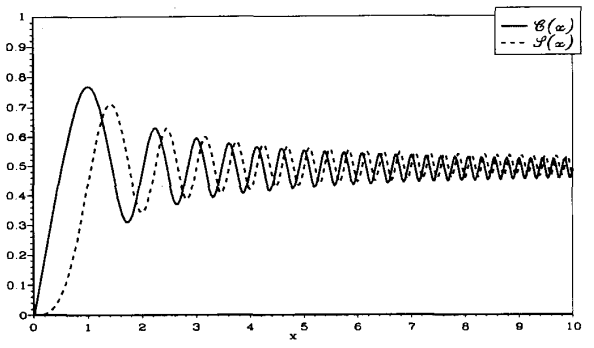


Fig. 1. Real and imaginary parts of the Fresnel integral $\mathcal{C}(x)$ and $\mathcal{S}(x)$.

$2\Delta\omega$ and a height proportional to $1/\sqrt{\Delta\omega}$. Additionally, the magnitude of the Fourier transform of the signal is symmetric around the mean frequency $\omega_0 + \Delta\omega$ of the signal, and it has its global maxima when $\omega \approx \omega_0$ and $\omega \approx \omega_0 + 2\Delta\omega$. This last property can be verified by examining the detailed shape of the Fresnel integral function involved in (17). As $\Delta\omega$ varies, the SNR remains constant. However, the probability of an outlier is clearly greater for bigger $\Delta\omega$ since the Fourier transform of the signal has a smaller amplitude for larger $\Delta\omega$.

In addition, we can consider the effect of varying SNR while holding $\Delta\omega$ fixed. At very high SNR's the mean squared error of the frequency estimate around the mean frequency will be approximately equal to $(\Delta\omega)^2$. This is because $|S(\omega)|$ is maximized for $\omega \approx \omega_0$ and $\omega \approx \omega_0 + 2\Delta\omega$. However, as SNR decreases slightly, some of the peaks inside the rectangular window will be detected falsely as the maxima, promoting decrease in the mean squared error. Afterwards, as the SNR decreases further, outliers (i.e., the frequencies far from the rectangular window) will be observed as the maximum of the Fourier transform of the measurement signal. (All these effects can be seen in the simulation examples presented later in this paper.)

If the sampling frequency is sufficiently high, then the discrete-time Fourier transform $S(\Omega)$ of $s[n] = s(nT)$ is equal to $\frac{1}{T}S(\omega)$, where $\Omega = \omega T$ and the DFT of the signal $s[n]$ is $S[k] \simeq \frac{1}{M}S(\Omega)\ |_{\Omega=2\pi k/M}$.

Now, in order to obtain the statistics of the DFT of the measurement signal, we need to find the statistics of the real and imaginary parts of the DFT of the noise term $w[\cdot]$. Since $W_R[\cdot] = \mathrm{DFT}\ \{w_e[\cdot]\}$ and $W_I[\cdot] = \mathrm{DFT}\{w_o[\cdot]\}$, then

$$E\{W_R[k]W_R^*[l]\} = \frac{1}{M^2}\sum_{n=0}^{M-1} E\{w_e[n]w_e^*[n]\}e^{j\frac{2\pi}{M}n(l-k)}. \tag{19}$$

When the DFT size M is equal to measurement signal size N, then (19) reduces to

$$E\{W_R[k]W_R^*[l]\} = \frac{\sigma^2}{N}\delta(l-k) \tag{20}$$

with $\delta(\cdot)$ being the Kronecker delta function. In this case, both $W_R[l]$ and $W_I[l]$ are white Gaussian random variables with mean 0 and variance σ^2/N, which are independent from each other for $l = 0, \ldots, N-1$. Hence, $Z_R[l]$ and $Z_I[l]$ are

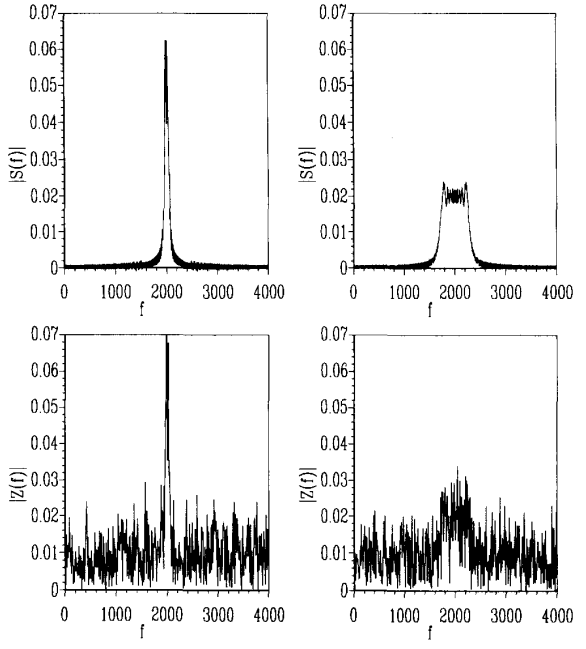


Fig. 2. $|\ S(f)\ |$ and realizations of $|\ Z(f)\ |$ for $b_0 = 1$, $f_s = 4000$ Hz $f_0 + \Delta f = 2000$ Hz, $\sigma = 1$ and $\Delta f = 50$, and 300 Hz.

independent Gaussian random variables with means $S_R[l]$ and $S_I[l]$ and variance $\frac{\sigma^2}{N}$. Thus, $\mathcal{Z}_l = |\ Z[l]\ |^2 = Z_R^2[l] + Z_I^2[l]$ has a noncentral Chi-Square (χ'^2) probability distribution with $\mu = 2$ degrees of freedom and noncentrality parameter $\lambda = \frac{(S_R[l]^2 + S_I^2[l])N}{\sigma^2}$. Then

$$F_{\mathcal{Z}_l}(x) := \Pr\{\mathcal{Z}_l \le x\} = F_{\chi'^2}\left(\frac{Nx}{\sigma^2}; \mu, \lambda\right) \tag{21}$$

where the probability density function for a noncentral Chi-Square distribution for $\mu = 2$ degrees of freedom and noncentrality parameter λ is given by [12]

$$f_{\chi'^2}(x; \mu, \lambda) = \frac{1}{2} I_0(\sqrt{\lambda x}) \exp\left[-\frac{1}{2}(\lambda + x)\right], \quad x \ge 0. \tag{22}$$

IV. Analysis of MSE of the Frequency Estimate of the Linear FM Signal

As indicated in the previous section, at high SNR's, the estimated frequency will be very close to ω_0 or $\omega_0 + 2\Delta\omega$. Then, as SNR decreases further, the local maxima inside $[\omega_0, \omega_0 + 2\Delta\omega]$ of the Fourier transform magnitude of the signal will be detected. At very low SNR's some frequency estimates will occur outside of the region $[\omega_0, \omega_0 + 2\Delta\omega]$. Only those frequency estimates that occur outside of this region will be termed outliers. Therefore, an outlier occurs if the amplitude of the DFT of the measurement signal outside the region $[\omega_0, \omega_0 + 2\Delta\omega]$ is greater than the amplitude of the DFT of the measurement signal inside this region. In other words, if we define the random variables

$$D_L := \max\left\{|\ Z[l]\ |^2 \mid\ \omega_0 \le l\frac{\omega_s}{N} \le \omega_0 + 2\Delta\omega\right\} \tag{23}$$

$$E_{N-L} := \max\left\{|\ Z[k]\ |^2 |\ 0 \le k\frac{\omega_s}{N} \le \omega_0 \text{ or } \omega_0 + 2\Delta\omega \le k\frac{\omega_s}{N} \le \omega_s\right\} \tag{24}$$

where L denotes the number of DFT bins in the region $[\omega_0, \omega_0 + 2\Delta\omega]$, then an outlier will occur if $E_{N-L} > D_L$. Therefore, by denoting the outlier probability as q and using the technique of Rife and Boorstyn [6], we have

$$\begin{aligned} q &= \Pr\{D_L < E_{N-L}\} \\ &= \int_0^\infty \Pr\{D_L < E_{N-L} \mid E_{N-L} = x\} f_{E_{N-L}}(x)dx \\ &= \int_0^\infty F_{D_L}(x) f_{E_{N-L}}(x)dx. \end{aligned} \tag{25}$$

Furthermore, the probability distribution and density functions of E_{N-L} can be written as

$$F_{E_{N-L}}(x) = [F_{\mathcal{Z}_k}(x)]^{N-L} \tag{26}$$

and

$$f_{E_{N-L}}(x) = (N - L)[F_{\mathcal{Z}_k}(x)]^{N-L-1} f_{\mathcal{Z}_k}(x) \tag{27}$$

where $\mathcal{Z}_k = |\ Z[k]\ |^2 = Z_R[k]^2 + Z_I[k]^2$, and

$$F_{\mathcal{Z}_k}(x) = \left(1 - e^{-\frac{Nx}{2\sigma^2}}\right) \quad x \ge 0 \tag{28}$$

$$f_{\mathcal{Z}_k}(x) = \frac{N}{2\sigma^2} e^{-\frac{Nx}{2\sigma^2}} \quad x \ge 0. \tag{29}$$

Similarly, since the $\mathcal{Z}_l$'s are independent random variables when the DFT size is equal to the data size (where l is such that $\omega_0 \le l\omega_s/N \le \omega_0 + 2\Delta\omega$), then the probability distribution function of D_L can be written as

$$F_{D_L}(x) = \prod_{l=1}^{L} [F_{\mathcal{Z}_l}(x)]. \tag{30}$$

Note that $F_{\mathcal{Z}_l}(x)$ is a noncentral Chi-Square probability distribution as given in (21).

Now, the outlier probability for the frequency estimate of the linear FM signal can be written as

$$\begin{aligned} q = \int_0^\infty \prod_{l=1}^{L} F_{\chi'^2}\left(\frac{Nx}{\sigma^2}; 2, \frac{(S_R[l]^2 + S_I[l]^2)N}{\sigma^2}\right) \\ \times (N - L)\left(1 - e^{-\frac{Nx}{2\sigma^2}}\right)^{N-L-1} \frac{N}{2\sigma^2} e^{-\frac{Nx}{2\sigma^2}} dx. \end{aligned} \tag{31}$$

By change of variables, (31) can be written as

$$\begin{aligned} q = \int_0^\infty \left\{\prod_{l=1}^{L} F_{\chi'^2}\left(2y; 2, \frac{(S_R[l]^2 + S_I[l]^2)N}{\sigma^2}\right)\right\} \\ \times (N - L)(1 - e^{-y})^{N-L-1} e^{-y} dy. \end{aligned} \tag{32}$$

It is important to analyze the performance of the ML constant frequency estimators in terms of bias and variance.

However, when these estimators are used to estimate the frequencies of the linear FM signals, how to define the bias is not clear since there is no "true" frequency for linear FM signals. This inherent difficulty can be overcome by defining the "true" frequency as the mean frequency of the signal, which is $\omega_0 + \Delta\omega$. Then, the bias is defined by $\bar{\hat{\omega}} - (\omega_0 + \Delta\omega)$, where $\bar{\hat{\omega}}$ is the mean of the frequency estimates. Note that we defined outliers as the frequency estimates that are outside of the frequency region $[\omega_0, \omega_0 + 2\Delta\omega]$. Let us define

$$\bar{\hat{\omega}}_{\text{outlier}} = E\{\hat{\omega} \mid \text{outlier}\} \tag{33}$$

$$\bar{\hat{\omega}}_{\text{No outlier}} = E\{\hat{\omega} \mid \text{No outlier}\}. \tag{34}$$

Since the spectrum of the signal is symmetric around $\omega_0 + \Delta\omega$, the mean squared error of the frequency estimates can be calculated around the mean frequency of the signal $\omega_0 + \Delta\omega$ so that

$$\bar{\hat{\omega}}_{\text{No outlier}} = \omega_0 + \Delta\omega. \tag{35}$$

Since the noise that causes the occurrence of an outlier has a uniform distribution outside the frequency region $[\omega_0, \omega_0 + 2\Delta\omega]$, the mean of the frequency estimate when an outlier occurs can be expressed as

$$\bar{\hat{\omega}}_{\text{outlier}} = \frac{\omega_s^2 - 4\omega_0\Delta\omega - 4\Delta\omega^2}{2(\omega_s - 2\Delta w)} \tag{36}$$

where ω_s is the sampling frequency in radians.

Evidently, the bias is

$$\begin{aligned} \text{BIAS} &= \bar{\hat{\omega}} - (\omega_0 + \Delta\omega) \qquad (37)\\ &= q\bar{\hat{\omega}}_{\text{outlier}} + (1-q)\bar{\hat{\omega}}_{\text{No outlier}} - (\omega_0 + \Delta\omega)\\ &= q\left[\frac{\omega_s^2 - 4\omega_0\Delta\omega - 4\Delta\omega^2}{2(\omega_s - 2\Delta\omega)}\right] - q(\omega_0 + \Delta\omega)\\ &= q\frac{\omega_s[\omega_s - 2(\omega_0 + \Delta\omega)]}{2(\omega_s - 2\Delta\omega)}. \qquad (38) \end{aligned}$$

If $\omega_s/2 = \omega_0 + \Delta\omega$, the bias is zero. Further, the variance (computed around $\bar{\hat{\omega}}$) is given by

$$\begin{aligned} &\text{VARIANCE} =\\ &\quad qE\{(\hat{\omega} - \bar{\hat{\omega}}_{\text{outlier}})^2 \mid \text{outlier}\} + q(\bar{\hat{\omega}}_{\text{outlier}} - \bar{\hat{\omega}})^2\\ &\quad + (1-q)E\{(\hat{\omega} - \bar{\hat{\omega}}_{\text{No outlier}})^2 \mid \text{No outlier}\}\\ &\quad + (1-q)(\bar{\hat{\omega}}_{\text{No outlier}} - \bar{\hat{\omega}})^2. \end{aligned} \tag{39}$$

It is more relevant to consider the mean square error computed around the mean signal frequency $\omega_0 + \Delta\omega$ that includes the mean square error due to the bias. In other words

$$E\{(\hat{\omega} - (\omega_0 + \Delta\omega))^2\} = (\text{BIAS})^2 + (\text{VARIANCE}). \tag{40}$$

Note that

$$E\{(\hat{\omega} - \bar{\hat{\omega}}_{\text{No outlier}})^2 \mid \text{No outlier}\} \leq (\Delta\omega)^2 \tag{41}$$

and $E\{(\hat{\omega} - \bar{\hat{\omega}}_{\text{outlier}})^2 \mid \text{outlier}\}$ can be calculated easily. If $\omega_s/2 = \omega_0 + \Delta w$, since $\bar{\hat{\omega}} = \bar{\hat{\omega}}_{\text{outlier}} = \bar{\hat{\omega}}_{\text{No outlier}}$, i.e., there is no bias, then (39) simplifies to

$$\text{MSE} \leq q\left(\frac{1}{3}\omega_0^2 + \omega_0\Delta\omega + \Delta\omega^2\right) + (1-q)\Delta\omega^2. \tag{42}$$

If $\Delta\omega$ is sufficiently large, then an approximate expression for the MSE of the frequency estimates can be given since the Fourier transform of the signal can be approximated by a rectangular window that has width $2\Delta\omega$. Again, when there is no bias, this expression becomes

$$\text{MSE} \approx q\left(\frac{1}{3}\omega_0^2 + \omega_0\Delta\omega + \Delta\omega^2\right) + (1-q)\frac{\Delta\omega^2}{3}. \tag{43}$$

V. Comparison of the Theoretical Results and Simulations

In order to test the performance of the maximum likelihood constant frequency estimator, linear FM signals that had different chirp rates were estimated under the assumption that they had constant frequencies. The sampling frequency f_s was selected as 4000 Hz. In order to concentrate on the variance part of the mean squared error of the frequency estimates, the bias effects were eliminated by selecting the mean frequencies of the signals as $f_s/2$, i.e., $f_0 + \Delta f = 2000$ Hz. Note that since the signals are complex, the aliasing effect due to the sampling rate is negligible. The data sizes for the discrete-time measurement signals were selected as $N = 256$ and $N = 1024$ for the simulations and $N = 256$ for the theoretical calculations.

In the simulations, the fine search was implemented by using a combination of modified Newton/Raphson and bisection methods [13]. The frequencies of the linear FM signals were estimated at 30 different SNR levels and for each signal, 1000 different realizations of the measurement signal were used at every SNR level. The mean squared errors of the frequency estimates of the linear FM signals obtained from the simulations versus SNR are plotted in Fig. 3, where $N = 1024$. As can be seen from this figure, for the constant frequency case the threshold for the frequency estimates occurs when SNR is approximately equal to -16 dB. Below the threshold region, the root mean squared errors (RMSE's) of the frequency estimates are approximately 1000 Hz for all frequency variation Δf. Above the threshold, the RMSE's of the frequency estimates of the linear FM signals that have small frequency variations are slightly greater than the RMSE of the frequency estimates when the signal has a constant frequency. On the other hand, when the frequency variation is large, the RMSE is almost constant above the threshold. From Fig. 3, it can be also observed that the threshold SNR and the RMSE above the threshold increase when the frequency variation increases.

As is pointed out in [6] for constant frequency signals, an increase in the data length causes a decrease in the threshold SNR. This is also valid for linear FM signals as can be understood by comparing Fig. 3 and the simulation results of Fig. 4, where the simulation and the theoretical results are plotted. In Fig. 4, the data size N is equal to the DFT size M, which is 256. Another interesting point that can be observed by the comparison of the simulation results in Figs. 3 and 4 is that the signal length is important for the RMSE value when the frequency variation Δf is quite small. For example, when Δf is equal to 20 Hz, the RMSE of the frequency estimates is almost constant above threshold when

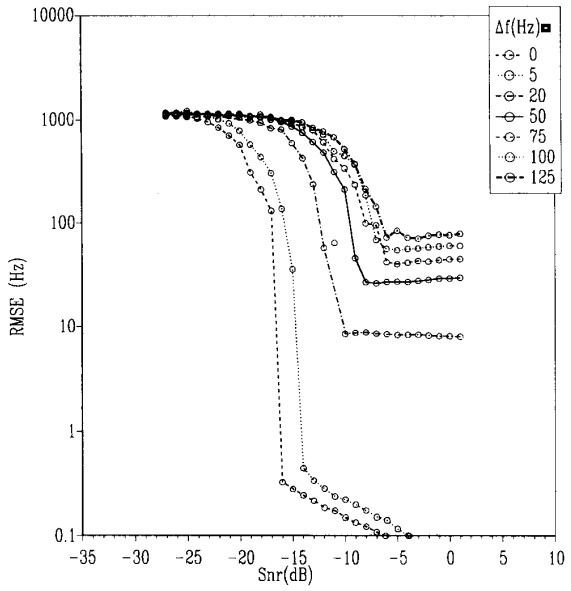


Fig. 3. Simulation results for the root mean squared error (RMSE) of the frequency estimates of the linear FM signals that have different chirp rates under the assumption that they have constant frequencies when the data size $N = 1024$, the DFT size $M = 1024$, the sampling frequency $f_s = 4000$ Hz, and the mean frequency $f_0 + \Delta f = 2000$ Hz.

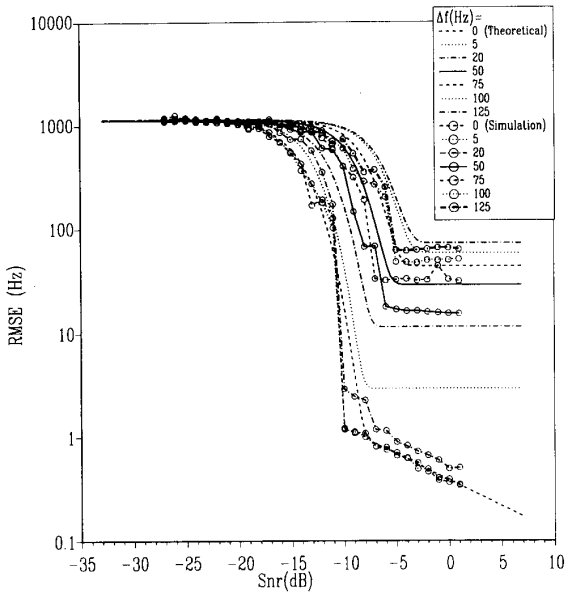


Fig. 4. Theoretical and simulation results for the root mean squared error (RMSE) of the frequency estimates of the linear FM signals that have different chirp rates under the assumption that thesignals have constant frequencies, when the data size $N = 256$, the DFT size $M = 256$, the sampling frequency $f_s = 4000$ Hz, and the mean frequency $f_0 + \Delta f = 2000$ Hz. For the theoretical RMSE calculation, the approximate MSE given in (43) is used. Note that if the bound for the MSE in (42) is used, the RMSE curves above the threshold SNR will shift up by $\sqrt{2}(\Delta f)/\sqrt{3}$.

$N = M = 1024$, whereas it is very close to the Cramer-Rao bound when $N = M = 256$. This dependence of the RMSE for small values of the frequency variation Δf can be better understood by Fig. 5. As can be seen from this figure, when the frequency variation Δf is relatively small, the maximum peaks of the absolute value of the DFT of the signal move to the edges of the rectangular window $[f_0, f_0 + 2\Delta f]$. However, for large values of Δf, the location of these peaks do not change very much.

A comparison of the simulation and theoretical results can be made using Fig. 4. From the theoretical curves of this figure, it can be observed that below the threshold, the RMSE is again approximately equal to 1000 Hz. Above the threshold, the RMSE increases as the frequency variation increases at a fixed SNR. The main difference between the theoretical curves and the simulation curves is the value of the RMSE of the frequency estimates above the threshold. As explained previously for the theoretical calculations, the RMSE above the threshold is approximately equal to $\Delta f/\sqrt{3}$, where this approximation is good when Δf is sufficiently large. Note that when the chirp rate is quite small, the discrete-time Fourier transform of the linear FM signal is very similar to the discrete-time Fourier transform of the signal when $\Delta f = 0$ Hz. In other words, the absolute value of the discrete-time Fourier transform of the linear FM signal cannot be approximated by a rectangle if the chirp rate is quite small. Thus, when the chirp rate is small, $\Delta f/\sqrt{3}$ is not a good approximation for the RMSE of the frequency estimates. A similar argument is valid for the bound given for the MSE of the frequency estimates in (42). In addition, this bound gets better for the signals that have larger data sizes since the peaks of the absolute value

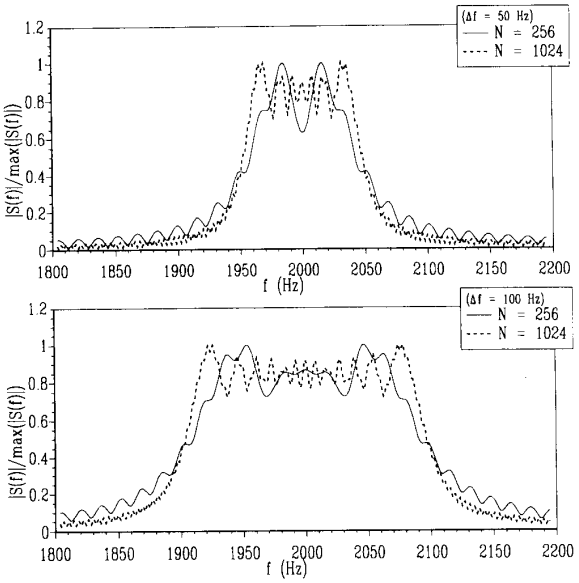


Fig. 5. Comparison of the normalized absolute value of the discrete Fourier transform (DFT) of linear FM signals that have different data lengths N and different chirp rates. The DFT size $M = 4096$, the sampling frequency $f_s = 4000$ Hz, and the mean frequency $f_0 + \Delta f = 2000$ Hz. Note that when the data length increases, the peaks of the absolute value of the DFT of the signal get closer to f_0 and $f_0 + 2\Delta f$.

of the discrete-time Fourier transform of a signal with a fixed Δf get closer to f_0 and $(f_0 + 2\Delta f)$ as the data size increases (see Fig. 5).

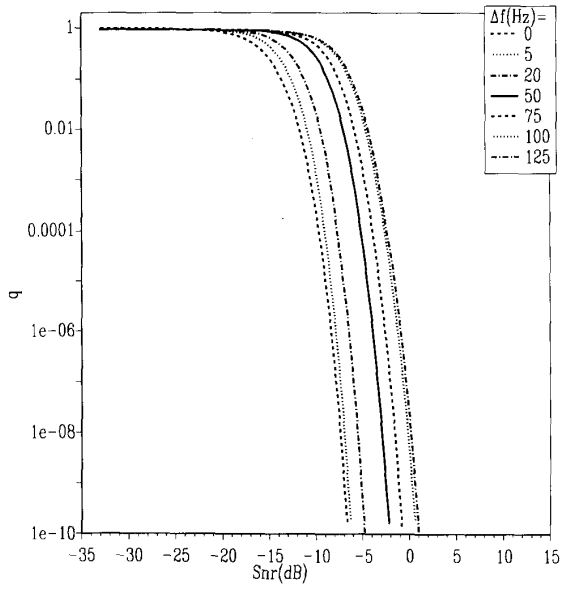


Fig. 6. Outlier probabilities of the linear FM signals for different chirp rates when the data size $N = 256$, the DFT size $M = 256$, the sampling frequency $f_s = 4000$ Hz, and the mean frequency $f_0 + \Delta f = 2000$ Hz.

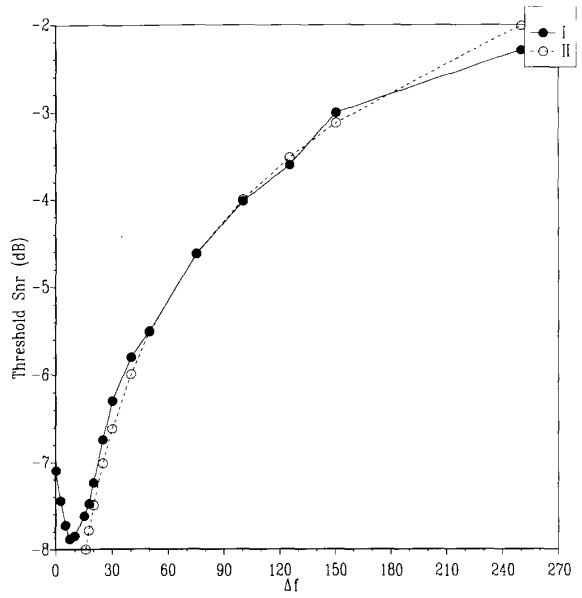


Fig. 7. Threshold SNR obtained from the theoretical calculations versus Δf plotted in curve I. ($10\log_{10}(\sqrt{\Delta f})$ + a constant) versus Δf is plotted in curve II, where the constant is -14. The parameters are selected such that the data size $N = 256$, the DFT size $M = 256$, the sampling frequency $f_s = 4000$ Hz, and the mean frequency$f_0 + \Delta f = 2000$ Hz.

Another difference between the theoretical and simulation curves is the threshold SNR's. In order to understand this, consider the outlier probability of the frequency estimates at different SNR's and different chirp rates (Fig. 6). Around the threshold SNR, the outlier probability is very small, for example, when $\Delta f = 0$ Hz, the outlier probability is approximately 10^{-6} (see Fig. 6). Thus, around 1 million realizations of the measurement signal have to be implemented just to observe one outlier. Since due to the computational limitations only 1000 realizations were implemented, the threshold SNR's of the simulation curves are expected to be smaller than the threshold SNR's of the theoretical curves. This also shows that simulations are impractical for finding the threshold SNR's accurately.

The increase of the RMSE (above the threshold) with an increase in the frequency variation is actually to be expected. Neglecting outliers, as we have agreed already, when we want to estimate $f_0+\Delta f$, the algorithm is likely to pick a frequency that is close to f_0 or $f_0 + 2\Delta f$ as the maximum (this being the approximate location of the maxima of $|S(f)|$). Thus, this error is likely to be Δf, even at very high SNR. Doubling of Δf doubles the root mean squared error of the frequency estimates if the frequency variation is sufficiently large. This can be observed in the simulations for the signals that have sufficiently large frequency variations.

The probability of occurrence of outliers is primarily influenced by the height of $\max_f |S(f)|$ above a notional noise floor, i.e., $[\max_f |S(f)|]/(\sigma^2/N)$. Thus, if Δf is doubled so that $\max_f |S(f)|$ is reduced by $\sqrt{2}$, one would expect the threshold to occur at an SNR that is $\sqrt{2}$ higher. Put another way, we should have a relation of the form

$$\text{SNR threshold in dB} = 10\log(\sqrt{\Delta f}) + \text{constant}. \quad (44)$$

This conjecture about the threshold SNR is supported by Fig. 7, where the first curve in the first graph corresponds to the threshold SNR's versus different frequency variations. The second curve in the first graph, which almost matches the first, is simply ($10\log(\sqrt{\Delta f})$ + a constant), where the constant for this data length is approximately -14. Of course, one cannot expect this relation to hold when Δf is very small. For $N = 256$, this relation holds for $\Delta f \geq 15$ Hz.

The constant in (44) clearly depends on the value of the data length N. However, this dependence is highly nonlinear and it is also related to the problem of finding the "optimal" data length for the ML constant frequency estimators. As shown in [3], the performance of a ML constant frequency estimator gets better as the data length increases if the signal has a constant frequency. However, if the signal is a linear FM signal with a constant chirp rate, then the frequency variation increases as the data length increases. Hence, there is a tradeoff between the asymptotic properties and the robustness of the ML constant frequency estimator. However, finding the "optimal" data length is not a trivial process. One should be able to obtain the RMSE versus SNR curves for different data sizes as is done in Fig. 4, where the data size N is equal to DFT size M, which is 256. Afterwards, by obtaining the threshold SNR versus Δf curves from the RMSE versus SNR curves, the dependence of the constant term in (44) on the data length can be found. Then, depending on the *a priori* selection of the desired threshold SNR and allowable frequency variation (i.e., the degree of robustness of the estimator), it is possible to find an optimal data length from these curves.

VI. Conclusion

In this paper, the robustness of the ML constant frequency estimator under wrong model assumptions was investigated. The motivation for the problem was to understand the performance of the HMM-ML tandem frequency tracker where the signal is divided into time blocks, and the frequency of the signal in each time block is estimated using the ML frequency estimation algorithm by assuming that the frequency of the signal is constant in those time blocks. For this purpose, a linear FM signal was estimated by assuming that it had a constant frequency over the measurement time.

In order to understand the performance of the ML constant frequency estimator, the outlier analysis of Rife and Boorstyn [6] for the constant frequency case was generalized for the linear FM signals. It has been shown that when the frequency variation is sufficiently large, even at very high SNR's, the RMSE of the frequency estimates of the linear FM signals is almost proportional to the frequency variation. Additionally, it is much greater than the RMSE of the frequency estimates for the constant frequency case. Using this fact, a bound and an approximate expression for the RMSE of the frequency estimates of the linear FM signals were given for above the threshold region.

Below the threshold, the RMSE for the linear FM signals was shown to be the same as the RMSE for the constant frequency case. A simple relation between the threshold SNR and the frequency variation was conjectured and supported by the theoretical curves. It was shown that the threshold SNR is a logarithmic function of the chirp rate of the linear FM signals. A discrepancy between the threshold SNR's of the simulation and the theoretical curves was observed and explained in terms of the very large number of realizations of the measurement signals required to determine the threshold region.

The work carried out in this paper has two aspects. The first one is robustness of the maximum-likelihood constant frequency estimators with respect to model errors in general. It has been shown that there is a tradeoff between the robustness and the asymptotic behavior of these estimators. The other aspect of this work is the performance of these estimators in the framework of the HMM-ML tandem frequency trackers, where the frequency of a signal is tracked by dividing the signal into segments, and the frequency of each signal segment is estimated using a ML constant frequency estimator. This analysis shows that the performances of ML constant frequency estimators depend on the deviation from the model assumptions. In [1], the transitions between the frequency estimates of the signal segments are modeled using a HMM. However, even though the ML constant frequency estimators may give bad frequency estimates, the HMM part of the tandem tracker may compensate these estimates depending on the performance of the ML estimators in the other signal windows. In particular, finding the overall threshold behavior of the HMM-ML tandem frequency tracker is an important open problem. The extension of this performance test to the HMM-ML tandem frequency trackers is an interesting research area.

Finally, in this paper, we have concentrated on the performance of the ML constant frequency estimators when they are used to estimate the frequencies of signals whose frequencies change linearly during the measurement intervals. It would be of interest to extend this work to other signals, the instantaneous frequencies of which change polynomially or randomly.

References

[1] R. L. Streit and R. Barrett, "Frequency line tracking using hidden Markov models," *IEEE Transactions Acoust, Speech, Signal Processing*, vol. 38, no. 4, pp. 586–598, 1990.

[2] A. M. Walker, "On the estimation of a harmonic component in a time series with stationary independent residuals," *Biometrika*, vol. 58, no. 1, pp. 21–36, 1971.

[3] E. J. Hannan, "The estimation of frequency," *J. Applied Probab.*, vol. 10, pp. 510–519, 1973.

[4] T. Hasan, "Nonlinear time-series regression for a class of amplitude modulated cosinusoids," *J. Time Series Anal.*, vol. 3, pp. 109–122, 1982.

[5] J. S. Rice and M. Rosenblatt, "On frequency estimation," *Biometrika*, vol. 75, pp. 477–484, 1988.

[6] D. C. Rife and R. R. Boorstyn, "Single-tone parameter estimation from discrete-time observations," *IEEE Trans. Inform. Theory*, vol. IT-20, no. 5, pp. 591–598, 1974.

[7] B. James, B. D. O. Anderson, and R.C. Williamson, "Characterization of threshold for single tone maximum likelihood frequency estimation," to be published in *IEEE Trans. Signal Processing*.

[8] P. M. Woodward, *Probability and Information Theory, with Applications to Radar*. New York: Pergamon, 1953.

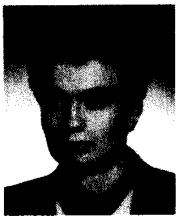
[9] P. J. Parker and B. D. O. Anderson, "Frequency tracking of non-sinusoidal periodic signals in noise," *Signal Processing*, vol. 20, pp. 127–152, 1990.

[10] B. Boashash, "Estimating and Interpreting the Instantaneous Frequency—Part 1: Fundamentals," *Proc. IEEE*, vol. 80, no. 4, pp. 520–538, 1992.

[11] A. Papoulis, *Systems and Transforms with Applications in Optics*. New York: McGraw-Hill, 1968.

[12] N. L. Johnson and S. Kotz, *Continuous Univariate Distributions-2*. New York: Wiley, 1972.

[13] W. H. Press, B. P. Flannery, S. A Teukolsky, and W. T. Vetterling, *Numerical Recipes in C, The Art of Scientific Computing*. Cambridge, UK: Cambridge University Press, 1988.



Mehmet Karan was born in Çankiri, Turkey, in 1964. He received the B.Sc. degree from the Middle East Technical University, Ankara, Turkey, in 1987 and the M.Sc. degree from Bilkent University, Ankara, Turkey, in 1990, both in electrical and electronics engineering. He is currently pursuing the Ph.D. degree at the Department of Systems Engineering, Australian National University.

His current research interests include signal processing and linear system theory.

Robert C. Williamson was born in Brisbane, Australia, in 1962. He received the B.E. degree from QIT in 1984 and the M.Eng.Sc. and Ph.D. degrees in 1986 and 1990, respectively, from the University of Queensland, all in electrical engineering.

In 1989, he was a lecturer at the University of Queensland. Since 1990, he has been at the Australian National University in the Department of Systems Engineering and the Department of Engineering, where he is a senior lecturer. His research interests include signal processing and neural networks.

Brian D. O. Anderson (S'62–M'66–SM'74–F'75) was born in Sydney, Australia, and received his undergraduate education at the University of Sydney, with majors in pure mathematics and electrical engineering. He subsequently received the Ph.D. degree in electrical engineering from Stanford University.

Following completion of his education, he worked in industry in Silicon Valley and served as an assistant professor in the Department of Electrical Engineering at Stanford. He was foundation professor of electrical engineering at the University of Newcastle, Australia, from 1967 till 1981 and is now professor of Systems Engineering at the Australian National University and Acting Director of the Research School of Information Sciences and Engineering. His interests are in control and signal processing.

Dr. Anderson is a Fellow of the Royal Society, the Australian Academy of Science, the Australian Academy of Technological Sciences and Engineering, and the Institute of Electrical and Electronic Engineers. He is an Honorary Fellow of the Institution of Engineers, Australia. He holds doctorates (honoris causa) from the Université Catholique de Louvain, Belgium and the Swiss Federal Institute of Technology, Zürich. He served a term as President of the International Federation of Automatic Control from 1990 to 1993.