THE EFFECT OF QUADRATURE HYBRID ERRORS ON A PHASE DIFFERENCE FREQUENCY ESTIMATOR AND METHODS FOR CORRECTION

John Kitchen, B.Sc.(Hons)

A thesis submitted in part fulfilment of the regulations for
the degree of Master of Applied Science (Communications)
in the Department of Applied Mathematics, University of Adelaide

November 1991

CONTENTS

TABLES

FIGURES

APPENDICES

SUMMARY

This thesis is concerned with the estimation of the frequency of a single sinusoid of fixed but unknown frequency and which has been corrupted by errors in a quadrature hybrid, coherent receiver system. The quadrature hybrid receiver has been employed to produce in-phase and quadrature baseband signal components which can be sampled and undergo an analogue to digital conversion. A simple, discrete frequency estimator is derived from the rate of change of signal phase between successive sampling instants after analogue to digital conversion. The statistical effects of the errors and imbalances, inherent in the quadrature hybrid, upon the discrete frequency estimator are studied. This study has been carried out in four stages forming the content of the four chapters : 2,3,4,5.

In chapter 2, an analytic study of the estimation of the frequency of a single sinusoid, which has passed through a quadrature hybrid system, is carried out. This study is further subdivided so that each of the quadrature hybrid errors is examined individually. A summary of the results derived for each error case is provided, in tabular form, in appendix XII.

In chapter 3, the quadrature hybrid and input signal are modelled as a computer simulation. This chapter is subdivided, as with chapter 2, so that each error case is simulated on an individual basis. In each case the error or imbalance is varied and the frequency of the input sinusoid is varied so that most of the possible error conditions and possible input frequencies are studied. Simulation results are presented in graphical form and compared with a similar graphical presentation of the theoretical results from chapter 2.

In chapter 4, a real quadrature hybrid, receiver system is examined and the inherent system errors are measured. These measurements serve to support both the simulations and the theory.

In chapter 5, techniques are derived in order to reduce the degradation of frequency estimation caused by the quadrature hybrid system errors and both simulation results and a real example are given. It is also demonstrated that the theoretical lower limit for frequency estimation in the presence of normally distributed noise (the Cramér-Rao lower variance bound) can be acheived for this system.

DECLARATION

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

J. KITCHEN

ACKNOWLEDGEMENTS

This work began as a project report but regulations required that a thesis be presented. I wish to express my sincere gratitude to Dr. Bruce Davis, for his role as project supervisor and for his help in making this transition from project report to thesis possible.

I am also indebted to Dr. Ken Sarkies for his efforts in clearing the way through the administrative quagmire which presented itself during the course of my study and project work.

Thanks are due to my employer, the Defence Science and Technology Organisation, Salisbury, for taking care of the financial aspects of this degree.

I would also like to thank Dr. D. Gray of EWD, DSTO, for some useful discussions in the early stages of this work.

This thesis was prepared using the LaTeX document preparation system which is a public domain computer typesetting program [1] which I believe to be the best of its kind currently available.

NOTATION

Symbol	Explanation
I or I(t)	The in-phase or REAL component of the analytic signal
Q or Q(t)	The quadrature or IMAGINARY component of the analytic signal
$\omega = 2\pi f$	Angular frequency measured in radians per second
θ or $\theta(t)$	The phase of the signal at time t
$\Delta\theta(t)$	The difference in signal phase between time t and time (t-T) where T is the sampling period
$y(t)$	The analytic signal at time t, $y(t) = I(t) + jQ(t)$
$\angle y(t)$	The phase or argument of the analytic signal at time t
$f(t)$	The frequency of the input sinusoid at time t, measured in Hz
θ_n	The phase of the signal at sample point n
$\Delta\theta_n$	The difference in signal phase between point n and point (n-1)
$\angle y_n$	The phase or argument of the analytic signal at point n
f_n	The frequency of the input sinusoid at point n
$TRUNC[x]$	An operator which removes the fractional part of a real value x leaving only the integer part
$\widehat{\Delta\theta}$	Expected value of phase difference taken over N phase difference data points
θ_f	Angle through which the complex signal vector rotates, ie phase change of signal, between sample instants
δ_i	DC offset in the I channel
δ_q	DC offset in the Q channel
γ	Relative amplitude gain error, $\frac{Q}{I} = 1 + \gamma$
ϵ	Quadrature phase error
N	The number of phase difference samples available
ρ	The signal-to-noise power ratio, $\rho = \frac{A^2}{\sigma^2}$
θ_e	The error in the signal phase
$E\{x\}$ or $\widehat{x}$	The mean or expected value of variable x, $E\{x\} = \int_{-\infty}^{\infty} xp(x)dx$
$p(x)$	The probability density of x
$Var\{x\}$	The variance or mean square error in variable x, $Var\{x\} = E^2\{x - \hat{x}\} = E\{x^2\} - E^2\{x\}$

$Var\{\Delta\theta\}$	The mean square error, or variance, in a single phase difference data point
$Var\{\widehat{\Delta\theta}\}$	The mean square error, or variance, in the expected value of phase difference taken over N data points. We use $mse\{\Delta\theta\}$ and $Var\{\widehat{\Delta\theta}\}$ in the figures as an aid to distinguish between the two variables : $\Delta\theta$ and $\widehat{\Delta\theta}$
σ_i^2	The variance of the noise in the I channel
σ_q^2	The variance of the noise in the Q channel
σ^2	Total variance due to the noise in the I and Q channels, $\sigma^2 = \sigma_i^2 + \sigma_q^2$
A	The signal amplitude after passing through the quadrature hybrid
θ_{emax}	The maximum phase error due to the system errors
R	The amplitude of resultant complex vector ie signal plus error
R_{max}	The maximum value of R when signal combined with error
R_{min}	The minimum value of R when signal combined with error
z	The amplitude of the complex noise vector
L	The order of the adaptive filter
μ	The convergence parameter of the adaptive filter

1 INTRODUCTION

1.1 The quadrature hybrid receiver

In many signal processing based systems, for example modern sonar, radar and communication systems using digital signal processing, an analogue operation is performed which produces an in-phase (I) and a quadrature (Q) component of a real input signal. In this case the input signal is a single (one frequency component) narrowband signal of fixed frequency. An intermediate frequency signal at frequency $\omega_0 + \omega_1$, with an amplitude A, is split and mixed with coherent I and Q reference sources given by $\cos\omega_0 t$ and $-\sin\omega_0 t$, respectively. After mixing the products are low-pass filtered to retain the difference frequency terms. This results in the signal being split into in-phase and quadrature components, which are given by

$$\begin{aligned} I(t) &= A\cos\omega_1 t \\ Q(t) &= A\sin\omega_1 t\ . \end{aligned} \tag{1}$$

This pair of signals can then be treated as a complex signal

$$y(t) = I(t) + Q(t) = A\exp\{j\omega_1 t\}\ . \tag{2}$$

This ideal signal is obtained only if the gains of the I and Q paths are equal and if the phase difference between channels is 90 degrees. The resulting pair of signals are then sampled and quantised by matched analogue-to-digital (ADC) converters. This representation of digitised signals by complex numbers has a certain mathematical convenience and allows us to resolve positive and negative frequencies, relative to the reference frequency. From these complex numbers we can derive information about the input signal such as amplitude, phase and frequency at any given sample instant. Figure 1 shows the main components involved in a receiver system employing a quadrature heterodyne process. These are :-

(a) Superhet receiver. The superhet has the ability to scan or sweep the radio frequency (RF) spectrum of interest. The superhet converts these RF signals to an intermediate frequency (IF) which can more easily be handled by IF analogue circuitry. Ideally a bandpass filter is inserted at this stage to constrain the bandwidth of the IF signals to be within the A/D sampling rate (Nyquist frequency). In the actual system we have relied upon the output bandwidth of the receiver to act as a suitable bandpass filter.

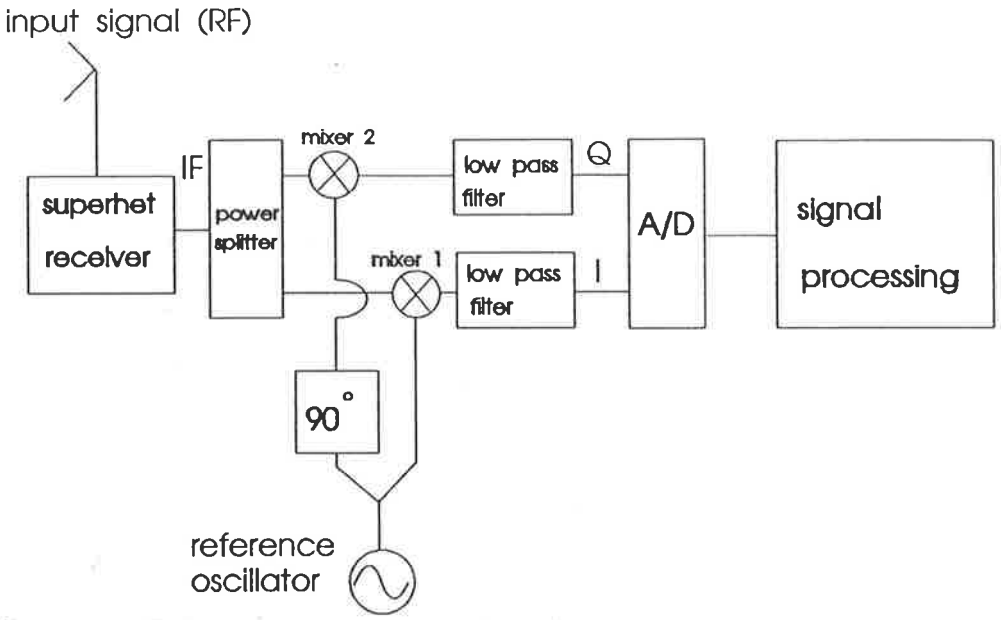

Figure 1 Block diagram showing main components comprising the quadrature heterodyne receiver and digital signal processing involved in this study.

(b) Quadrature mixing unit. In this frequency mixing stage the IF signal output from the superhet receiver passes through a power splitter whereupon each output is mixed with : 1 a fixed frequency signal from a stable reference oscillator, 2 a fixed frequency signal from the same reference oscillator but delayed by 90 degrees phase with respect to that of 1. The output from mixer 1 contains frequency components involving the sum and difference between the IF signal frequency and the reference oscillator signal frequency (in-phase or I channel). The output from mixer 2 is the same as 1 but with the 90 degree relative phase shift (quadrature or Q channel, this is strictly -Q since a phase delay has been incurred). Low-pass filters follow each of the two mixers in order to remove the summation frequency components prior to sampling. The output from mixer 1 is commonly referred to as the REAL component of the IF signal and the output from mixer 2 is commonly referred to as the IMAGINARY component of the IF signal.

(c) Digital sampling and recording. The I channel and Q channel are next sampled, in this case a high speed digital sampling oscilloscope (DSO) is used to carry out the A/D conversion on each channel. The DSO sampling rate for each channel is 400MHz, which implies a sampling period of 2.5 nanoseconds. 8-bit A/D converters are employed thus dividing the amplitude range into 256 levels.

1.2 Discrete frequency estimation

Having sampled and recorded waveforms of interest we can carry out the appropriate signal processing methods in order to determine the IF signal frequency. The method applied here is to calculate the phase of the signal at each sampling instant in time, then calculate the change in phase between each successive sampling instant

$$\theta(t) = \angle y(t) = \tan^{-1}\left[\frac{Q(t)}{I(t)}\right] . \tag{3}$$

The phase difference between successive sampling instants is calculated as

$$\Delta\theta(t) = \theta(t) - \theta(t - T) \tag{4}$$

where T is the intersample period. Hence frequency is given by

$$f(t) = \frac{1}{2\pi}\left[\frac{\Delta\theta(t)}{T}\right] . \tag{5}$$

For our discrete sampling system we shall use n in place of t and normalise T to be unity, hence

$$\theta_n = \angle y_n = \tan^{-1}\left[\frac{Q_n}{I_n}\right] \tag{6}$$

$$\text{and } \Delta\theta_n = \theta_n - \theta_{n-1} \tag{7}$$

$$\text{and } f_n = \frac{1}{2\pi}[\Delta\theta_n] . \tag{8}$$

One problem which might be encountered in the above calculation of phase difference occurs when the absolute value of phase difference between successive samples is greater than π radians. This problem may be overcome if the Nyquist criterion can be satisfied, namely that the sampling rate should be at least twice the frequency of the highest-frequency component

of the signal. Alternatively we can avoid this condition by restricting the signal to be recorded to be within the sampling bandwidth of the system. This is achieved through the use of low pass filters before the A/D stage. For example, if the A/D sampling rate is 400MHz, then each low pass filter should have a cut-off point at 200MHz. Thus, in our example, we sample an input signal frequency range of ±200MHz with an A/D sampling rate of 400MHz. We note that in practical filter design the low pass filters may still pass some frequencies beyond 200MHz, though rapidly attenuated as frequency increases. If it is desired to filter out signal frequencies above 200MHz then the filters might have to be designed to have a lower pass frequency eg 175MHz leading to a reduction in the input signal frequency range to ±175MHz.

1.3 Applications of the quadrature hybrid reciever

The quadrature hybrid can be found in many radar and communications applications. For examples of radar applications see Hovanessian [2], Tsui [3]. In [3] the quadrature hybrid is used as part of a digital microwave receiver for electronic warfare purposes. In communications systems the quadrature hybrid can be used as a frequency detector see Bellanger [4]. The quadrature hybrid behaves as a Hilbert transformer which is described in most good books on signal processing applications eg Papoulis [5]. In 1988, Naegeli and McHenry [6] presented some applications of IQ vector modulation for signal simulation and analysis.

1.4 Aims of the thesis

In any practical quadrature heterodyne system there will be system errors and imbalances which will inhibit the accuracy with which the signal parameters can be measured. In this study we are particularly interested in obtaining an accurate estimate of the frequency of a input sinusoid, of fixed but unknown frequency, using the very simple frequency estimator described in section 1.2 and also the relatively simple frequency estimator developed by Kay [7]. The effects of the main system errors : dc offsets, amplitude imbalance, phase error, noise and quantisation, upon the statistics of the phase difference frequency estimator, are studied both analytically and via computer simulations. The phase error is the error between the intended phase difference of 90 degrees and the actual phase difference between the I and Q signal paths. Real system errors have been measured and used to validate the analytic results.

1.5 Previous work on quadrature hybrid systems

There are several articles which report work carried out on quadrature hybrid systems and sampling of the I and Q waveforms. In 1974, Sinsky and Wang [8] looked at the output signal distortion with phase and gain errors and derived the statistics of the output signal voltage envelope. The following year, Persons [9] determined the ratio of the power in the error frequency bands to the power in the signal frequency band, for any sampling frequency. Brown [10], in 1979, introduced a quadrature sampling theorem and looked at ways of minimising the sampling rate for quadrature sampling. In 1981, Churchill et al [11] described a technique for correcting the gain, phase imbalances and bias errors of a quadrature system. This technique requires an input test signal and the derivation of digital filter coefficients. The residual errors using this technique are derived from measurements of the output of an image filter. Waters and Jarrett, in 1982, developed a method of using a single A/D converter to digitise the input signal IF directly and producing I and Q components thus eliminating baseband conversion. Though producing phase errors of the order of 2 degrees, this method was designed to operate at low frequencies ie 2.5MHz. Rice and Wu [12] , in 1982, employed a Hilbert transformer to produce quadrature outputs from an input bandpass signal and developed a hardware implementation with a 60kHz bandwidth.

1.6 Main results

Much work has been previously carried out in investigating the effects of quadrature hybrid system errors. These investigations are based on theoretical expressions or measurements of the output frequency spectrum, in particular the image frequency components. The main aim of this thesis is to derive expressions for the statistical accuracy obtainable in estimating the frequency of an input sinusoid, of fixed but unknown frequency, and which is corrupted by errors in a quadrature hybrid, coherent receiver system. These expressions are derived analytically and computer simulations are carried out in order to support the theory. A hardware implementation of the quadrature hybrid is also studied, measurements of the inherent system errors are made and a comparison with the theoretical results for the variance in the estimated are made.

The second aim of this thesis is to investigate methods for reducing the effects of the inherent system errors, with a view to real time implementation.

1.7 Note on figure labelling

Due to a limitation in the in software package [13] used to plot the results, it is not possible to correctly label the plots. To circumvent this problem we use the following representations

(i) mse or mse{pd} $\equiv Var\{\Delta\theta\}$.

(ii) var or var{pd} $\equiv Var\{\widehat{\Delta\theta}\}$.

(iii) sn or snr $\equiv$signal-to-noise ratio ρ.

2 ANALYTIC STUDY OF ERRORS

In this chapter we examine the mean value of the phase difference error, the variance in the phase difference and the variance of the phase difference frequency estimator in the presence of each of the quadrature heterodyne system errors. This is most readily achieved by considering the effects upon the complex vector representing the IF signal. The detailed analyses for all of the error cases can be found in the appendices.

2.1 Signal plus noise case

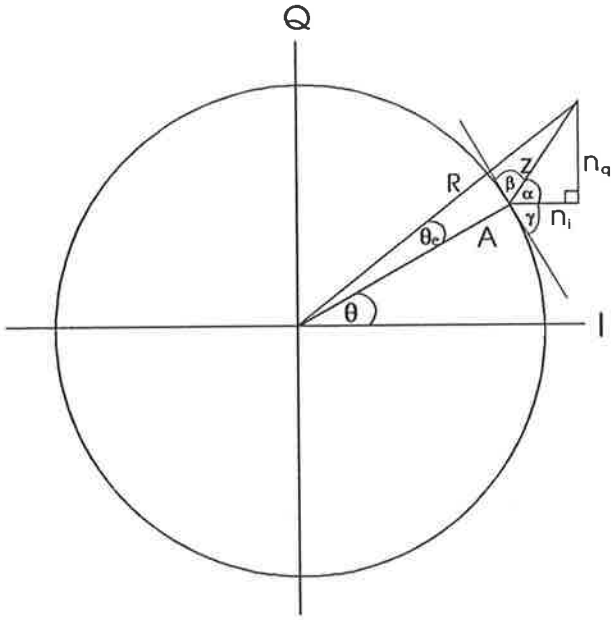

Figure 2 Complex signal vector - showing effect of additive complex noise vector.

We consider the case where each of the I and the Q channels are independently corrupted by normally distributed white noise, shown as the complex vector z in figure 2. The two channel waveforms may then be represented as

$$\begin{aligned} I &= A\cos\theta + n_i \\ Q &= A\sin\theta + n_q \end{aligned} \tag{9}$$

where $\theta = \omega_1 t$ is the signal phase at sample instant t and A is the signal amplitude, n_i is the noise in the I channel and n_q is the noise in the Q channel. We assume firstly that the noise input to the receiver has come from a multiplicity of independent sources which may have different probability distributions for the noise amplitude. Similarly sources of noise within the receiver system itself (mixer diodes, receiver components etc) are assumed to be independent and their noise distributions may be different. We invoke the central limit theorem which states that if different random variables are independent then the probability density function of their sum tends to a normal curve as the number of sources tends to ∞. In our receiver system we consider the number of noise sources to be sufficiently large that the central limit theorem is approached. Hence n_i and n_q are treated as white Gaussian noise.

In previous studies eg [7] n_i and n_q are treated as real Gaussian random variables with zero means and uncorrelated with each other. However the assumption that n_i and n_q are uncorrelated is dubious in this case and, almost certainly, the noise is not independent. The complex noise vector formed from n_i and n_q will still have a zero mean and be uniformly distributed in [0,2π] for high enough signal-to-noise ratios. Hence our derivations are unaffected by the correlation of n_i and n_q.

In figure 2 the amplitude of the input signal vector is A and θ is the phase of the signal vector. The complex vector due to the noise in each of the I and Q channels is shown as z. The instantaneous phase error due to the error in each of the I and Q channels is given by θ_e. The results are summarised in table XII.1. We note that :-

(a) the phase difference error is a random variable with zero mean
(b) $Var\{\Delta\theta\}$ is inversely proportional to the signal-to-noise ratio ρ
(c) for a set of N consecutive phase difference data points the variance
is also inversely proportional to N^2.

2.2 Quantisation case

We consider the case where each of the I and the Q channels are independently quantised by the A/D sampling process.

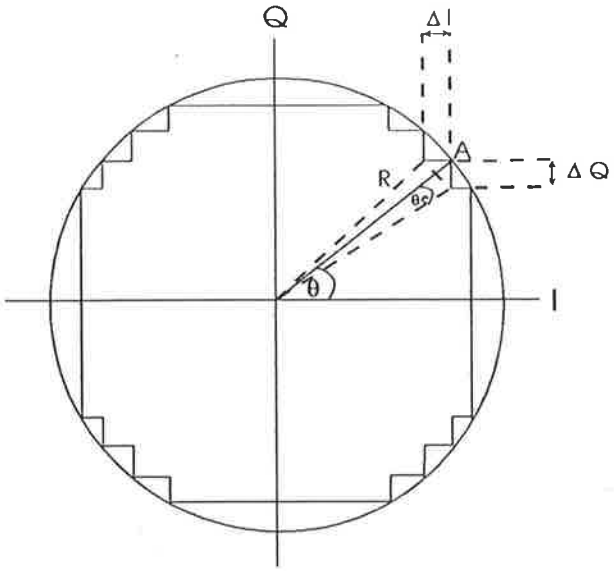

Figure 3 Complex signal vector - showing discretisation due to A/D process.

The two channel waveforms may then be represented as

$$\begin{aligned} I &= TRUNC[A\cos\theta] \\ Q &= TRUNC[A\sin\theta] \end{aligned} \tag{10}$$

where $TRUNC$ is an operator which the truncates the waveform sample values to the nearest integer value below the signal level. Hence the complex signal vector can only assume certain discrete values in I and Q (see figure 3) and so discrete values of phase. This will have the effect of producing a phase error at each sample instant which, in turn, will produce an error in the phase difference estimator. The results are summarised in table XII.2. We note that :-

(a) the phase difference error is a zero mean process

(b) $Var\{\Delta\theta\}$, for a single phase difference data point is inversely proportional to the square of the input signal amplitude A

(c) for a set of N consecutive phase difference data points the variance is also inversely proportional to N^2.

2.3 Phase error case

We consider the case where the Q channel is not exactly 90 degrees relative phase to the I channel, leading to distortion of the signal complex circle as shown in figure 4. One source of phase error is the path difference between the two reference oscillator channels. The 90 degree phase shift in the Q channel may be acheived by designing this channel to have a path length one quarter wavelength longer than the I channel or by using a 90 degree hybrid power splitter designed to operate at the reference oscillator frequency. In either case the signal path lengths may not be correct leading to a fixed phase error. For example, if the reference oscillator has a frequency of 1GHz and there is a path length error of 1cm then there will be a phase error of 12 degrees or 0.21 radians. Another source of phase error is the difference between the low pass filters. Each filter will have a phase error which is dependent on the signal frequency and, if the two filters are not properly matched, each filter will have a different error versus frequency response. The two channel waveforms may then

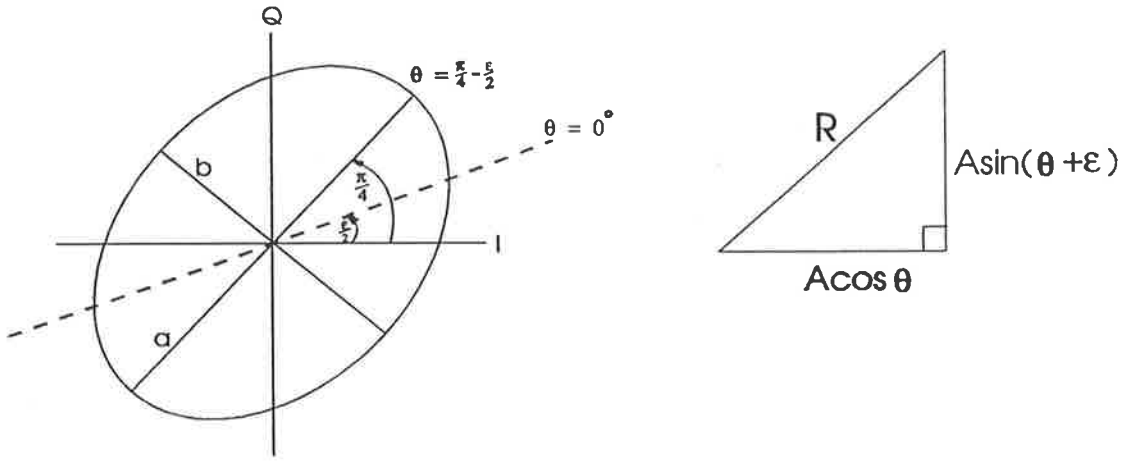

Figure 4 Complex signal vector - showing distortion due to quadrature phase error.

be represented as

$$\begin{aligned} I &= A\cos\theta \\ Q &= A\sin(\theta+\epsilon) \end{aligned} \tag{11}$$

where ϵ is the quadrature phase error. The error in the signal phase is given by

$$\theta_e = \theta - \tan^{-1}\left[\frac{sin(\theta+\epsilon)}{cos\theta}\right] . \tag{12}$$

This is the error in phase for the signal vector in any of the four complex quadrants.

If we plot I versus Q, for each sample instant, then the quadrature phase error produces a characteristic ellipse with its major axis tilted at an angle of 45 degrees to the I axis. The lengths of the major and minor axes of the ellipse are determined by the quadrature phase error. The signal phase angle and angles on the IQ plot are no longer the same. In fact the signal phase angle $\theta = 0$ degrees now corresponds to an angle of $\frac{\epsilon}{2}$ relative to the I axis. The results are summarised in table XII.3. We note that :-

(a) the phase difference error is a deterministic function with zero mean
(b) $Var\{\Delta\theta\}$ depends upon both the maximum phase error θ_{emax}
and the signal frequency θ_f
(c) for a set of N consecutive phase difference data points the variance also
depends upon the number of data points.

2.4 Amplitude imbalance case

We consider the case where the Q channel is not equal in amplitude to the I channel. With this type of error the signal complex vector describes an ellipse with its major axis aligned with the axis of highest relative gain, as shown in figure 5. The two channel waveforms may then be represented as

$$\begin{aligned} I &= A\cos\theta \\ Q &= A(1+\gamma)\sin\theta . \end{aligned} \tag{13}$$

The error in the signal phase is given by

$$\theta_e = \theta - \tan^{-1}\left[\frac{(1+\gamma)\sin\theta}{\cos\theta}\right] . \tag{14}$$

This is the error in phase for the signal vector in any of the four complex quadrants. The results are summarised in table XII.4. We note that :-

(a) the phase difference error is a deterministic function with zero mean

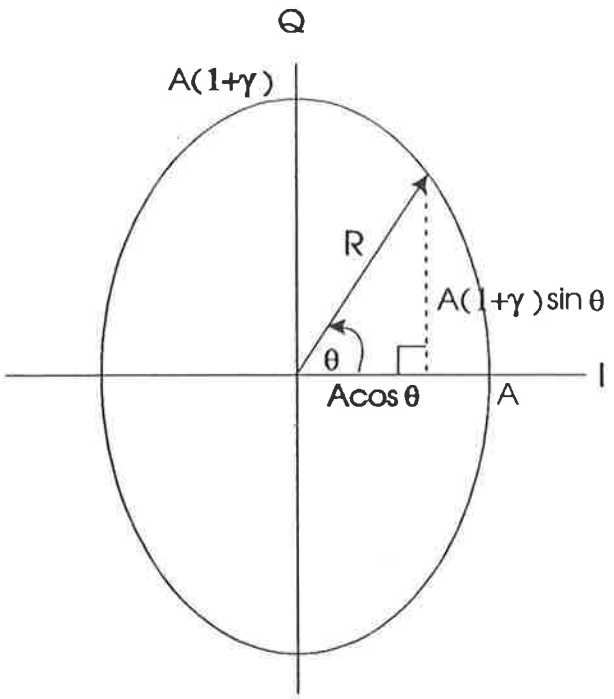

Figure 5 Complex signal vector - showing distortion due to amplitude imbalance between I and Q channels.

(b) $Var\{\Delta\theta\}$ depends upon both the maximum phase error θ_{emax} and the signal frequency θ_f

(c) for a set of N consecutive phase difference data points the variance depends upon the number of data points.

2.5 DC offset case

We consider the case where the Q channel and the I channel are offset from the origin. In this case the origin of the signal complex vector is shifted (see figure 6) thus causing an error in the calculated signal amplitude and phase. The two channel waveforms may then be represented as

$$\begin{aligned} I &= A\cos\theta + \delta_i \\ Q &= A\sin\theta + \delta_q \ . \end{aligned} \tag{15}$$

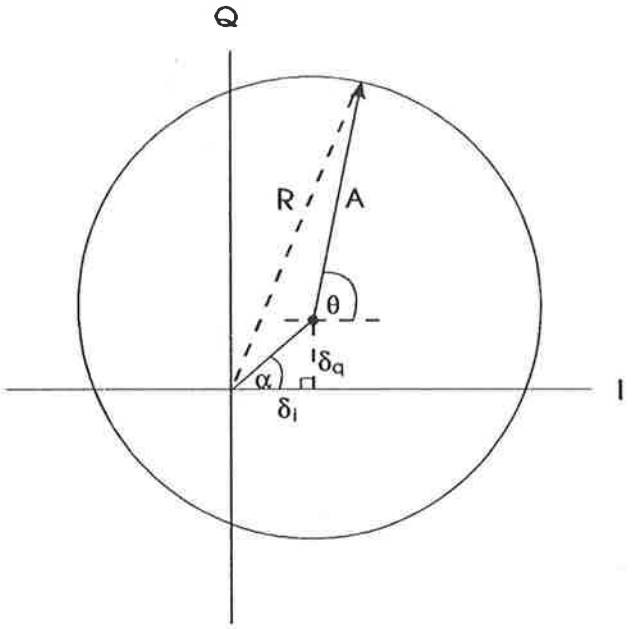

Figure 6 Complex signal vector - showing effect of dc offsets in I and Q channels.

The error in the signal phase is given by

$$\theta_e = \theta - \tan^{-1}\left[\frac{A\sin\theta + \delta_q}{A\cos\theta + \delta_i}\right] . \tag{16}$$

This is the error in phase for the signal vector in any of the four complex quadrants. The results are summarised in table XII.5. We note that :-

(a) the phase difference error is a deterministic function with zero mean

(b) $Var\{\Delta\theta\}$ depends upon both the maximum phase error θ_{emax} and the signal frequency θ_f

(c) for a set of N consecutive phase difference data points the variance also depends upon the number of data points.

3 COMPUTER SIMULATIONS

Simulations were carried out in the following manner :-

(i) 400 point vectors for each of the REAL signal component and the IMAG signal component were generated with the appropriate error incorporated in each vector. The initial phase for the signal was randomly chosen using a Pascal function for uniformly distributed number generation. The initial phase for the signal was therefore uniformly distributed in $[0, 2\pi]$.

(ii) the signal phase at each sample point was calculated then the difference in phase between successive sample points was calculated forming the phase difference vector.

(iii) the mean value of the phase difference vector was next calculated

(iv) the variance, $Var\{\Delta\theta\}$, in the phase difference vector was calculated

(v) steps (i) to (iv) were repeated thirty times for each parameter change such as signal-to-noise level or frequency, thus giving thirty values of mean phase difference. Following this the variance in the mean value of the phase difference, $Var\{\widehat{\Delta\theta}\}$, could be calculated.

3.1 Signal plus noise case

Figure 7 shows the results from simulating a signal with various levels of normally distributed random noise added to each of the two signal vector components. The routine used to generate the Gaussian noise was based upon the Box-Muller method listed in *Numerical Recipes* [14] (see also appendix VIII). The random noise was added independently to each of the signal vector components. Shown also are the theoretical values for the mse or variance. Note that the variance is a good fit even for low signal-to-noise ratios but the variance is a poor fit below 6dB. The Cramer-Rao lower variance bound (CRLB) gives the smallest error variance that can be attained by an unbiased estimator - an introduction to this can be found in lesson 6 of the book by Mendel [15]. Rife and Boorstyn [16] have derived the CRLB for a maximum-likelihood (ML) estimator of frequency in noise and Kay [7] has shown that his phase difference estimator achieves the same CRLB at moderately high SNR. Kay [7] shows

that the phase difference values form a moving average (MA) process with coefficients 1 and -1. Next the covariance matrix for the moving average process is formed and Kay derives an estimator which minimises the variance for this matrix. The frequency estimator is equivalent to a set of weights applied to the phase difference data vector. The Kay weighting function has been applied to both the simulations and to the experimental results (see later). Figure 8 shows that the Cramer-Rao lower variance bound is indeed achieved for this simulation of signal plus noise.

3.2 Quantisation case

Figure 9 shows the results from simulations of a fixed frequency signal which has had its REAL and IMAG components quantised or truncated to the nearest integer values. The frequency of the signal was 0.157 rads/sample interval (where frequency band is defined between $-\pi$ and $+\pi$ rads per sample interval) or equivalently, for a sampling rate of 400MHz the frequency is 10MHz. Both $Var\{\Delta\theta\}$ and $Var\{\widehat{\Delta\theta}\}$ are a good fit, even for low signal amplitude. The Kay weighting function was applied in this case to study its effect, even though it was an optimisation developed for the noise only case. In the simulations an improvement by a factor of approximately $\frac{1}{N}$ was obtained (figure 10).

3.3 Phase error case

Figures 11, 12, 15, 16 show the results of simulating a phase error in the 90 degree quadrature hybrid. Comparing these plots with the theoretical values shown in figures 13, 14, 17, 18 we can see there is a very good match in both $Var\{\Delta\theta\}$ and $Var\{\widehat{\Delta\theta}\}$. The Kay weighting function has also been applied in the simulations (figure 35), even though it is not appropriate for this type of error. An improvement in the variance in the mean phase difference was achieved except at very low signal frequencies.

3.4 Amplitude imbalance case

Figures 19, 20, 23, 24 show the results of simulating an amplitude imbalance between the REAL and IMAG signal vectors . Comparing these plots with the theoretical values shown in figures 21, 22, 25, 26 we can see there is a very good match in both $Var\{\Delta\theta\}$ and $Var\{\widehat{\Delta\theta}\}$.

The Kay weighting function has also been applied in the simulations (figure 36), even though it is not appropriate for this type of error. An improvement in the variance in the mean phase difference was achieved except at very low signal frequencies.

3.5 DC offset case

Figures 27, 28, 31, 32 show the results of simulating a fixed dc offset in both the REAL and IMAG signal vectors . Comparing these plots with the theoretical values shown in figures 29, 30, 33, 34 we can see there is a very good match in both the $Var\{\Delta\theta\}$ and $Var\{\widehat{\Delta\theta}\}$. The Kay weighting function has also been applied in the simulations (figure 37), even though it is not appropriate for this type of error. An improvement in the variance in the mean phase difference was achieved except at very low signal frequencies.

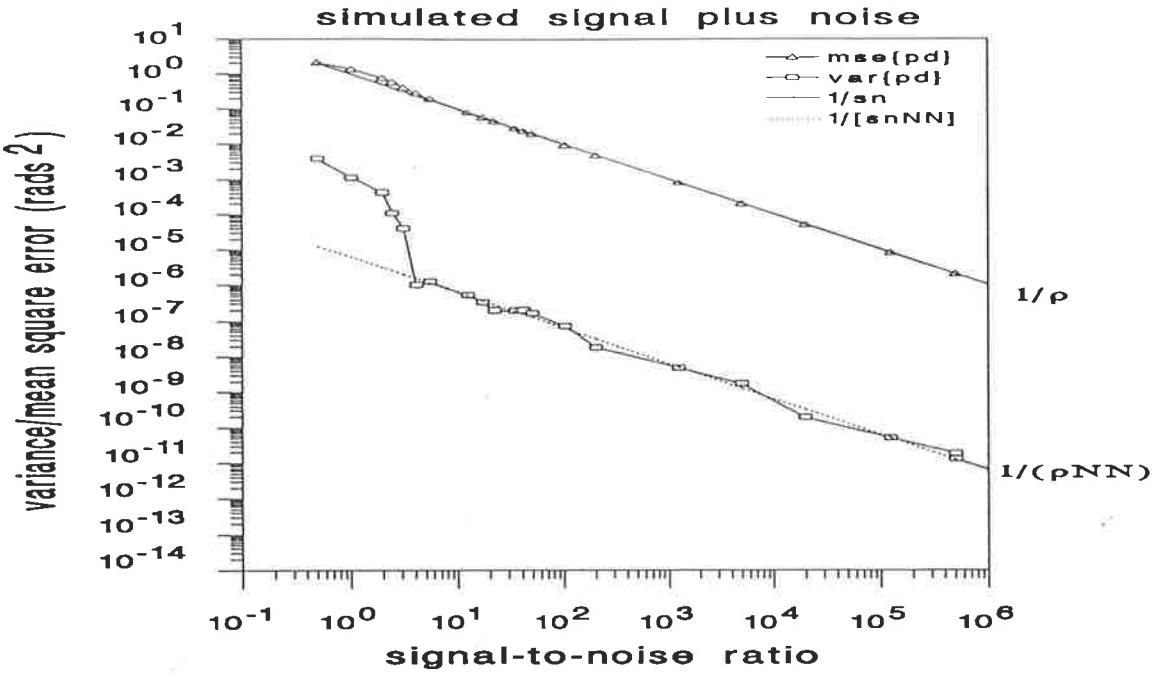

Figure 7 Simulation results, signal plus noise case - showing $Var\{\Delta\theta\}$ and $Var\{\widehat{\Delta\theta}\}$ compared with theoretical results.

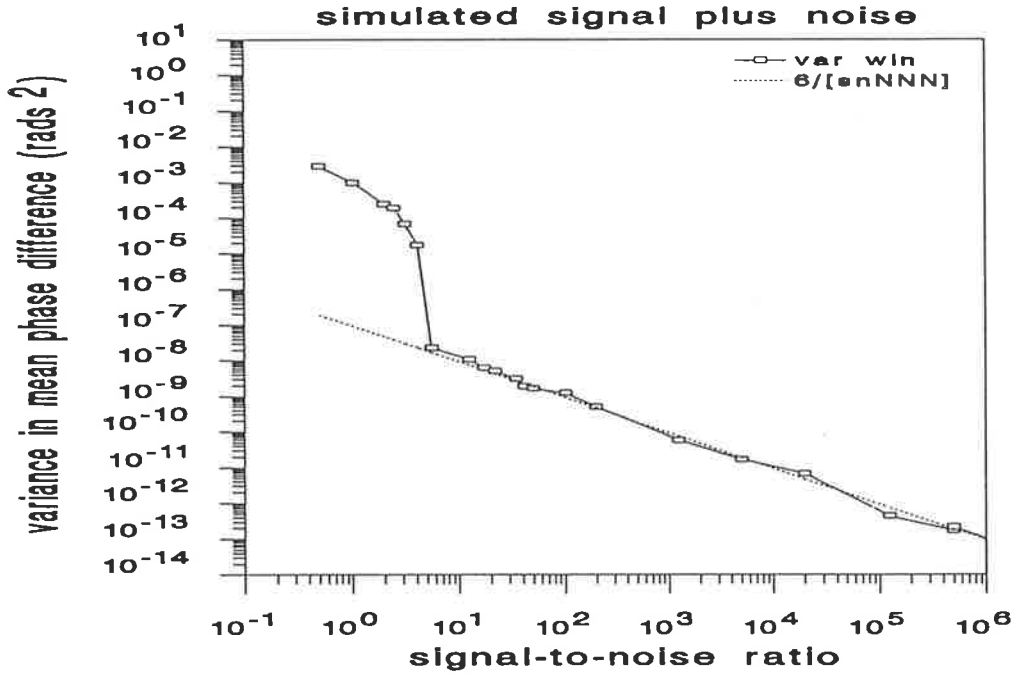

Figure 8 Simulation results, signal plus noise case with Kay weighting function applied to obtain minimum variance estimate for $\widehat{\Delta\theta}$ - showing $Var\{\widehat{\Delta\theta}\}$ compared with CRLB.

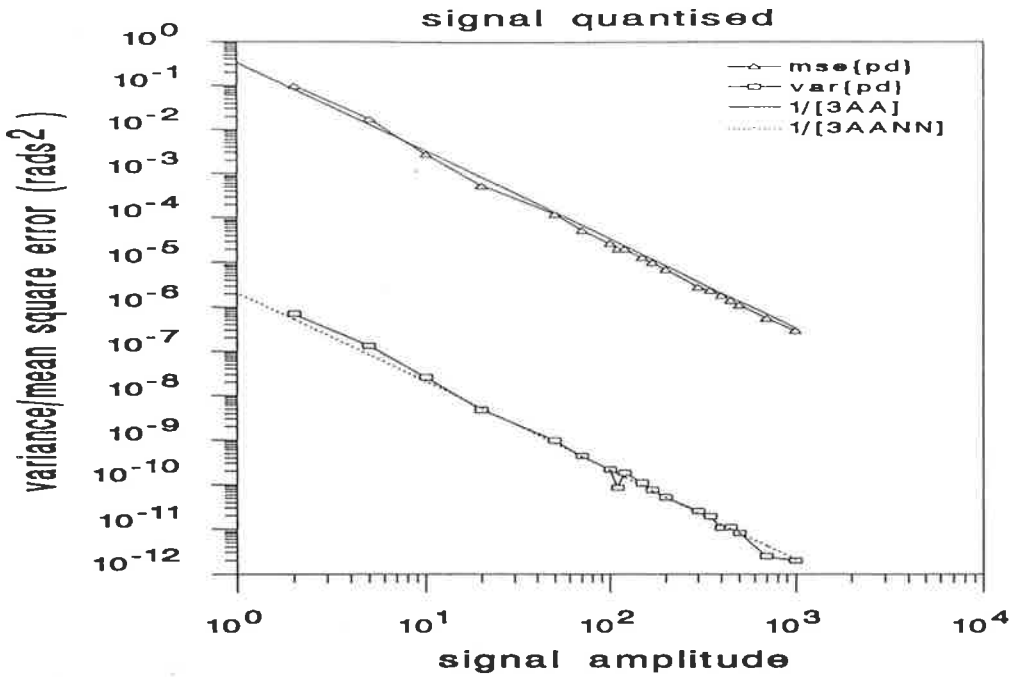

Figure 9 Simulation results, signal quantised case - showing $Var\{\Delta\theta\}$ and $Var\{\widehat{\Delta\theta}\}$ compared with theoretical results.

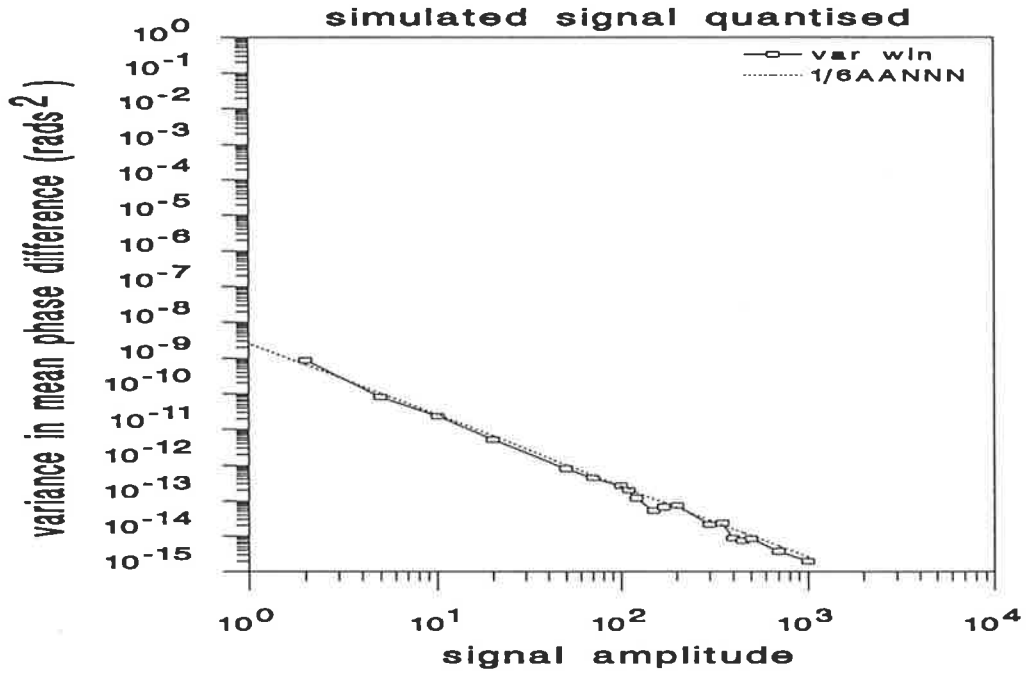

Figure 10 Simulation results, signal quantised case with Kay weighting function applied - showing $Var\{\widehat{\Delta\theta}\}$ compared with $\frac{1}{6A^2N^3}$.

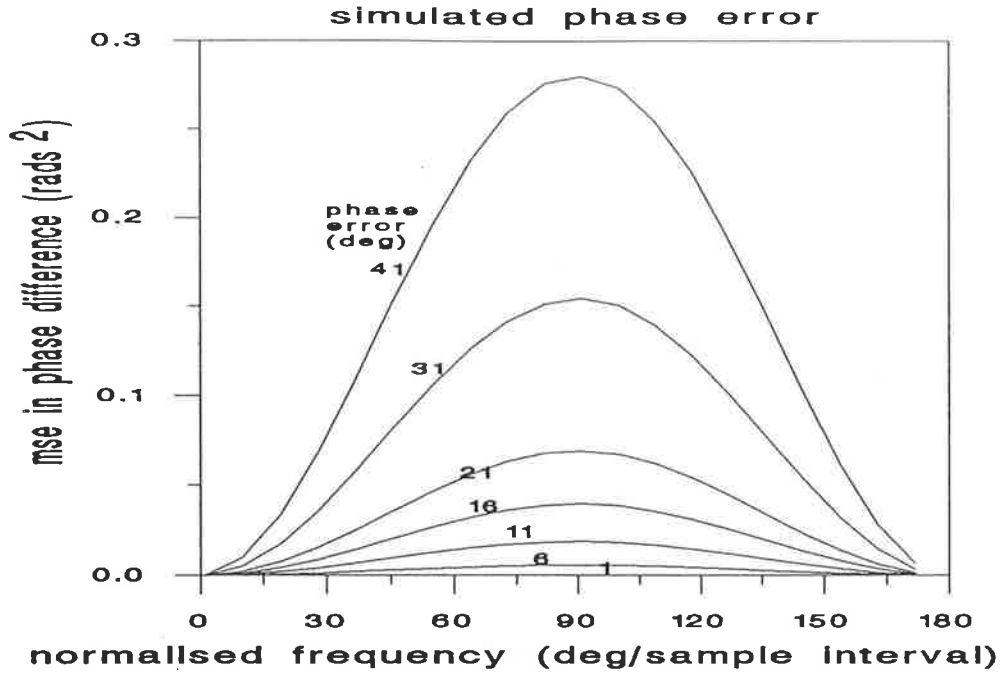

Figure 11 Simulation results, quadrature phase error case - showing $Var\{\Delta\theta\}$ as a function of signal frequency and phase error.

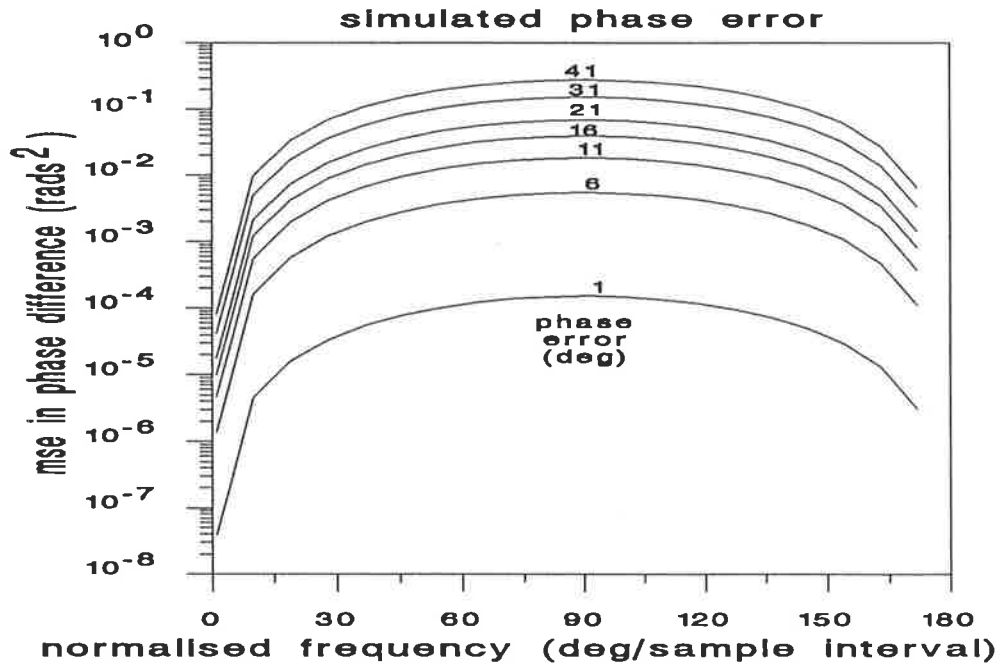

Figure 12 Simulation results, quadrature phase error case - showing $Var\{\Delta\theta\}$ as a function of signal frequency and phase error, plotted on a log scale.

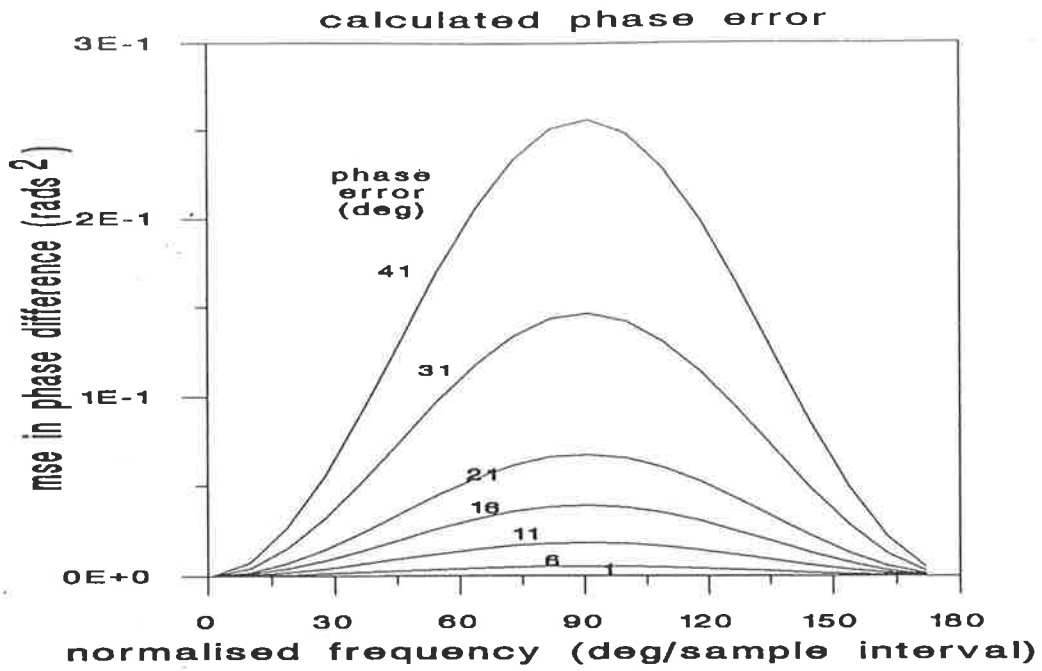

Figure 13 Theoretical results, quadrature phase error case - showing $Var\{\Delta\theta\}$ as a function of signal frequency and phase error.

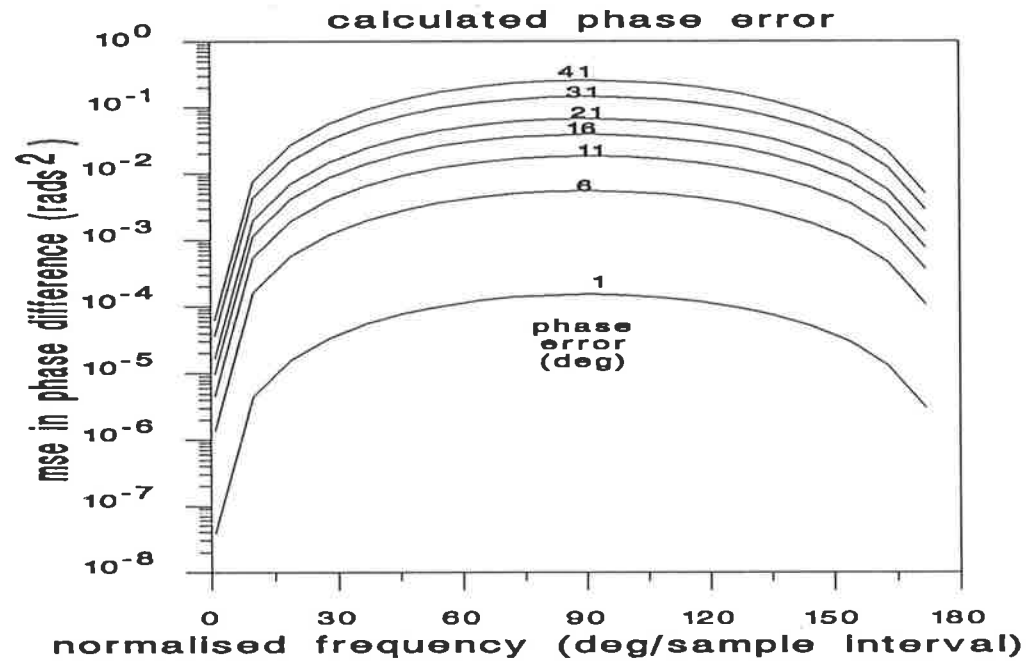

Figure 14 Theoretical results, quadrature phase error case - showing $Var\{\Delta\theta\}$ as a function of signal frequency and phase error, plotted on a log scale.

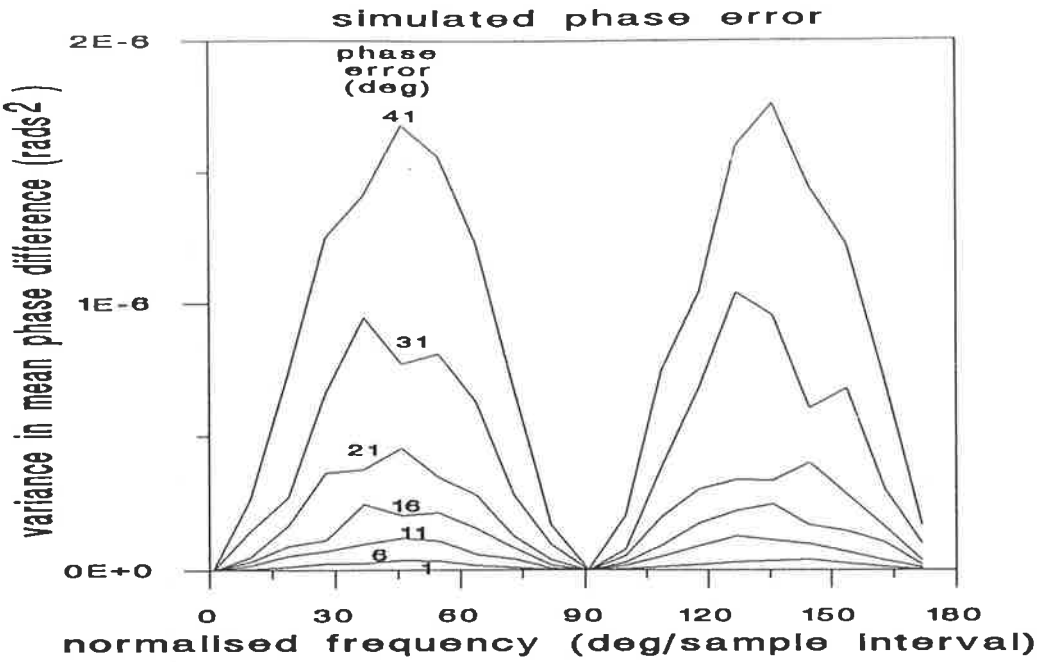

Figure 15 Simulation results, quadrature phase error case - showing $Var\{\widehat{\Delta\theta}\}$ as a function of signal frequency and phase error.

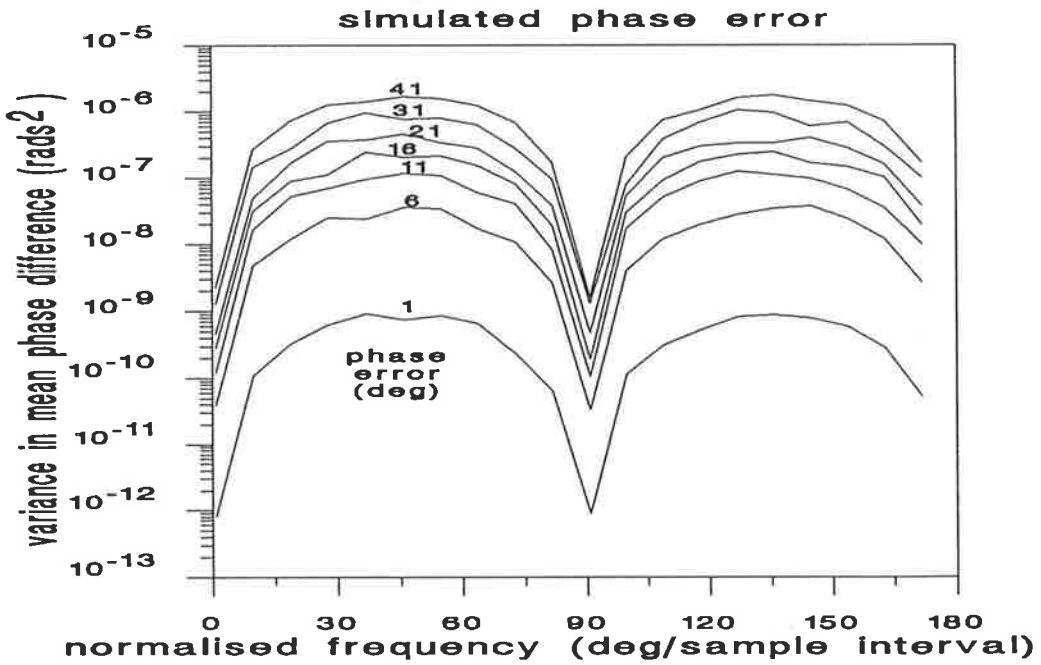

Figure 16 Simulation results, quadrature phase error case - showing $Var\{\widehat{\Delta\theta}\}$ as a function of signal frequency and phase error, plotted on a log scale.

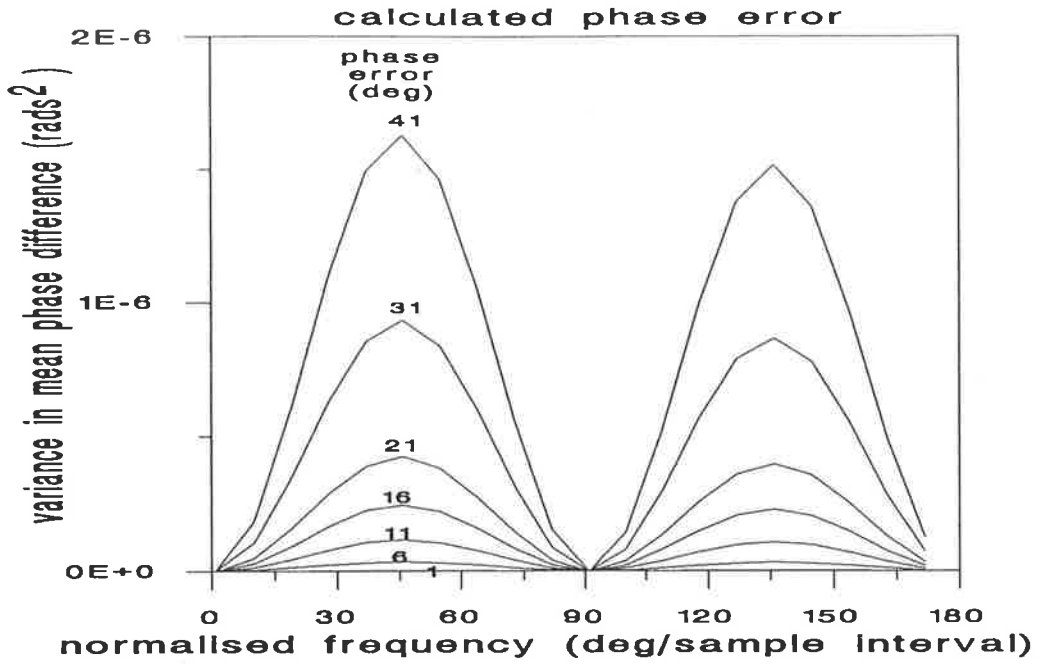

Figure 17 Theoretical results, quadrature phase error case - showing $Var\{\widehat{\Delta\theta}\}$ as a function of signal frequency and phase error.

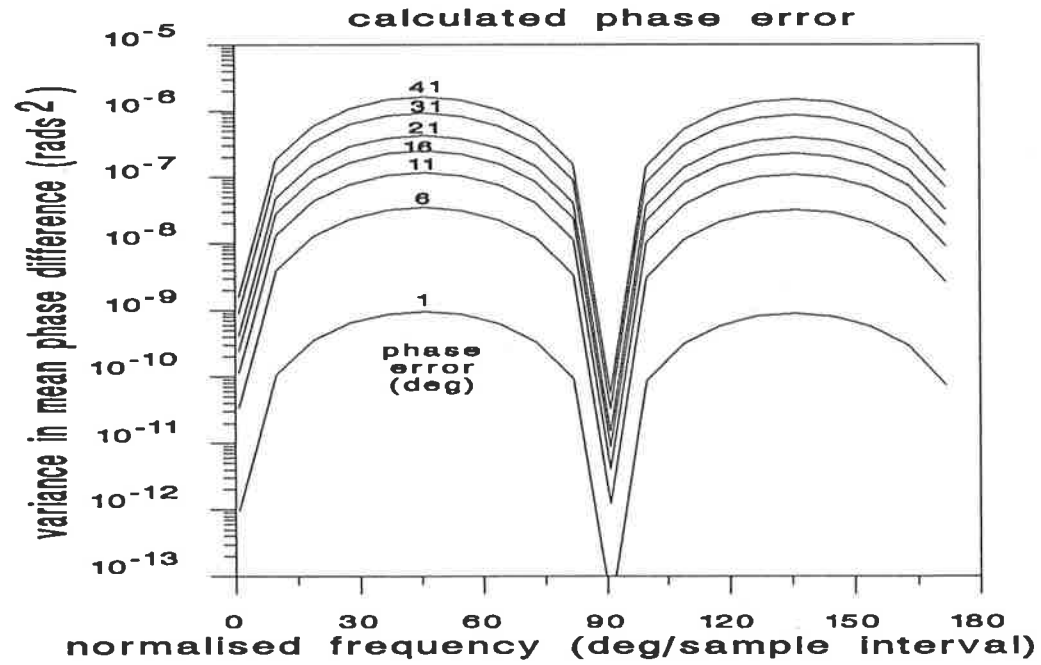

Figure 18 Theoretical results, quadrature phase error case - showing $Var\{\widehat{\Delta\theta}\}$ as a function of signal frequency and phase error, plotted on a log scale.

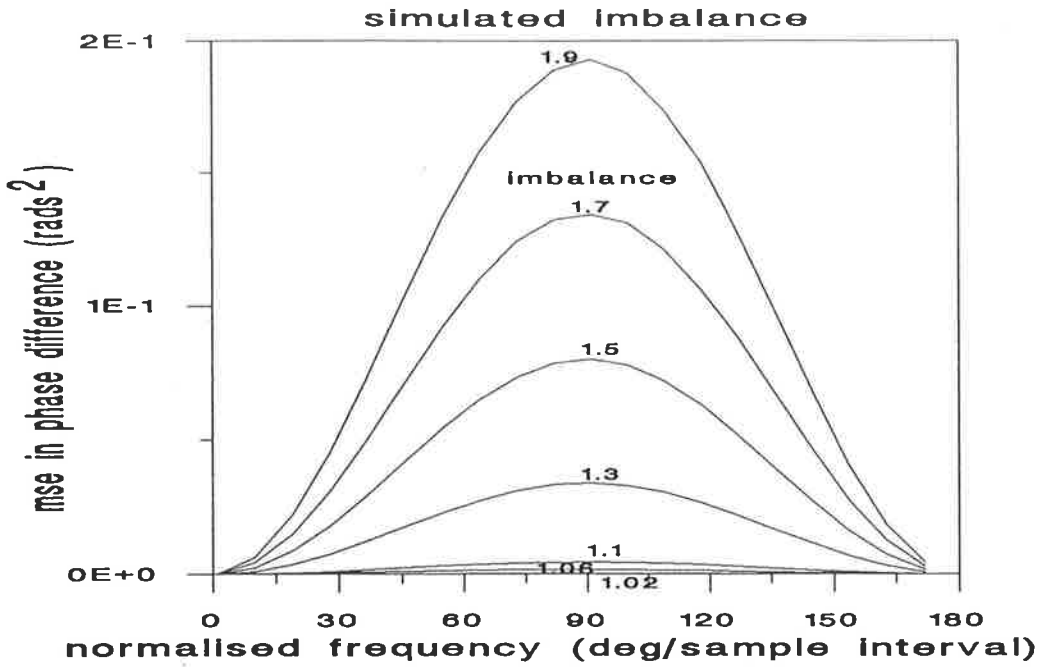

Figure 19 Simulation results, IQ amplitude imbalance case - showing $Var\{\Delta\theta\}$ as a function of signal frequency and imbalance.

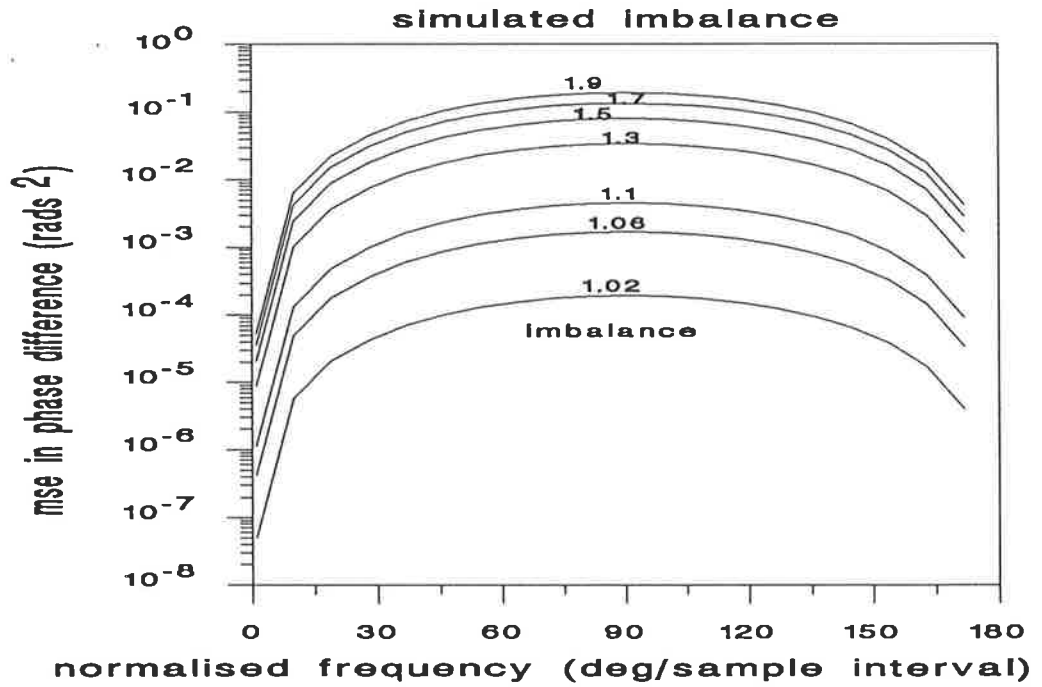

Figure 20 Simulation results, IQ amplitude imbalance case - showing $Var\{\Delta\theta\}$ as a function of signal frequency and imbalance, plotted on a log scale.

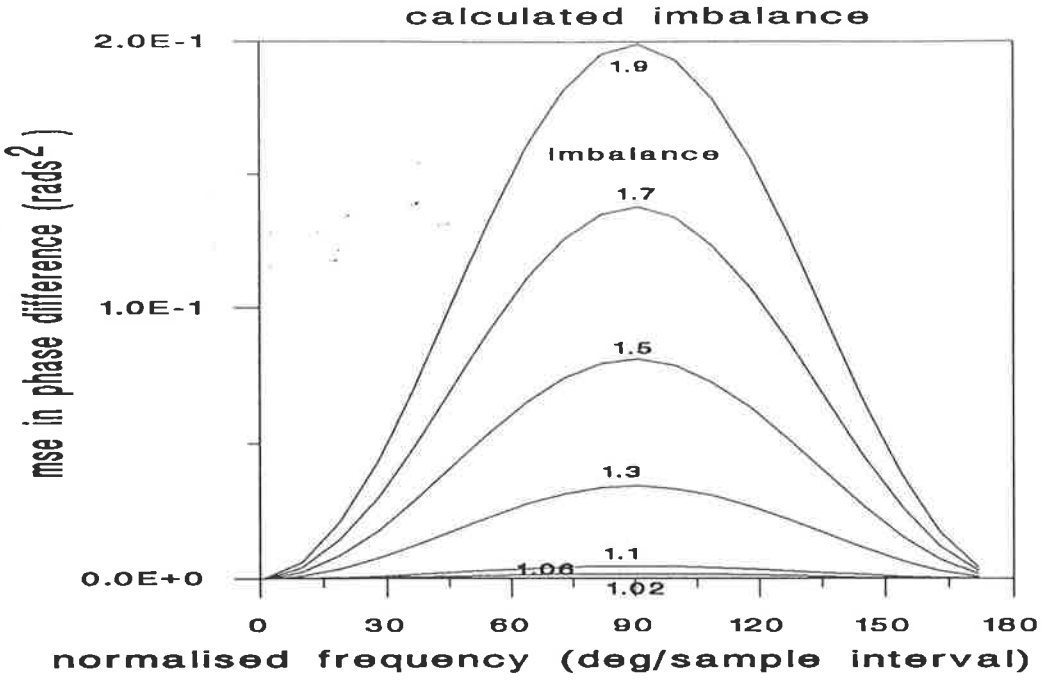

Figure 21 Theoretical results, IQ amplitude imbalance case - showing $Var\{\Delta\theta\}$ as a function of signal frequency and imbalance.

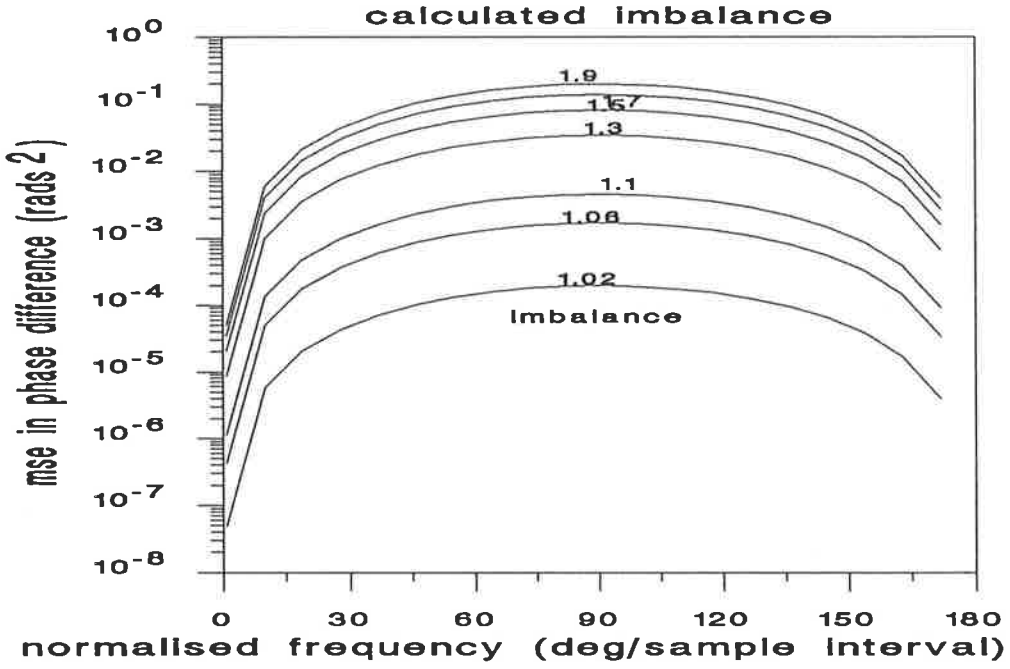

Figure 22 Theoretical results, IQ amplitude imbalance case - showing $Var\{\Delta\theta\}$ as a function of signal frequency and imbalance, plotted on a log scale.

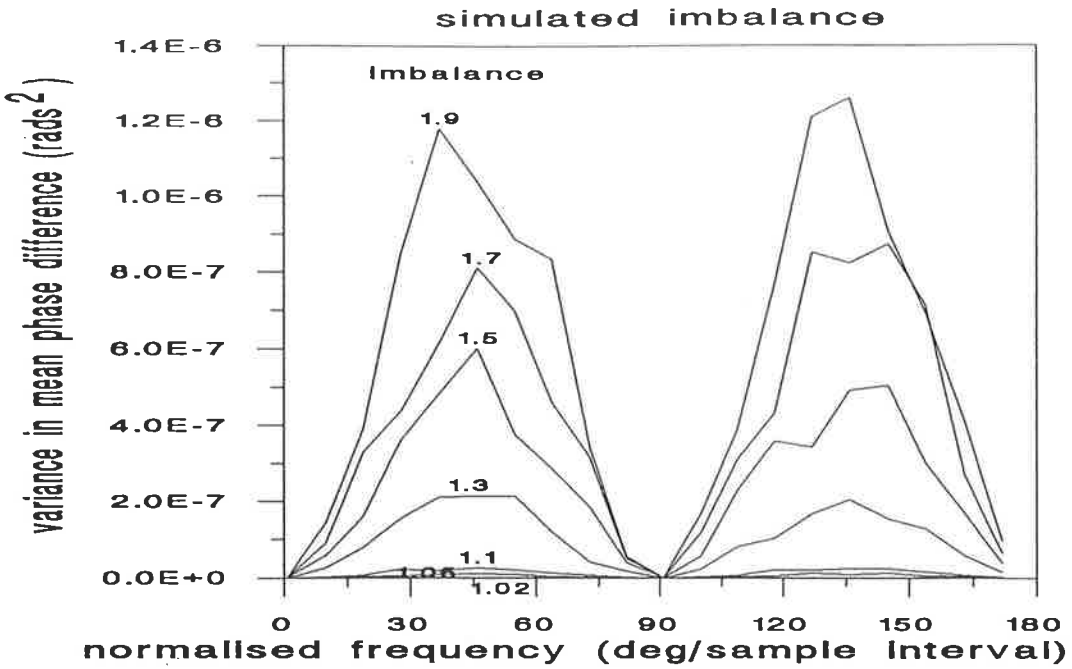

Figure 23 Simulation results, IQ amplitude imbalance case - showing $Var\{\widehat{\Delta\theta}\}$ as a function of signal frequency and imbalance.

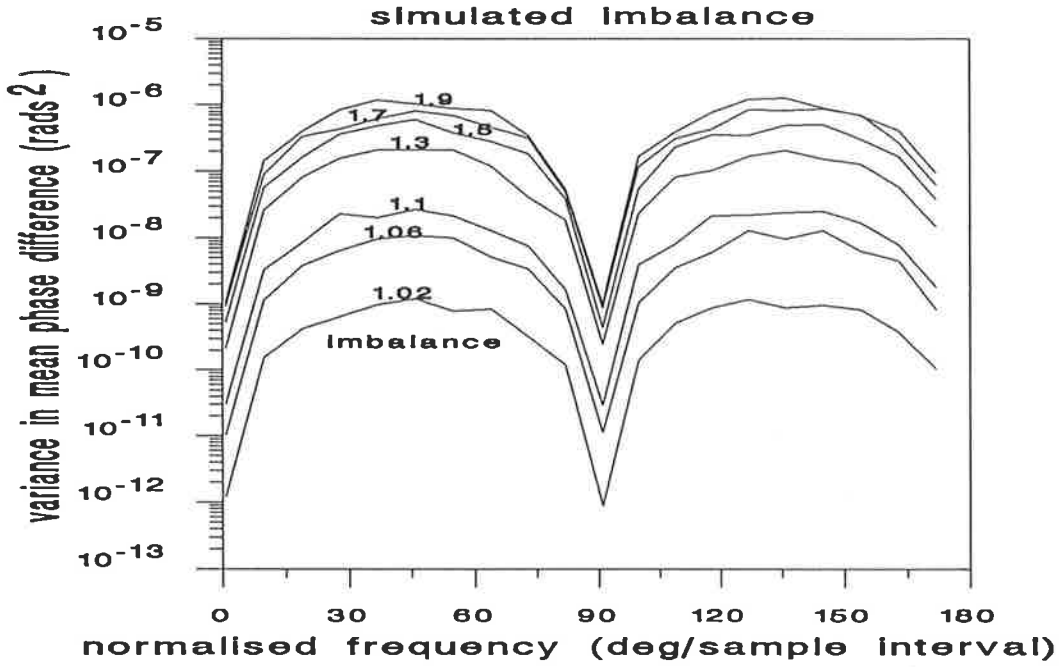

Figure 24 Simulation results, IQ amplitude imbalance case - showing $Var\{\widehat{\Delta\theta}\}$ as a function of signal frequency and imbalance, plotted on a log scale.

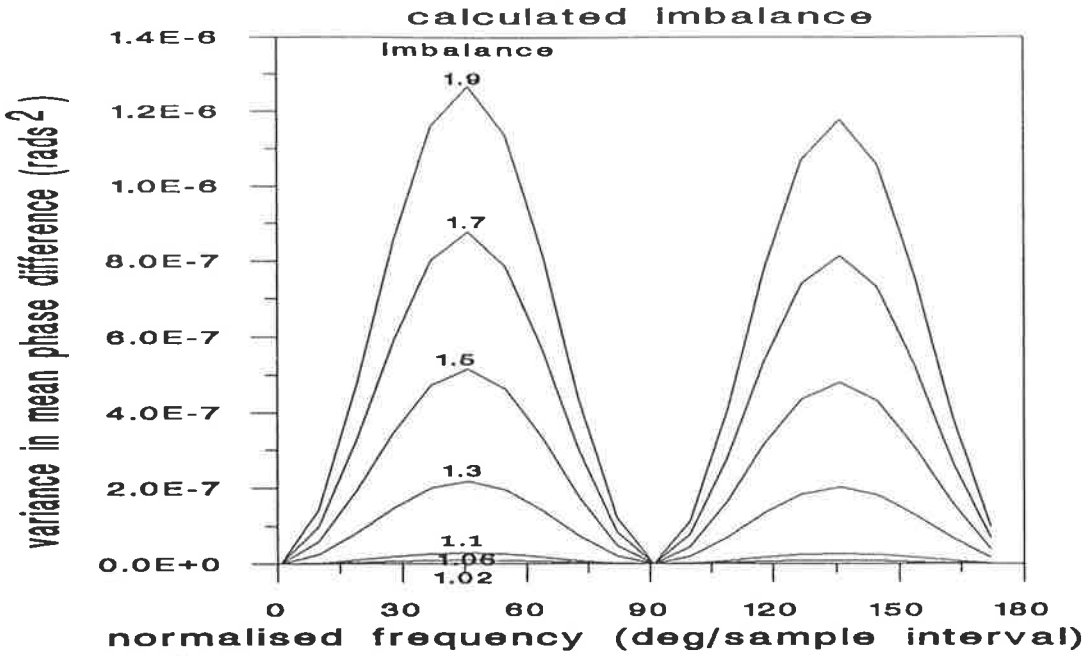

Figure 25 Theoretical results, IQ amplitude imbalance case - showing $Var\{\widehat{\Delta\theta}\}$ as a function of signal frequency and imbalance.

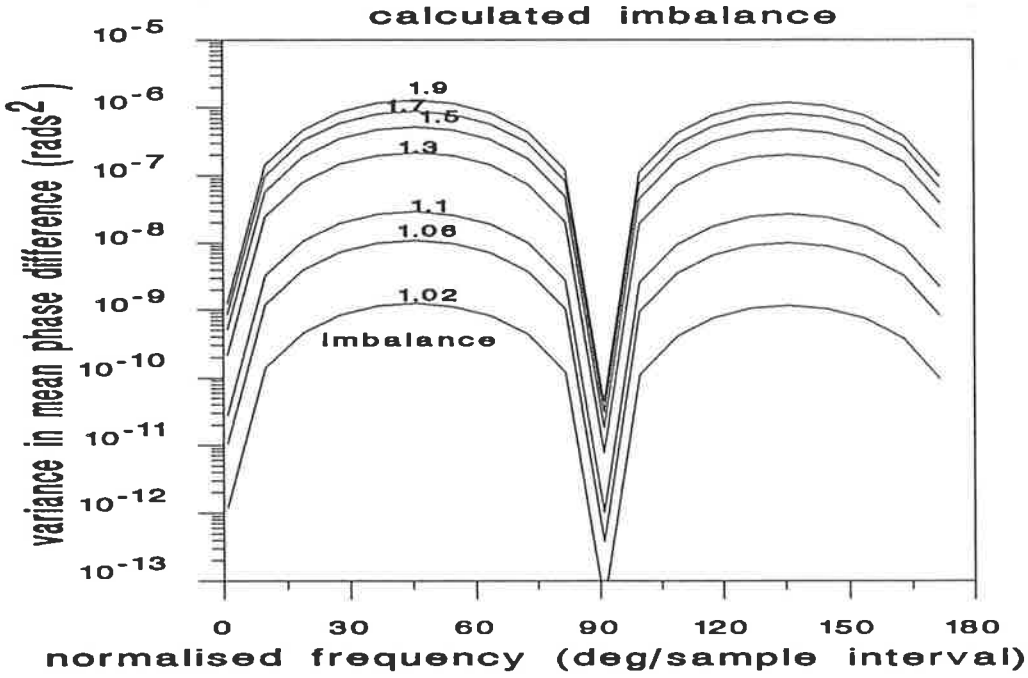

Figure 26 Theoretical results, IQ amplitude imbalance case - showing $Var\{\widehat{\Delta\theta}\}$ as a function of signal frequency and imbalance, plotted on a log scale.

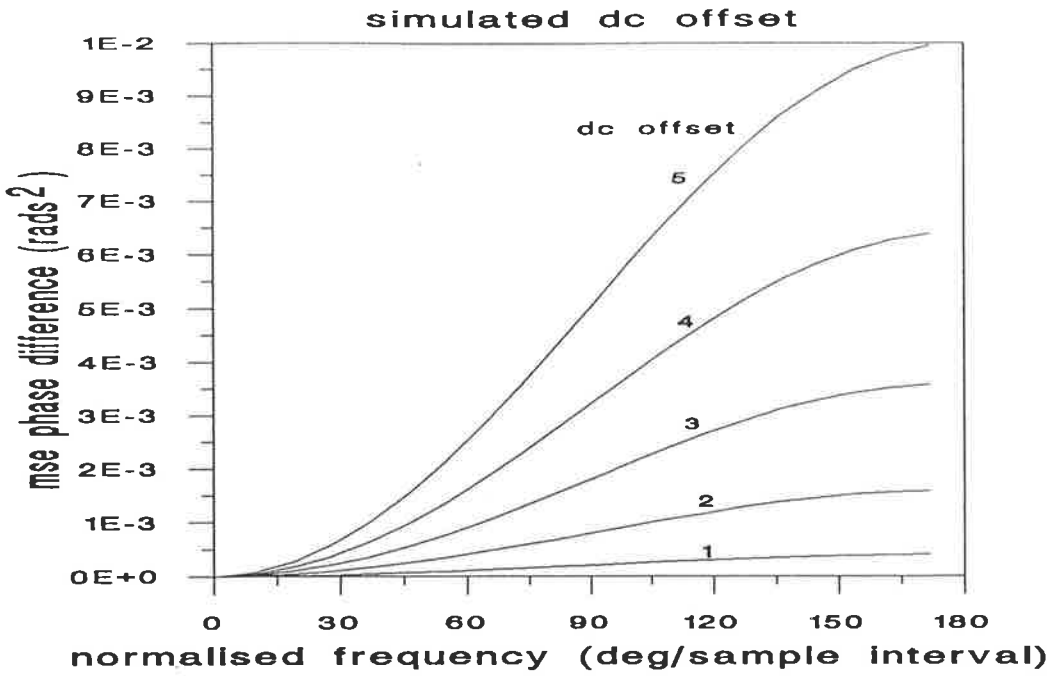

Figure 27 Simulation results, dc offset case - showing $Var\{\Delta\theta\}$ as a function of signal frequency and dc offset.

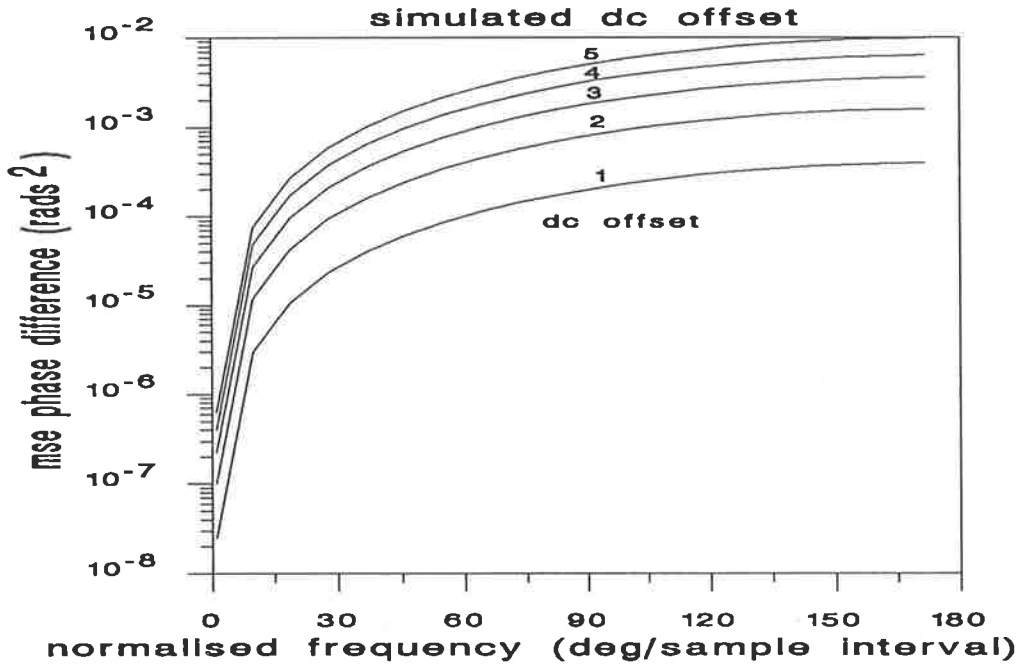

Figure 28 Simulation results, dc offset case - showing $Var\{\Delta\theta\}$ as a function of signal frequency and dc offset, plotted on a log scale.

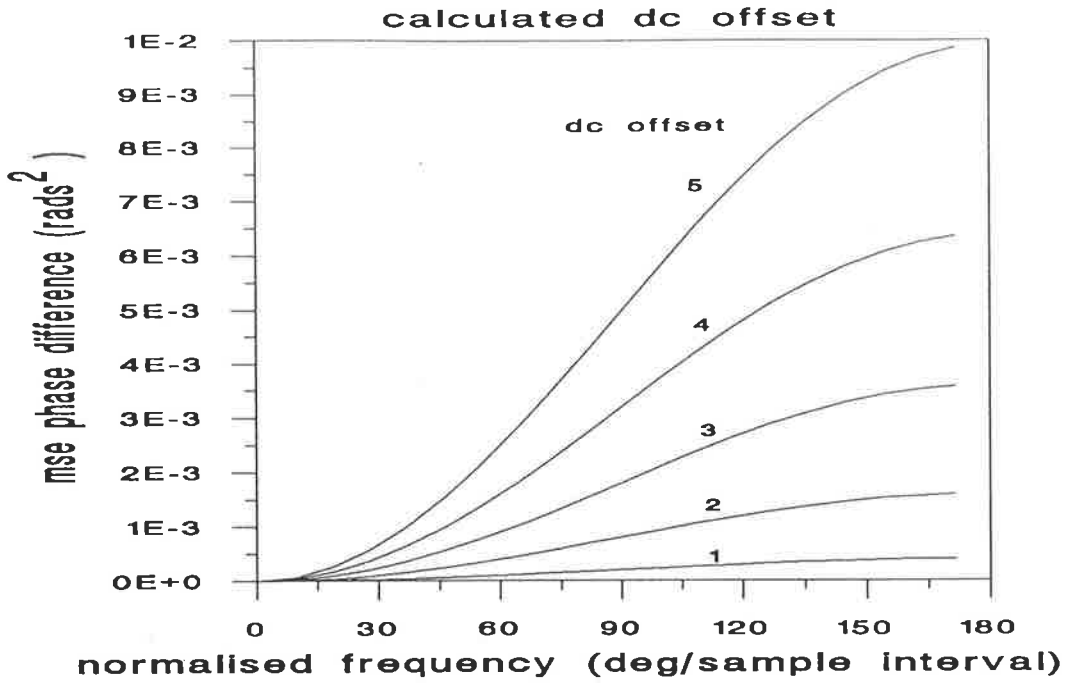

Figure 29 Theoretical results, dc offset case - showing $Var\{\Delta\theta\}$ as a function of signal frequency and dc offset.

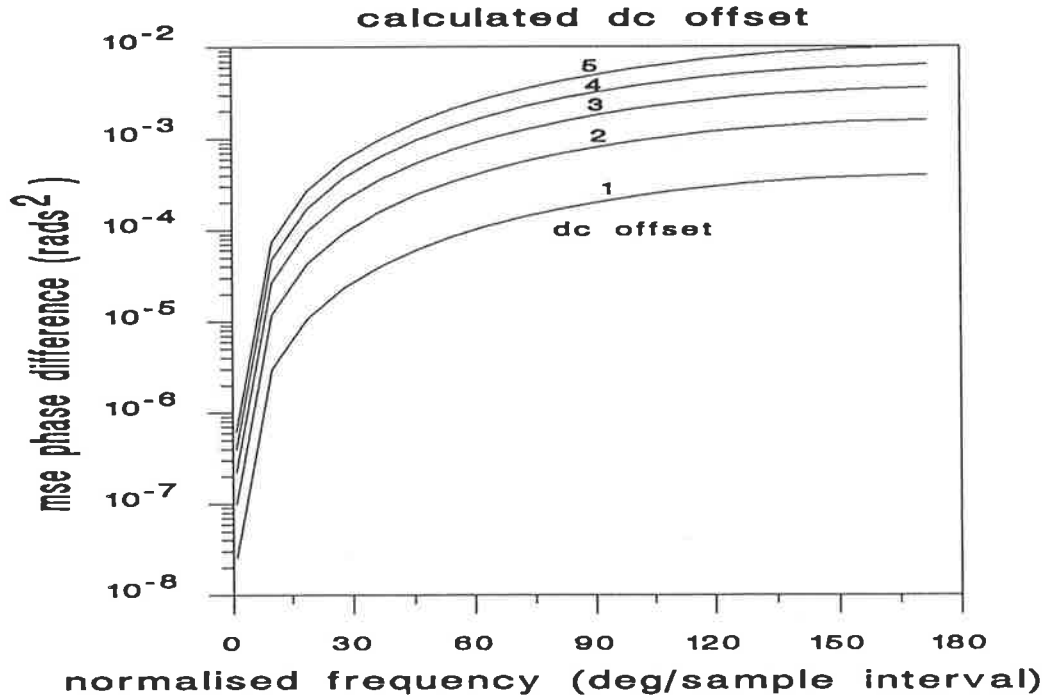

Figure 30 Theoretical results, dc offset case - showing $Var\{\Delta\theta\}$ as a function of signal frequency and dc offset, plotted on a log scale.

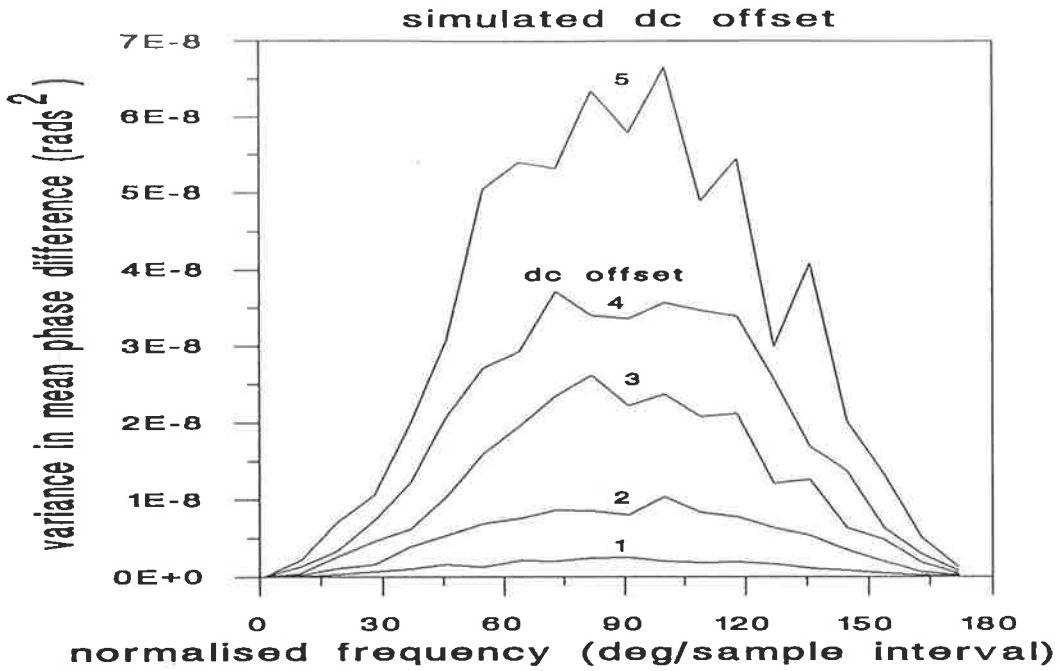

Figure 31 Simulation results, dc offset case - showing $Var\{\widehat{\Delta\theta}\}$ as a function of signal frequency and dc offset.

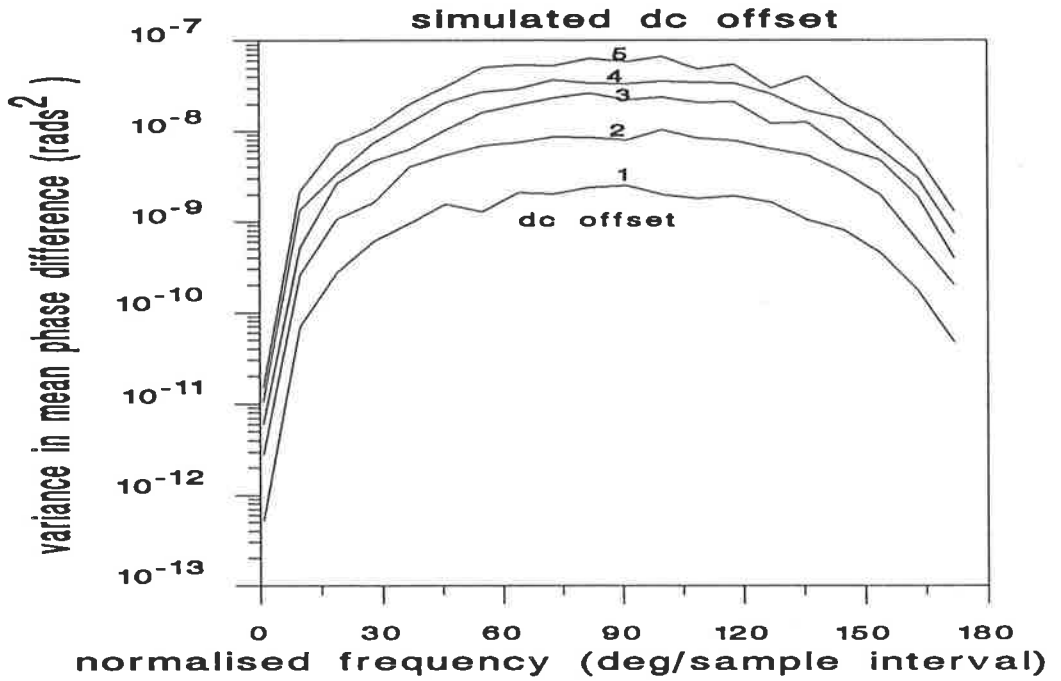

Figure 32 Simulation results, dc offset case - showing $Var\{\widehat{\Delta\theta}\}$ as a function of signal frequency and dc offset, plotted on a log scale.

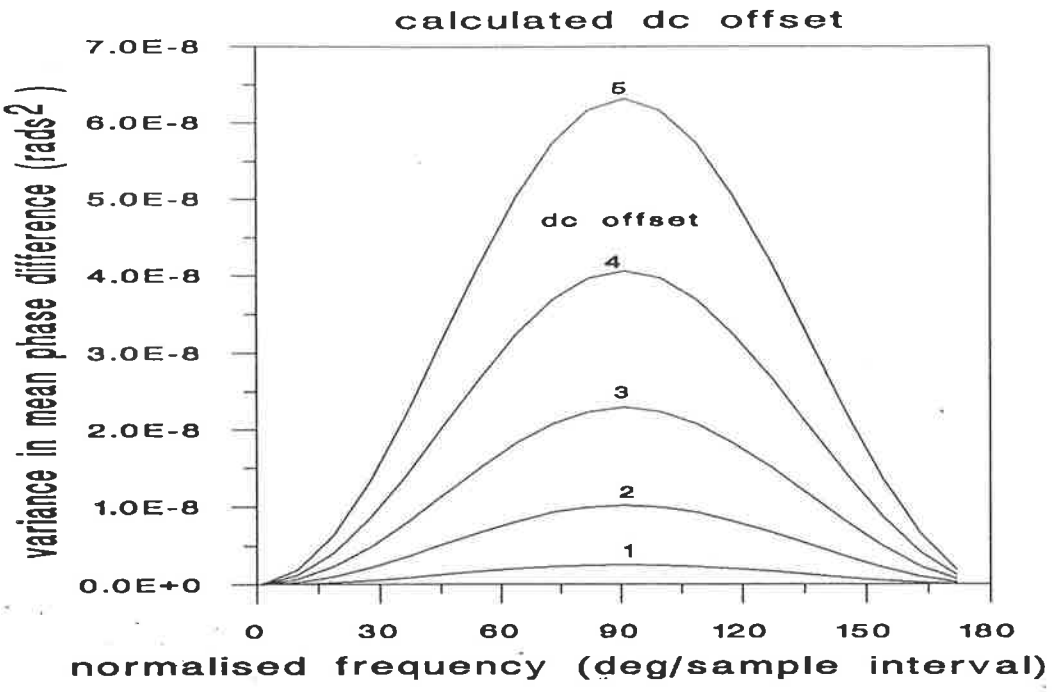

Figure 33 Theoretical results, dc offset case - showing $Var\{\widehat{\Delta\theta}\}$ as a function of signal frequency and dc offset.

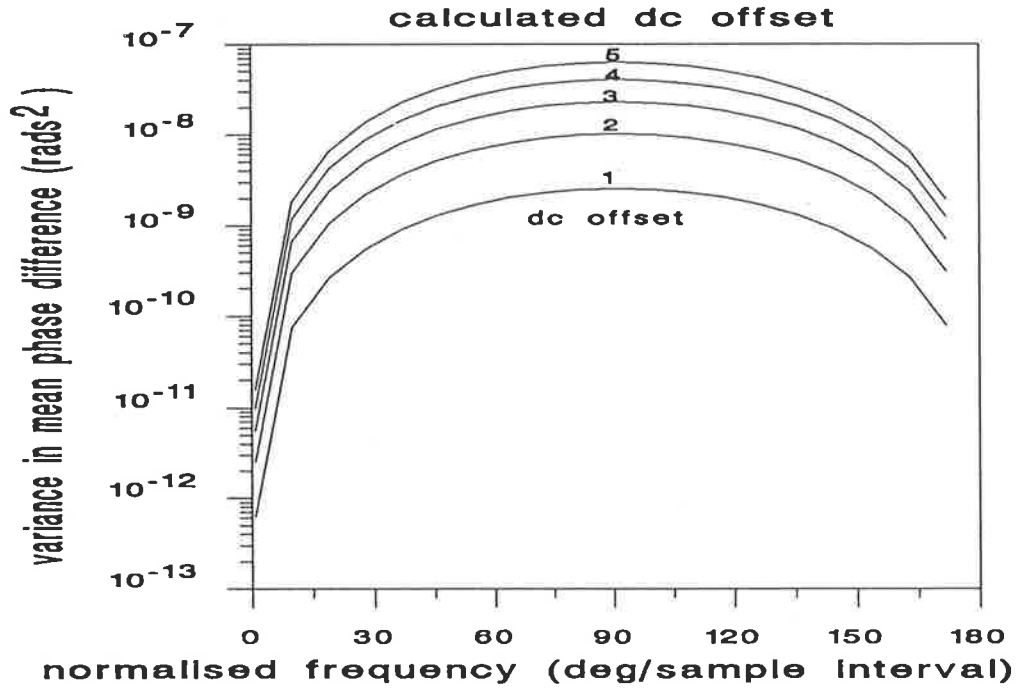

Figure 34 Theoretical results, dc offset case - showing $Var\{\widehat{\Delta\theta}\}$ as a function of signal frequency and dc offset, plotted on a log scale.

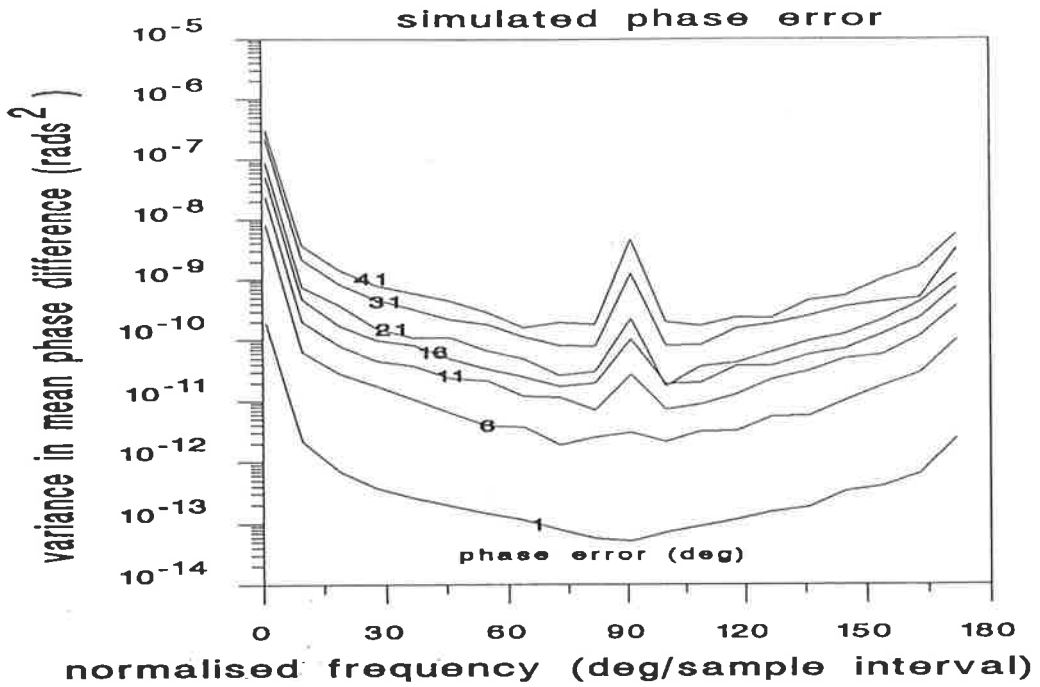

Figure 35 Simulation results, phase error case with Kay weighting function applied - showing $Var\{\widehat{\Delta\theta}\}$ as a function of signal frequency and phase error.

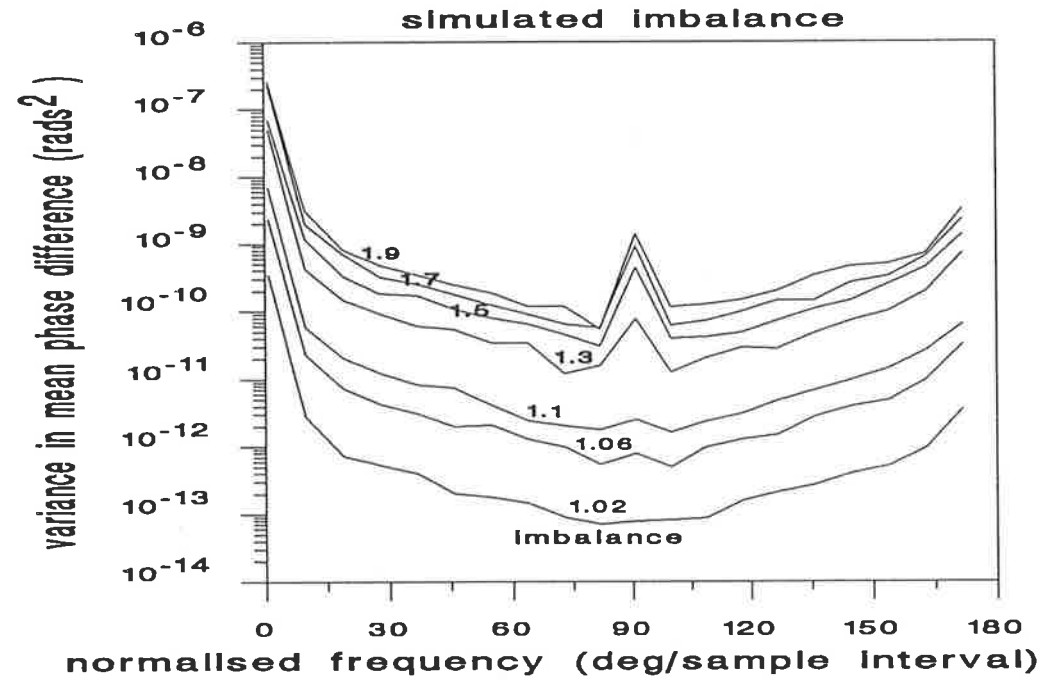

Figure 36 Simulation results, IQ amplitude imbalance case with Kay weighting function applied - showing $Var\{\widehat{\Delta\theta}\}$ as a function of signal frequency and imbalance.

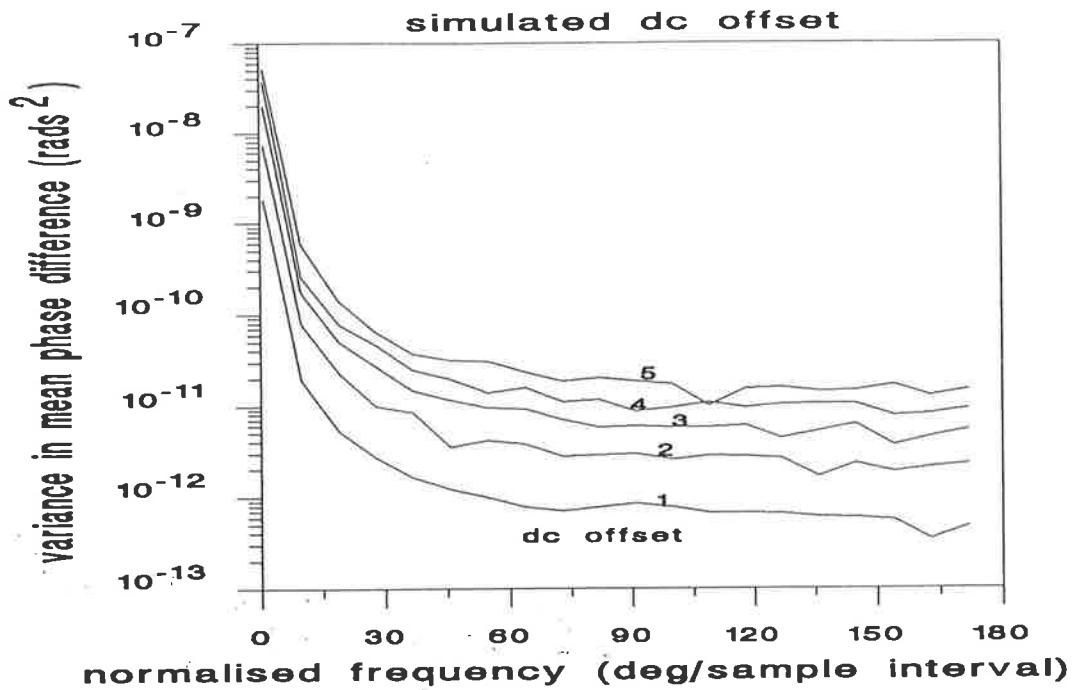

Figure 37 Simulation results, dc offset case with Kay weighting function applied - showing $Var\{\widehat{\Delta\theta}\}$ as a function of signal frequency and dc offset.

4 MEASURING SYSTEM ERRORS

The errors in the real system were measured using the simple correlation methods suggested by Harris [17] (see also appendix X) and were found to be of the order shown in table 1 below.

Table 1 Real system results - measured quadrature hybrid errors.

dc offset I channel	1 unit
dc offset Q channel	1 unit
amplitude imbalance $\frac{Q}{I}$	1.1
quadrature phase error	-2 degrees

where 1 unit = 1 least significant bit (LSB) in the A/D conversion.

Measurements were made with a fixed frequency signal (0.157 rads/sample interval or, equivalently, 10MHz) and various levels of signal-to-noise ratio and the results are shown in figure 38. In the real system a digital sampling rate of 400MHz was used hence the figure 0.157 rads per intersample period corresponds to a signal frequency of 10MHz. The Kay weighting function has been applied and figure 39 shows that the lower variance bound can be achieved for moderate to high SNR. These figures show that the theory fits well with experiment for the signal plus noise case. Note that the variance of all other system errors are include in this experimental case but these produce values much less than those due to the noise alone.

Recordings were also made of various levels of signal amplitude, with minimal noise present in order to study the quantisation effects. Figure 40 shows the results obtained, where $Var\{\Delta\theta\}$ and $Var\{\widehat{\Delta\theta}\}$ without the weighting function are quite a reasonable fit. However application of the weighting function (see figure 41) has not provided the same scale of improvement as for the simulations. In this case the variances due to the phase error, amplitude imbalance and dc offsets are sufficient (even with the weighting function applied) to prevent the combined measured variance from dropping below the bounds obtained (ie approximately 1e-9 to 1e-10) and hence the theoretical lower limit for the pure quantisation case cannot be obtained from this experiment.

In the experimental setup it proved impracticable to carry out measurements of phase error, amplitude imbalance and dc offsets in the real system with the precision required, however the results obtained from the analytical study and from simulations are believed to be satisfactory for the purposes of this study.

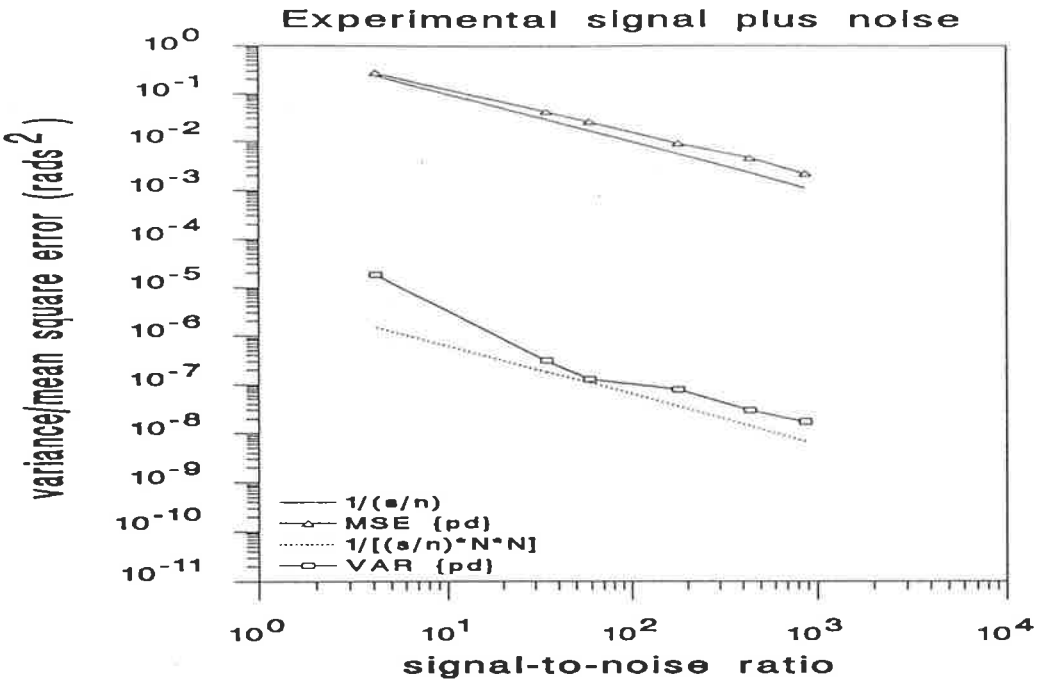

Figure 38 Experimental results, signal plus noise case - showing $Var\{\Delta\theta\}$ and $Var\{\widehat{\Delta\theta}\}$ compared with theoretical results.

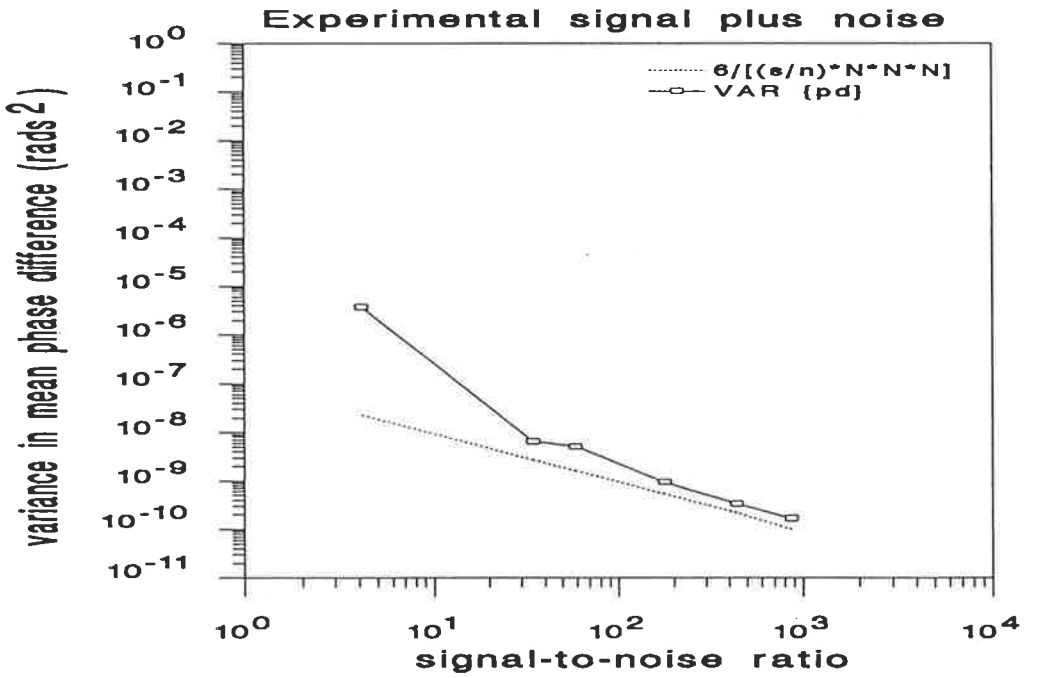

Figure 39 Experimental results, signal plus noise with Kay weighting function applied to obtain minimum variance estimate for $\widehat{\Delta\theta}$ - showing $Var\{\widehat{\Delta\theta}\}$ compared with CRLB.

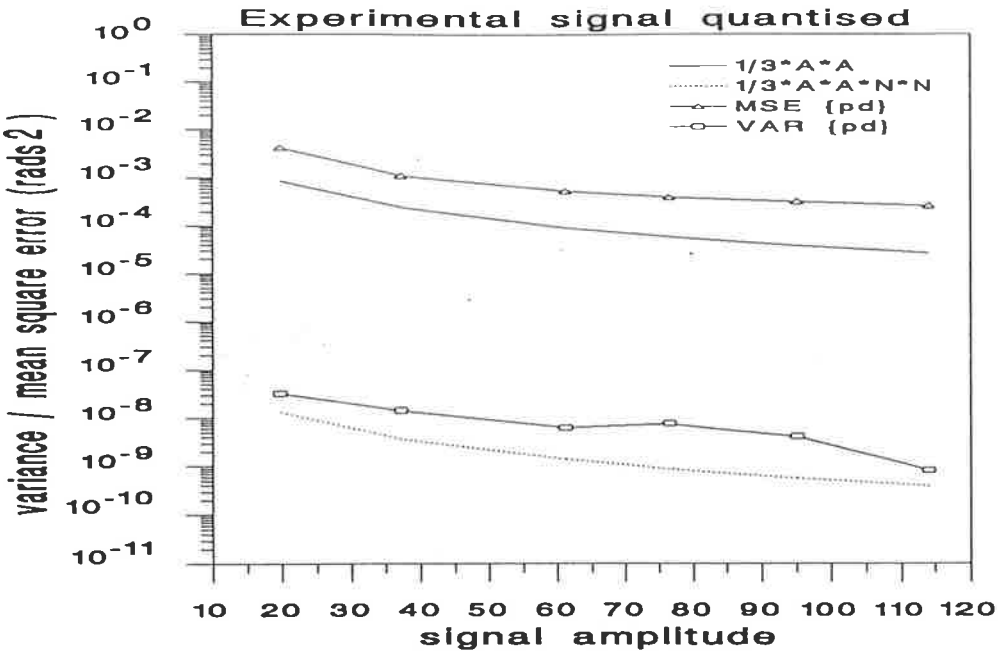

Figure 40 Experimental results, signal quantised case - showing $Var\{\Delta\theta\}$ and $Var\{\widehat{\Delta\theta}\}$ compared with theoretical results.

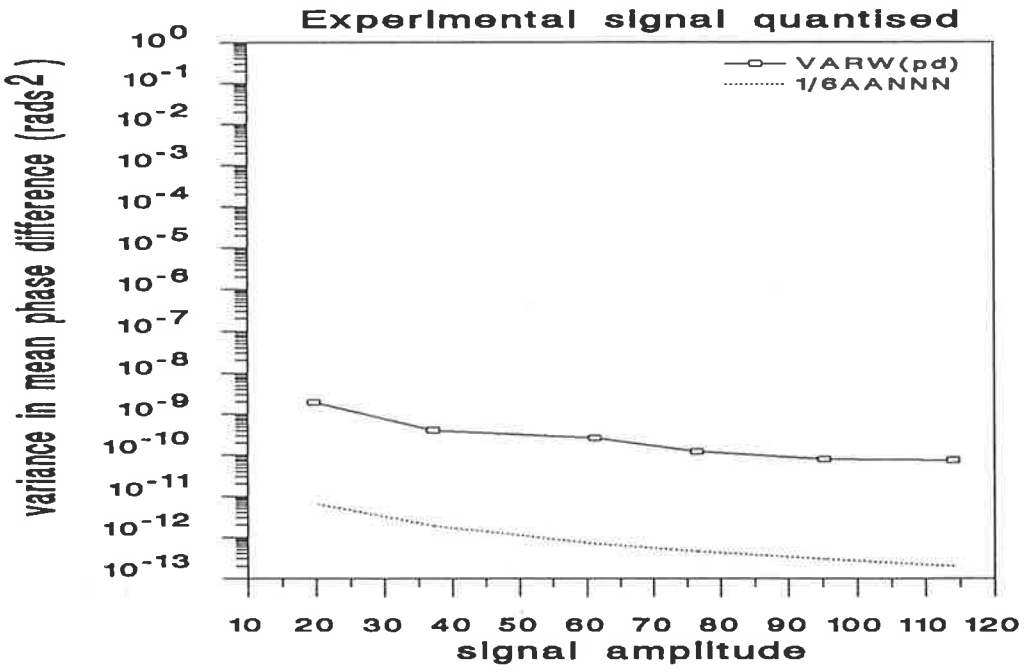

Figure 41 Experimental results, signal quantised case with Kay weighting function applied - showing $Var\{\widehat{\Delta\theta}\}$ compared with $\frac{1}{6A^2N^3}$.

5 TECHNIQUES TO REDUCE SYSTEM ERRORS

Using the auto correlations and the cross correlations of the two signal sample vectors (see appendix X and XI) it has been shown that a good estimate can be obtained for the quadrature hybrid errors : phase error, amplitude imbalance, dc offsets. Clearly we require a number of sample points to be able to carry out the correlations and the more points that are available the better the estimates will be. Hence we must consider some form of block processing for real time applications. To reduce the noise level a form of linear predictive filter has been selected which also has the ability to adapt to a signal which has a time varying frequency. This filter has the added benefit that it reduces the quantisation noise.

5.1 Removing dc offset and channel imbalance

Removing the channel imbalance and dc offsets is straightforward but readjusting the phase error is more complicated - the methods are detailed in appendix XI. Figures 42,43 show the effects on $Var\{\Delta\theta\}$ and $Var\{\widehat{\Delta\theta}\}$ respectively of simulations with a phase error followed by correction for phase error. Both $Var\{\Delta\theta\}$ and $Var\{\widehat{\Delta\theta}\}$ have been reduced. Figures 44,45 show the effects on the $Var\{\Delta\theta\}$ and $Var\{\widehat{\Delta\theta}\}$ respectively of simulations with an amplitude imbalance followed by correction for amplitude imbalance. Both $Var\{\Delta\theta\}$ and $Var\{\widehat{\Delta\theta}\}$ have been reduced. Figures 46,47 show the effects on $Var\{\Delta\theta\}$ and $Var\{\widehat{\Delta\theta}\}$ of simulations with a fixed dc offset in each channel followed by dc offset correction. $Var\{\widehat{\Delta\theta}\}$ has been greatly reduced and $Var\{\Delta\theta\}$ has been reduced though not to the same extent as for the phase and amplitude cases.

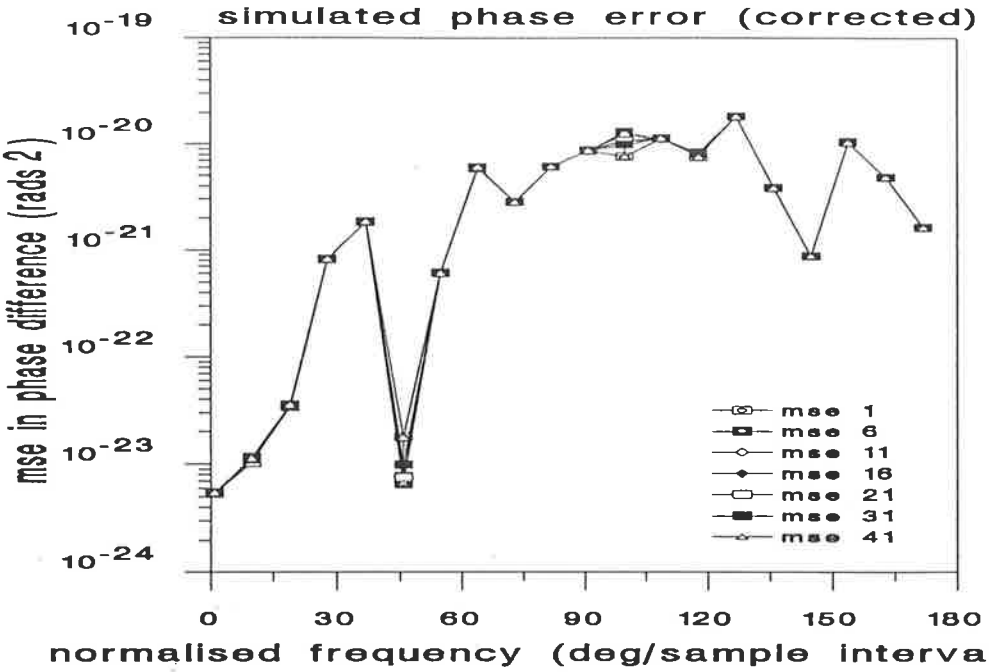

Figure 42 Simulation results, quadrature phase error case - showing $Var\{\Delta\theta\}$, after correction applied, as a function of signal frequency and phase error.

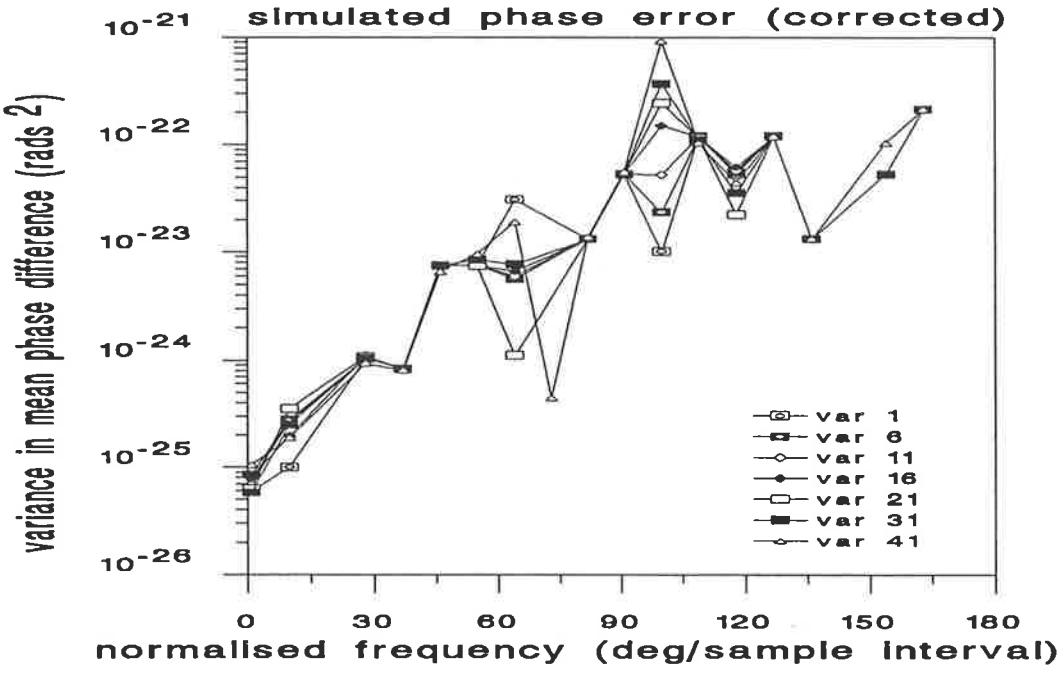

Figure 43 Simulation results, quadrature phase error case - showing $Var\{\widehat{\Delta\theta}\}$, after correction applied, as a function of signal frequency and phase error.

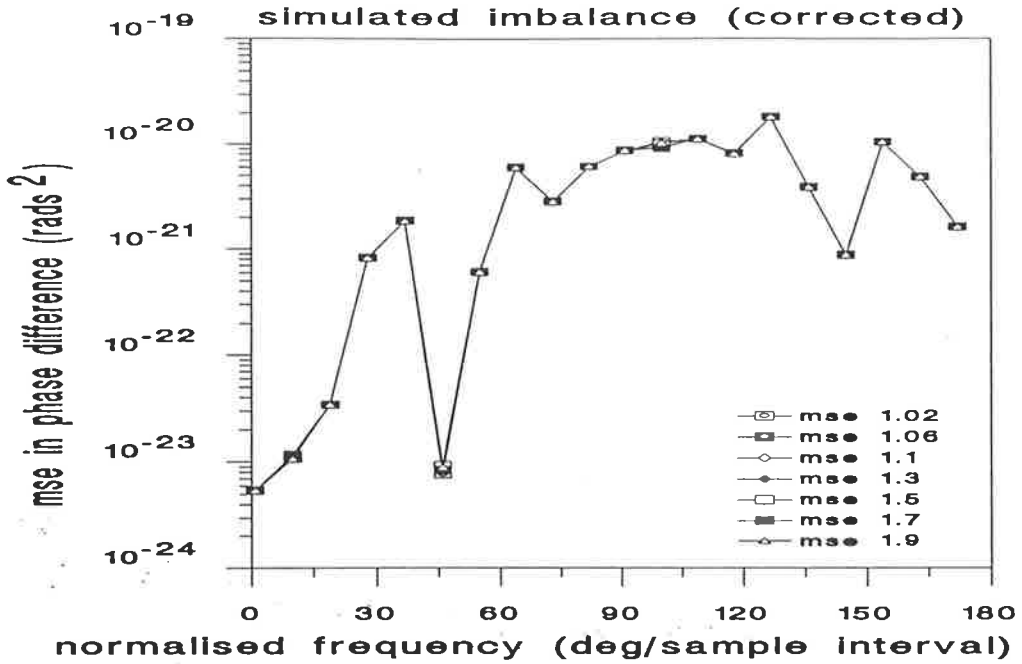

Figure 44 Simulation results, IQ amplitude imbalance case - showing $Var\{\Delta\theta\}$, after correction applied, as a function of signal frequency and imbalance.

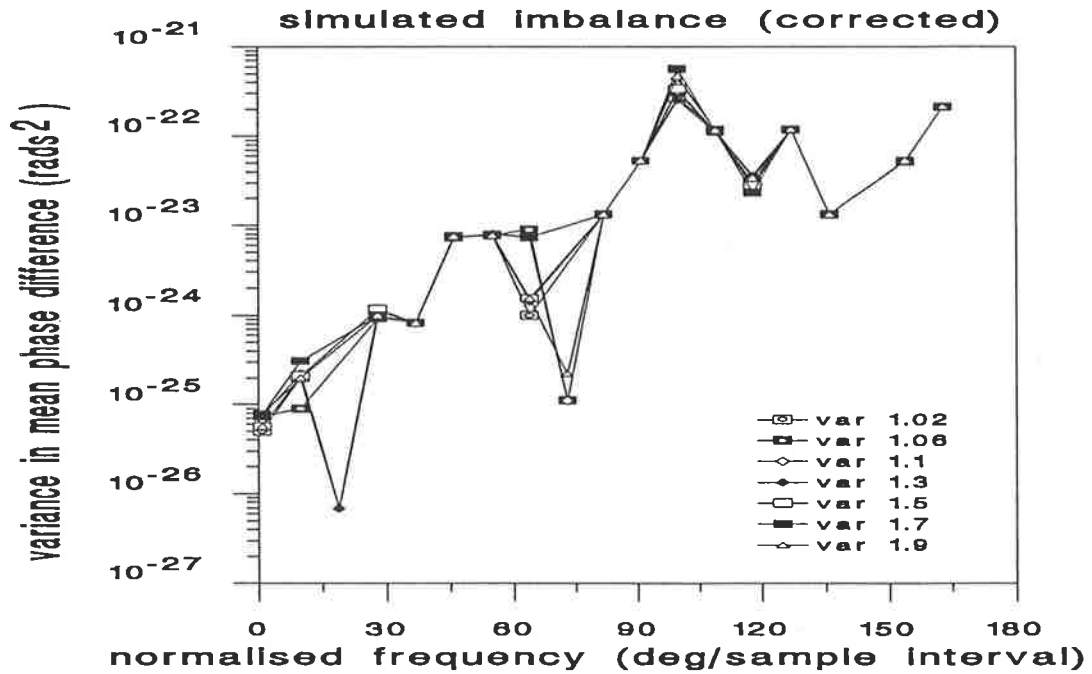

Figure 45 Simulation results, IQ amplitude imbalance case - showing $Var\{\widehat{\Delta\theta}\}$, after correction applied, as a function of signal frequency and imbalance.

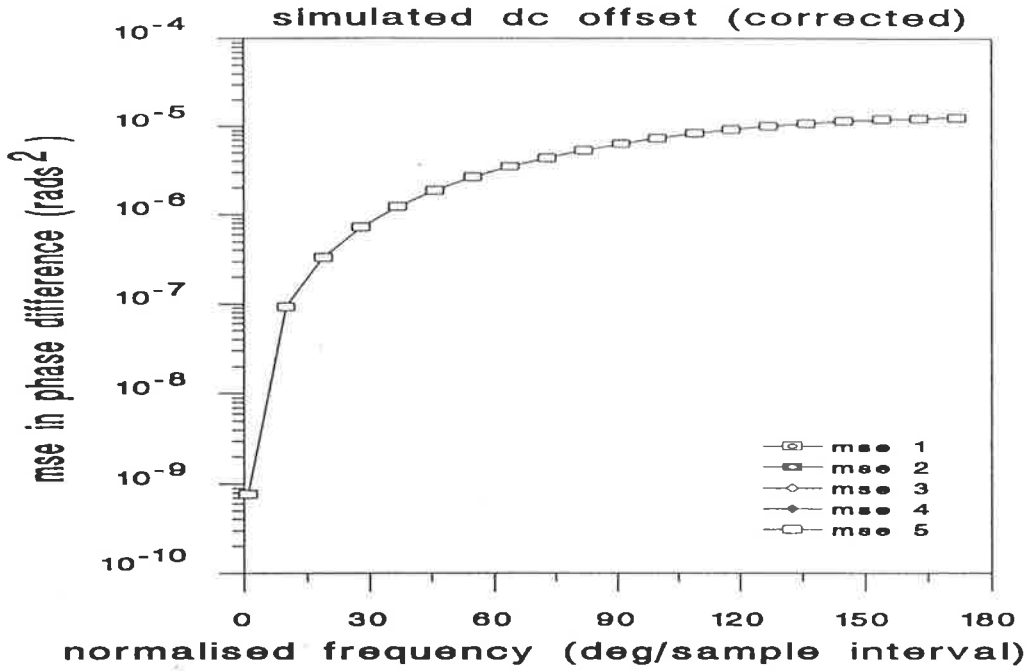

Figure 46 Simulation results, dc offset case - showing $Var\{\Delta\theta\}$, after correction applied, as a function of signal frequency and dc offset.

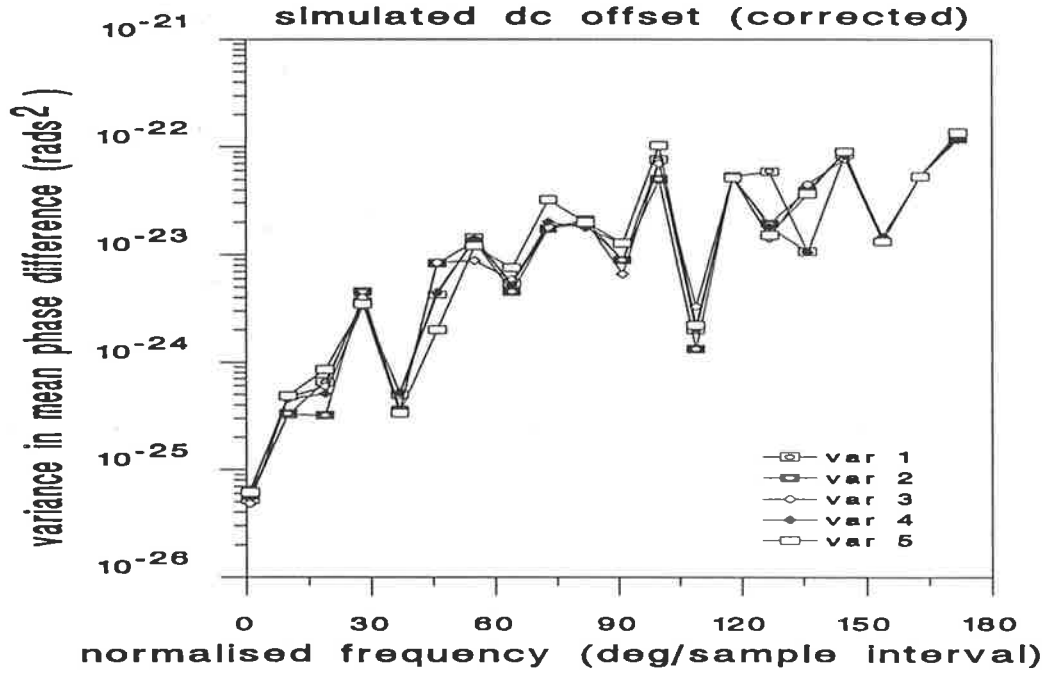

Figure 47 Simulation results, dc offset case - showing $Var\{\widehat{\Delta\theta}\}$, after correction applied, as a function of signal frequency and dc offset.

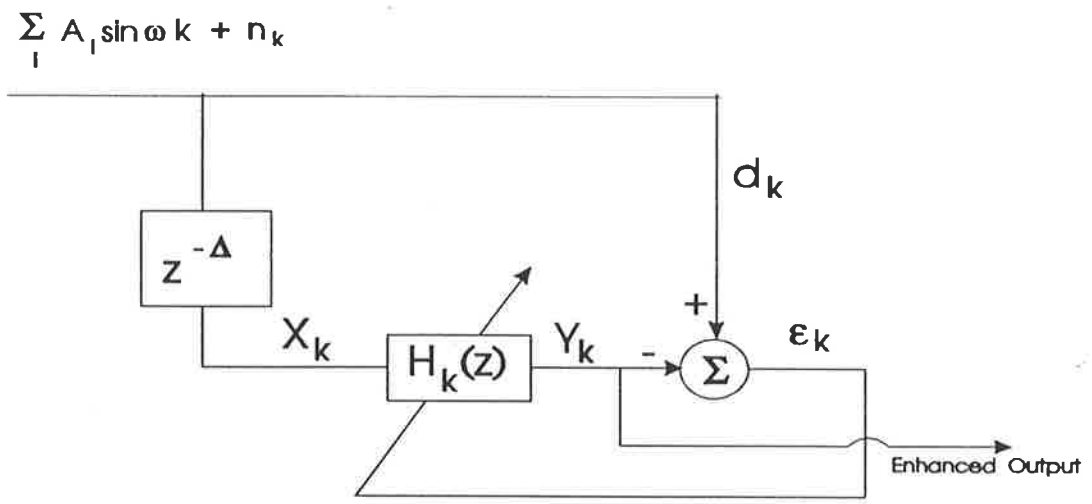

Figure 48 Block diagram for selected method of adaptive line enhancement.

An adaptive line enhancement type filter (a block diagram of which is shown in figure 48) was selected from the book by Stearns and David [18] with filter coefficient updating based upon minimising the mean square signal error by the method of steepest descent. By applying a filter of this type in each of the REAL and IMAG channels of the quadrature hybrid we can enhance the signal output ie improve the snr. The adaptive filter was applied with the filter order (L) set to 16 and the convergence parameter (μ) equal to 0.1 and various levels of snr simulated. In the simulations a delay of 10 times the filter order was imposed before samples were used to calculate phase etc - this was done in order to allow the filter coefficients time to adapt sufficiently. Figure 49 shows the results of these simulations for a fixed frequency signal (0.157 rads/sample interval or, equivalently, 10MHz) plus Gaussian noise. In figure 49 we see that $Var\{\Delta\theta\}$ has been reduced by approximately 10dB except for snr less than 0dB.$Var\{\widehat{\Delta\theta}\}$ has been improved by approximately 5dB except for snr less than 0dB. With the Kay weighting function applied after the adaptive filter we still obtain the Cramer-Rao lower variance bound (see figure 50). The advantage in using the filter is that it acts upon the signal waveform in each channel and reduces the snr. This then reduces any problems encountered in calculating the signal phase and phase difference ie the phase unwrapping problem is reduced and the effective signal input bandwidth can be maintained.

The same filter as above was applied in simulations involving a quantised signal for various signal amplitudes with the results as shown in figure 51. Again $Var\{\Delta\theta\}$ has been improved

by approximately 10dB and $Var\{\widehat{\Delta\theta}\}$ by about the same except for low SNR.

Simulations were performed with all types of error incorporated with error correction for each type of error (see table 2 below).

Table 2 Summary of errors incorporated in system simulations.

signal frequency	0.157 rads/sample interval or 10MHz
90 degree phase error	5 degrees
amplitude imbalance $\left(\frac{Q}{I}\right)$	1.3
dc offset in I channel	5 units = 5LSB
dc offset in Q channel	5 units = 5LSB
SNR	0 to 60dB

Error correction was applied in the following manner :-

(i) dc offsets calculated for each channel waveform then subtracted.

(ii) adaptive filter applied to reduce the mse in each waveform.

(iii) amplitude imbalance calculated and corrected in the Q channel.

(iv) relative phase error calculated and corrected for in the Q channel.

Figure 52 shows the results when the filter adaptation coefficient is 0.01.

Figure 53 shows the results when the filter adaptation coefficient is 0.05.

Figure 54 shows the results when the filter adaptation coefficient is 0.1.

A low value of μ (0.01) means that the filter coefficients adapt more slowly but should yield better results. A high value of μ (0.1) allows the filter coefficients to adapt more quickly but yield less of an improvement in the results.

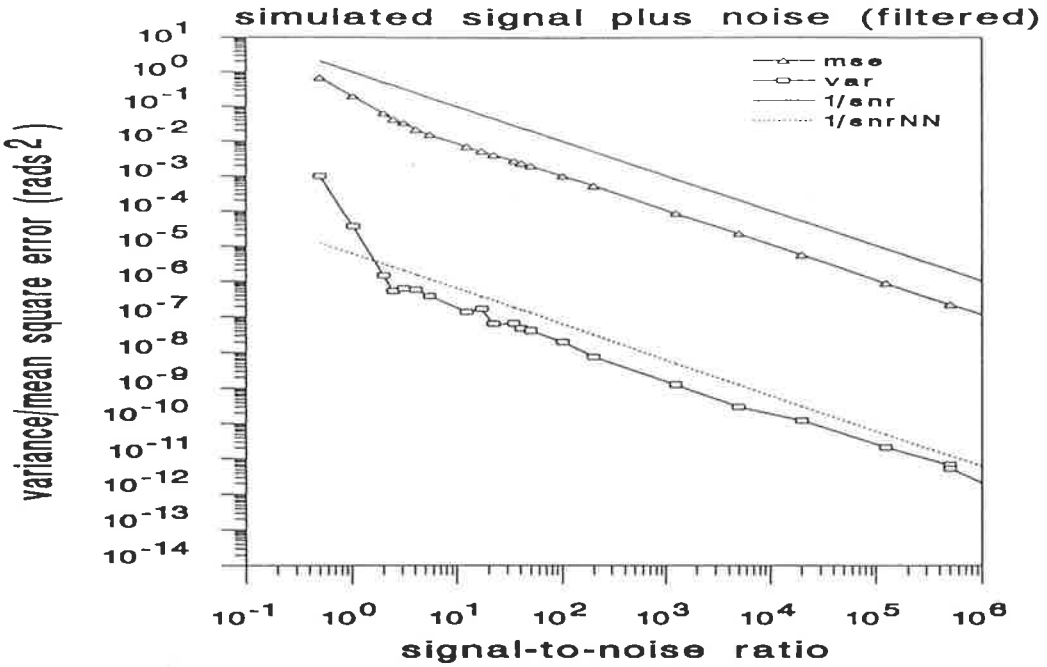

Figure 49 Simulation results, signal plus noise case - showing $Var\{\Delta\theta\}$ and $Var\{\widehat{\Delta\theta}\}$, after filtering applied, compared with theoretical results.

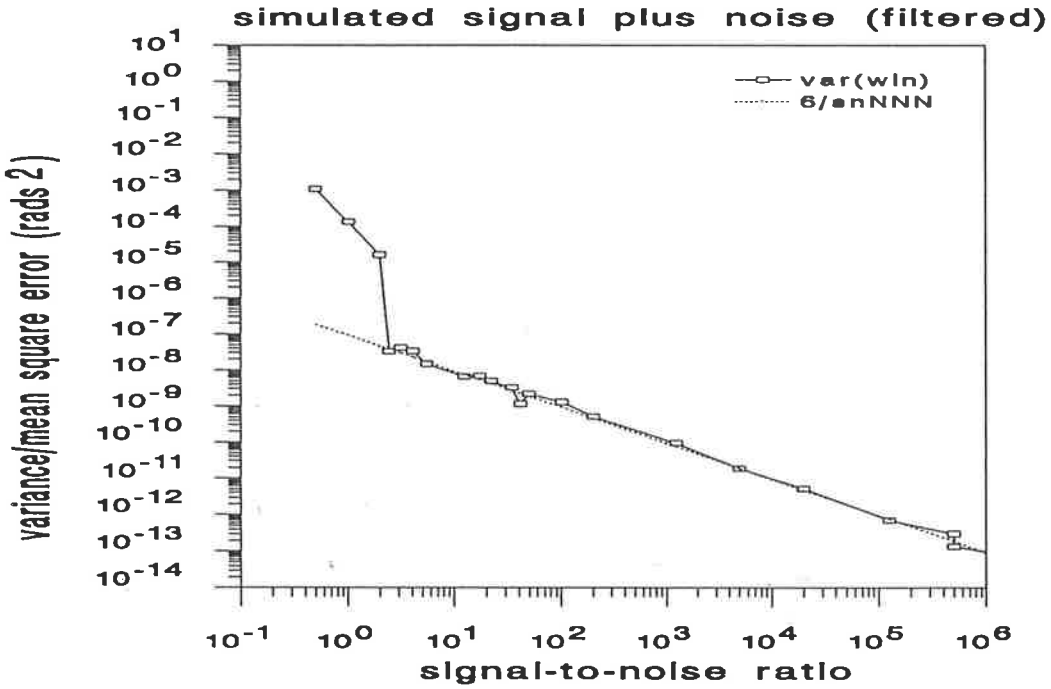

Figure 50 Simulation results, signal plus noise case with Kay weighting function applied - showing $Var\{\widehat{\Delta\theta}\}$, after filtering applied, compared with CRLB.

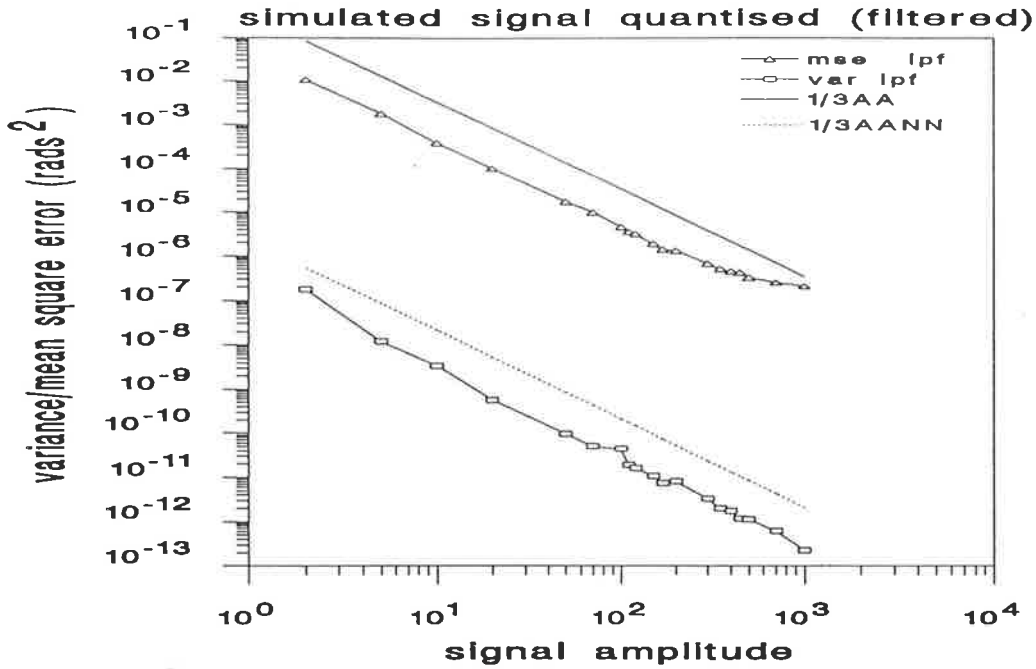

Figure 51 Simulation results, signal quantised case - showing $Var\{\Delta\theta\}$ and $Var\{\widehat{\Delta\theta}\}$, after filtering applied, compared with theoretical results.

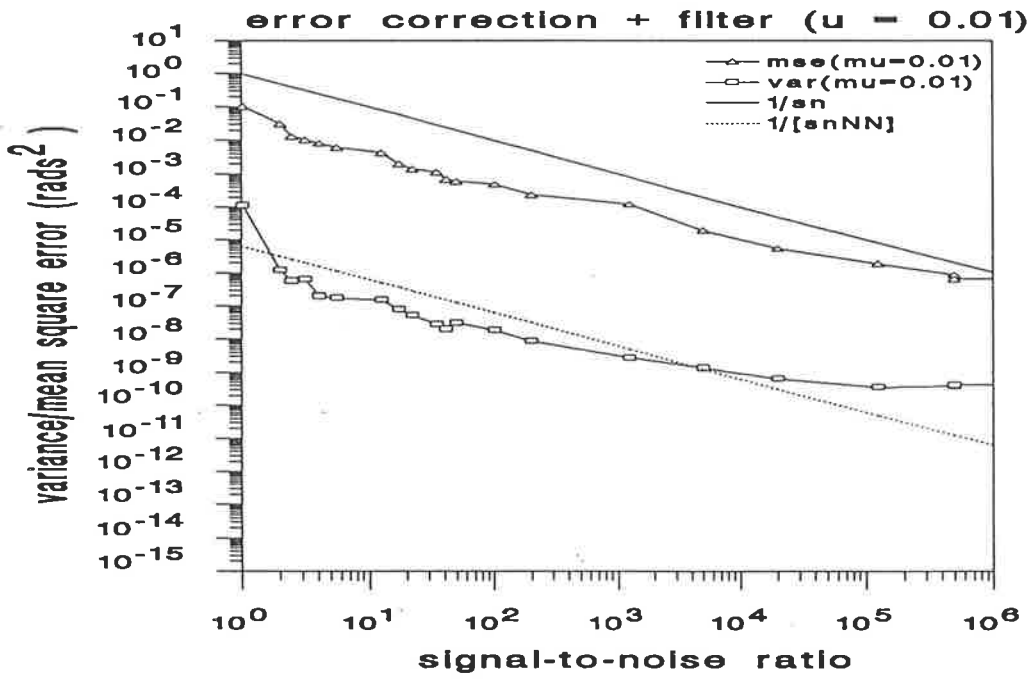

Figure 52 Simulation results, all error types incorporated case - showing $Var\{\Delta\theta\}$ and $Var\{\widehat{\Delta\theta}\}$, after filtering applied ($\mu = 0.01$), compared with theoretical results.

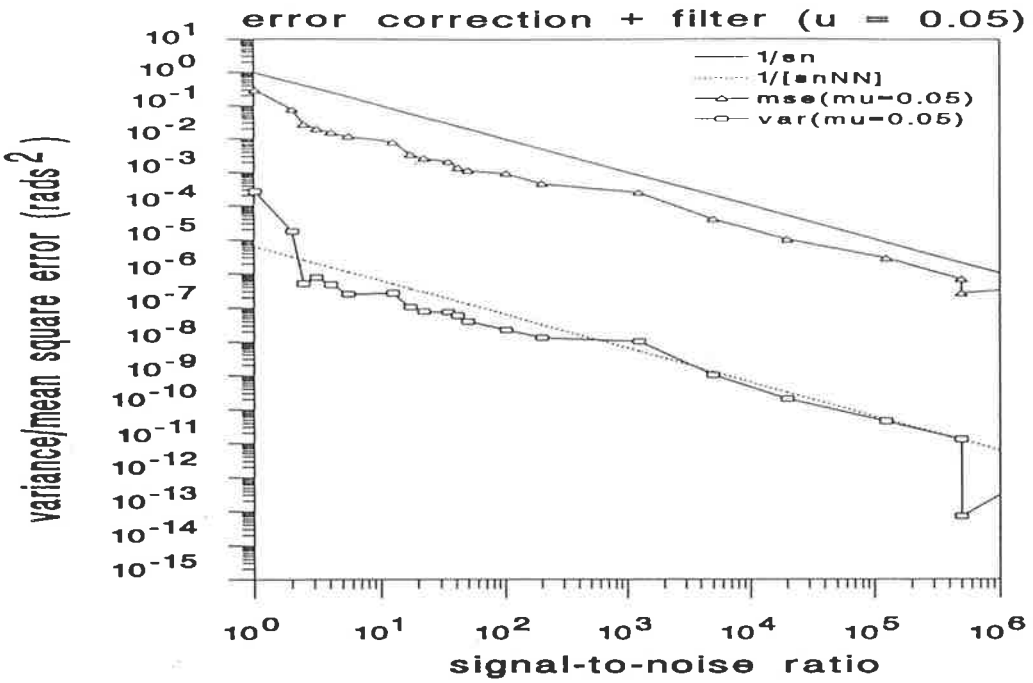

Figure 53 Simulation results, all error types incorporated case - showing $Var\{\Delta\theta\}$ and $Var\{\widehat{\Delta\theta}\}$, after filtering applied ($\mu = 0.05$), compared with theoretical results.

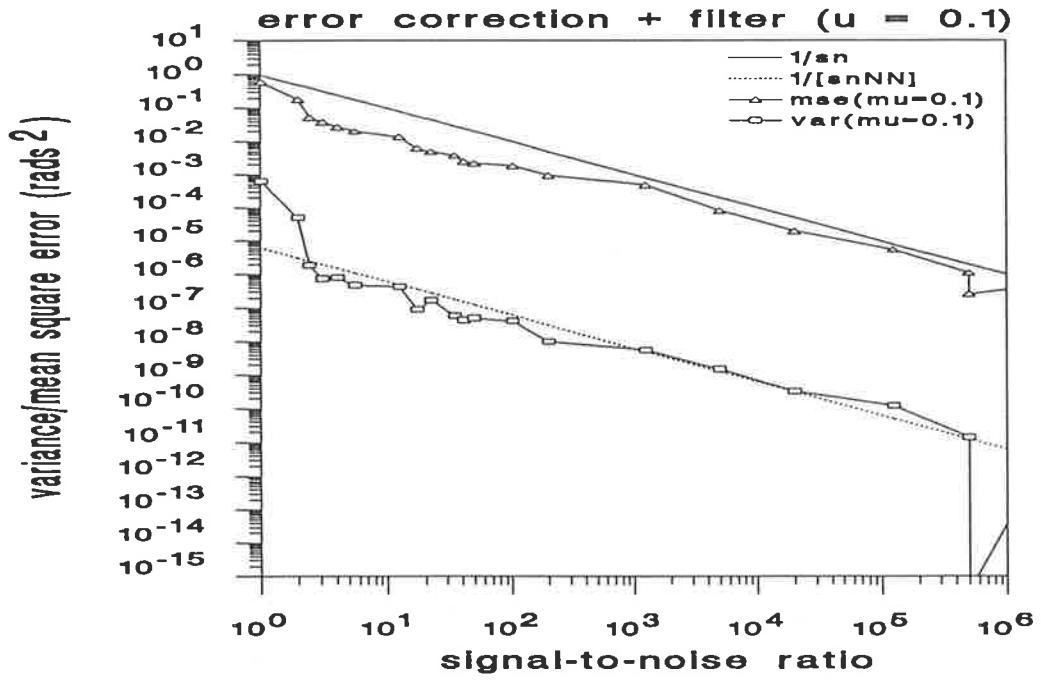

Figure 54 Simulation results, all error types incorporated case - showing $Var\{\Delta\theta\}$ and $Var\{\widehat{\Delta\theta}\}$, after filtering applied ($\mu = 0.1$), compared with theoretical results.

To demonstrate the usefulness of the techniques developed herein, real signals have been collected by a real system and studied. In the following example a $1.5\mu s$ pulsed waveform has been downconverted from the microwave frequency range 8-12GHz to an IF of 1GHz, passed through the quadrature hybrid stage and sampled by a high speed DSO. In figure 55 the signal power, as a function of time, has been calculated as the square of the signal amplitude A at each sample instant. Note the large amount of ripple or variance due to the system errors. Figure 56 shows the signal power after passing though the error correction algorithms. Here we see that the adaptive filter, in adapting the coefficients, has caused an increase in the pulse rise time and has smoothed the overall plot. Figure 57 shows the Discrete Fourier Transform of the original signal data. Note the spikes at 0Hz (ie dc) and at other image components due to the system errors. Figure 58 shows the DFT after correcting the signal waveform for errors. The plot is now much 'cleaner' ie the signal frequency is now more obvious with the image frequencies, noise etc reduced. Figure 59 shows a plot of phase difference versus time for the example pulse. In this case the signal IF frequency was offset from the reference oscillator frequency by almost 200MHz equivalent to $+\pi$ rads per sample period. Ripple is evident and, where the noise has caused the signal vector to cross the $+\pi$ boundary, spikes have appeared close to the $-\pi$ boundary. This crossing of the $\pm\pi$ boundaries can be considered as a phase difference unwrapping problem and is more likely to occur as the errors in the system increase thus restricting the input signal bandwidth. Figure 60 shows the same signal after error correction. Again start-up effects, due to the adaptive filter, are clear at the beginning of the pulse. The remainder of the phase difference plot is much smoother with no problems encountered with being close to the $+\pi$ phase difference boundary. Hence the signal input bandwidth can be maintained even for quite severe system errors.

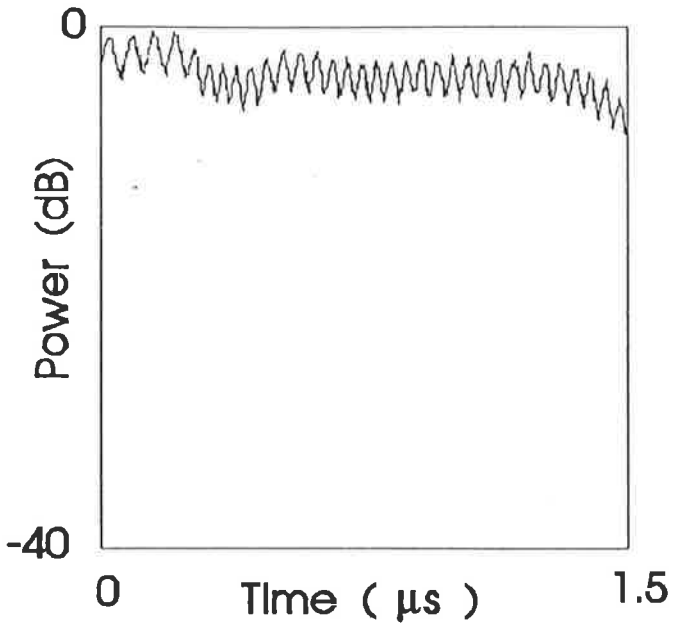

Figure 55 Real signal results, original pulse data - showing signal power versus time.

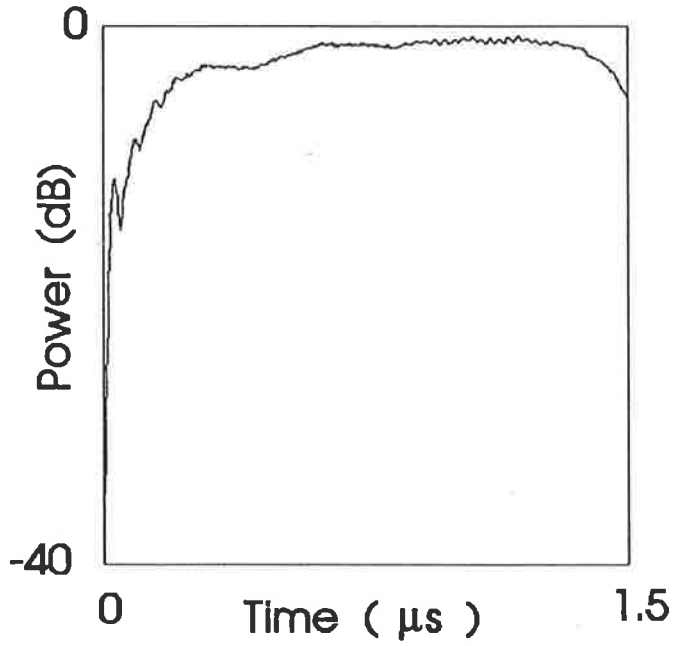

Figure 56 Real signal results, corrected pulse data - showing signal power versus time.

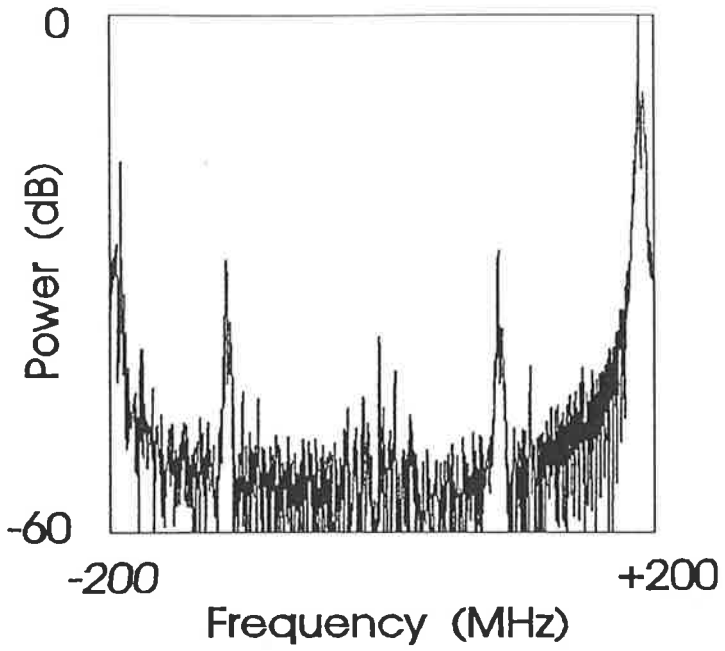

Figure 57 Real signal results, original pulse data - showing signal spectral content.

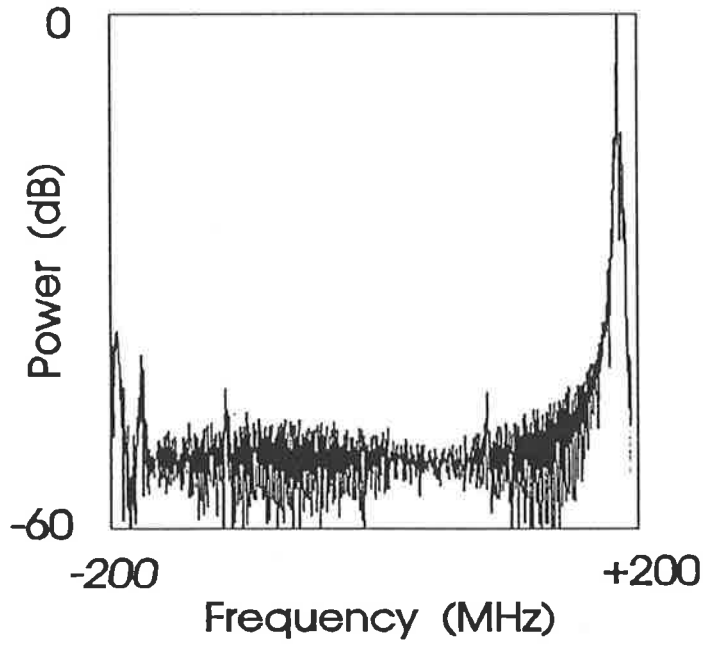

Figure 58 Real signal results, corrected pulse data - showing signal spectral content.

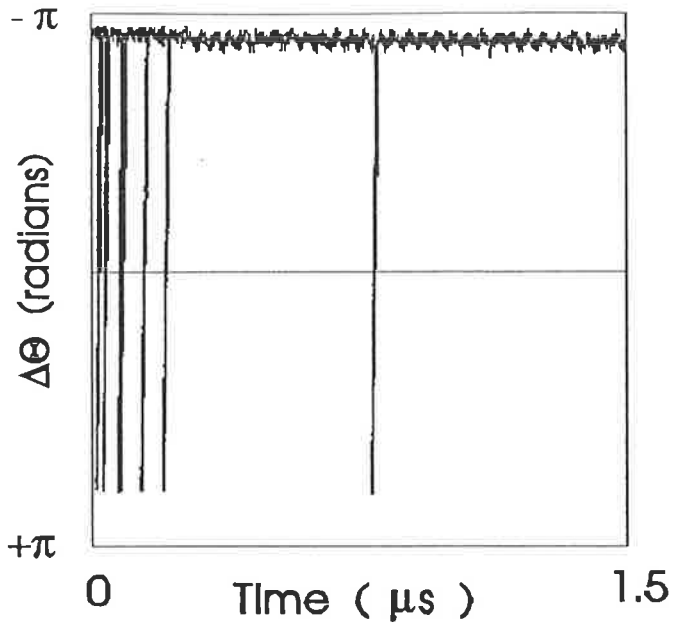

Figure 59 Real signal results, original pulse data - showing signal $\Delta\theta$ versus time.

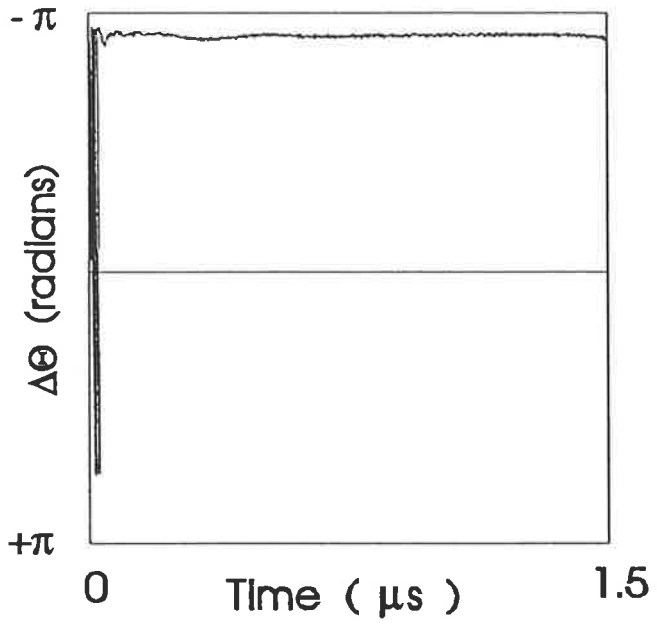

Figure 60 Real signal results, corrected pulse data - showing signal $\Delta\theta$ versus time.

6 CONCLUSIONS

In the analytic parts of this study we derived equations which related the mean square error in the phase difference values and the variance in the mean phase difference, for a number of points, to each of the major errors which might occur in a quadrature hybrid process. Of the five error types considered the signal plus noise case involved random noise whilst the other four cases were deterministic (for fixed errors or imbalances). In order to validate the analytic results a large number of computer simulations were carried out and the resulting means, mean square errors and variances recorded. These results have been plotted graphically and compare well with the theoretically derived results in all cases. Some work by Kay [7] has been incorporated in that his method for deriving the phase difference frequency estimator via a parabolic weighting function has been used. It has been shown, in both the computer simulations and experiment, that this method for frequency estimation performs efficiently (attains the Cramer-Rao lower variance bound) at signal-to-noise levels above approximately 5dB. Although developed for the signal-to-noise case, the effect of the weighting function on the other error classes has been examined. With knowledge of the effects of the errors, the next stage was to develop techniques for reducing these and some work on measuring these Harris [17] proved useful. The actual errors in the real system were measured which were in agreement with the mean square error and variance results from the system. The dc offset and amplitude imbalance cases are corrected very simply and the phase error correction is straightforward. The noise level has been effectively reduced by application of an adaptive linear predictive filter in each of the I and Q signal channels. The adaptive filter has the added benefit that it reduces the quantisation error also. We could of course have used many other methods for frequency estimation since many of these have been in existence for many years but the phase difference frequency estimator is simple to implement and is ideal for digital signal processing systems. Since we have corrected the errors in the I and Q signal waveforms before carrying out any processing, all the information about the signal is still available and has been enhanced. With many other techniques for frequency estimation the algorithms have been optimised for frequency only and any other information contained in the signal is lost. A simple example of a real signal - a single pulse from a pulsed radar system also demonstrates the value of the algorithms developed in this study which can now be used

to process real signals and eventually process these in real time digital receiver systems.

APPENDIX I

SIGNAL + NOISE CALCULATIONS

We consider first the case where each of the I and Q channels are independently corrupted by normally distributed white noise with the variance of the noise in each channel being the same and given by : $\sigma_i^2 = \sigma_q^2 = \frac{\sigma^2}{2}$. The two channel waveforms are represented by

$$\begin{aligned} I &= A\cos\theta + n_i \\ Q &= A\sin\theta + n_q \ . \end{aligned} \tag{17}$$

where $\theta = \omega_1 t$ is the signal phase at sample instant t and A is the signal amplitude. From figure 2 the instantaneous phase error due to the noise is given by θ_e. For large snr ($A \gg \sigma$), we may approximate θ_e as

$$\theta_e \approx \frac{r}{A} \tag{18}$$

where $r = |z|\cos\beta$. $|z|$ and $\cos\beta$ are independent and, from figure 2

$$\alpha + \beta + \gamma = \pi \tag{19}$$

$$\gamma = \theta - \frac{\pi}{2} \tag{20}$$

hence

$$\beta = \frac{3\pi}{2} - \alpha - \theta \ . \tag{21}$$

The expected value of $\cos\beta$ is :

$$E\{\cos\beta\} = \frac{1}{2\pi}\int_0^{2\pi} \cos\left(\frac{3\pi}{2} - \alpha - \theta\right) d\alpha = 0 \ . \tag{22}$$

The expected value of r is :

$$E\{r\} = |z|E\{\cos\beta\} = 0 \ . \tag{23}$$

The expected value of $\cos^2\beta$ is :

$$E\{\cos^2\beta\} = \frac{1}{2\pi}\int_0^{2\pi}\left[\frac{1}{2} + \cos\left(3\pi - 2\alpha - 2\theta\right) d\alpha\right] = \frac{1}{2} \tag{24}$$

and

$$E\{|z|^2\} = E\{n_i^2\} + E\{n_q^2\} = \sigma^2 \tag{25}$$

$$E\{r^2\} = E\{|z|^2\}E\{\cos\beta^2\} = \frac{\sigma^2}{2} \ . \tag{26}$$

The expected value of the instantaneous phase error is :

$$E\{\theta_e\} = \frac{E\{r\}}{A} = 0 \tag{27}$$

and the variance in the phase error is :

$$Var\{\theta_e\} = E\{\theta_e^2\} = \frac{E\{r^2\}}{A^2} = \frac{\sigma^2}{2A^2} \ . \tag{28}$$

We now consider the value of the phase difference estimator as the mean value taken over all data points,

$$\widehat{\Delta\theta} = \frac{1}{N}\sum_{i=1}^{i=N}\Delta\theta_i = \frac{1}{N}\left[\theta_{eN} - \theta_{e0}\right] + \theta_f \tag{29}$$

where θ_f is the value of the signal phase difference between successive data points.

$$\begin{aligned} Var\left\{\widehat{\Delta\theta}\right\} &= \frac{1}{N^2}\left[Var\left\{\theta_{eN}\right\} - Var\left\{\theta_{e0}\right\}\right] \\ &= \frac{1}{N^2}\left[2Var\left\{\theta_e\right\}\right] \\ &= \frac{1}{\rho N^2} \end{aligned} \tag{30}$$

where $\rho = \dfrac{A^2}{\sigma^2}$ = the signal to noise ratio .

APPENDIX II

QUANTISATION CALCULATIONS

We consider the case where each of the I and Q channels are independently quantised by the A/D sampling process. The two channel waveforms may then be represented as

$$\begin{aligned} I &= TRUNC\,[A\cos\theta] \\ Q &= TRUNC\,[A\sin\theta]\ . \end{aligned} \tag{31}$$

For large signal amplitudes ($A \gg 1\ LSB$), we may approximate θ_e as : $\theta_e \approx \frac{r}{A}$ (32)

where $r = \Delta I^2 + \Delta Q^2 = 2\Delta^2$ (33)

The phase error will be zero mean due to the symmetry around the complex circle, see fig.3. Hence we may write

$$E\{\theta_e\} = \frac{E\{r\}}{A} = 0\ . \tag{34}$$

Using the well known result for A/D quantisation noise we obtain, with q=1

$$E\{r^2\} = 2E\{\Delta^2\} = 2\left[\frac{q^2}{12}\right] = \frac{1}{6}\ . \tag{35}$$

Hence the variance in the phase error is given by : $Var\,\{\theta_e\} = \frac{1}{A^2}E\{r^2\} = \frac{1}{6A^2}$. (36)

The variance in the phase difference error is then : $Var\,\{\Delta\theta_e\} = \frac{1}{3A^2}$. (37)

The value of the phase difference estimator is : $\widehat{\Delta\theta} = \frac{1}{N}\sum_{i=1}^{i=N}\Delta\theta_i = \frac{1}{N}\left[\theta_{eN} + \theta_{e0}\right] + \theta_f$ (38)

where θ_f is the value of the signal phase difference between successive data points.

The variance in the phase difference estimate is : $Var\left\{\widehat{\Delta\theta}\right\} = \frac{2}{N^2}E\{\Delta\theta_e^2\} = \frac{1}{3A^2N^2}$ (39)

APPENDIX III

QUADRATURE PHASE ERROR CALCULATIONS

We consider the case where the Q channel is not exactly 90 degrees relative to the I channel. The two channel waveforms may then be represented as

$$\begin{aligned} I &= A\cos\theta \\ Q &= A\sin(\theta+\epsilon)\ . \end{aligned} \tag{40}$$

The amplitude of the complex vector is : $R = A\sqrt{\cos^2\theta + \sin^2(\theta+\epsilon)}\ .$ (41)

In order to find the maximum and minimum amplitude we look for the turning points in the function for R,

$$\frac{dR}{d\theta} = \left(\frac{A}{2}\right)\frac{-2\cos\theta\sin\theta + 2\sin(\theta+\epsilon)\cos(\theta+\epsilon)}{\sqrt{\cos^2\theta+\sin^2(\theta+\epsilon)}}\ . \tag{42}$$

For min and max R, $\frac{dR}{d\theta} = 0$, hence : $-2\cos\theta\sin\theta + 2\sin(\theta+\epsilon)\cos(\theta+\epsilon) = 0$ (43)

or : $-\sin(2\theta) + \sin(2\theta+2\epsilon) = 0\ .$ (44)

This reduces to : $\cos(2\theta+\epsilon)\sin\epsilon = 0$ (45)

Which is satisfied when : $\epsilon = 0, \pm n\pi$

or : $\theta = \pm\frac{n\pi}{4} - \frac{\epsilon}{2}$ n odd (46)

Substituting for phase angle θ back into the equation for R we find

$$\begin{aligned} R_{max} &= A\sqrt{1+\sin\epsilon} \quad \text{when } \theta = \frac{\pi}{4} - \frac{\epsilon}{2} \\ R_{min} &= A\sqrt{1-\sin\epsilon} \quad \text{when } \theta = \frac{-\pi}{4} - \frac{\epsilon}{2}\ . \end{aligned} \tag{47}$$

We obtain the maximum error in the signal phase by finding the turning points in the function for the phase error, where the signal phase error is given by

$$\theta_e = \theta - \tan^{-1}\left[\frac{\sin(\theta+\epsilon)}{\cos\theta}\right]\ . \tag{48}$$

Now let $x = \frac{\sin(\theta+\epsilon)}{\cos\theta}$ and substitute into the phase error function thus

$$\theta_e = \theta - \tan^{-1}[x]\ . \tag{49}$$

Using the standard identity for the derivative of $\tan^{-1} x$ we get

$$\frac{d\theta_e}{d\theta} = 1 - \frac{1}{1+x^2}\left[\frac{dx}{d\theta}\right] \tag{50}$$

which becomes, after substituting for x^2 and $\frac{dx}{d\theta}$

$$\frac{d\theta_e}{d\theta} = 1 - \frac{\sin\theta\sin(\theta+\epsilon) + \cos\theta\cos(\theta+\epsilon)}{\cos^2\theta + \sin^2(\theta+\epsilon)} . \tag{51}$$

$$\frac{d\theta_e}{d\theta} = 0 \quad \text{when } \sin\theta\sin(\theta+\epsilon) + \cos\theta\cos(\theta+\epsilon) = \cos^2\theta + \sin^2(\theta+\epsilon) \tag{52}$$

$$\text{which reduces to :} \qquad -\tan\left(\frac{\epsilon}{2}\right) = \sin(2\theta+\epsilon) \tag{53}$$

Hence the maximum phase error occurs when the signal phase

$$\theta = \frac{1}{2}\left[\sin^{-1}\left(-\tan^{-1}\frac{\epsilon}{2}\right) - \epsilon\right] \approx \frac{-3\epsilon}{4} \qquad \text{for small phase errors .} \tag{54}$$

The maximum phase error will be

$$\begin{aligned}
\theta_{emax} &= \frac{-3\epsilon}{4} - \tan^{-1}\left[\frac{\sin\frac{\epsilon}{4}}{\cos\frac{3\epsilon}{4}}\right] \\
&= \frac{-3\epsilon}{4} - \tan^{-1}\left[\frac{\sin\frac{\epsilon}{4}}{\cos\frac{\epsilon}{2}\cos\frac{\epsilon}{4} - \sin\frac{\epsilon}{2}\sin\frac{\epsilon}{4}}\right] \\
&= \frac{-3\epsilon}{4} - \tan^{-1}\left[\frac{\sin\frac{\epsilon}{4}}{\left(1-2\sin^2\frac{\epsilon}{4}\right)\cos\frac{\epsilon}{4} - 2\sin^2\frac{\epsilon}{4}\cos\frac{\epsilon}{4}}\right] \\
&= \frac{-3\epsilon}{4} - \tan^{-1}\left[\frac{\tan\frac{\epsilon}{4}}{1-4\sin^2\frac{\epsilon}{4}}\right]
\end{aligned} \tag{55}$$

and so $\theta_{emax} \approx -\epsilon$ for small phase errors.

$$\text{The maximum phase error occurs when :} \qquad \theta = \pm n\pi - \frac{3\epsilon}{4} \quad n = 0, 1, 2, \ldots \tag{56}$$

$$\text{Zero phase error occurs when :} \qquad \theta = \frac{n\pi}{2} \qquad \text{n odd .} \tag{57}$$

APPENDIX IV

AMPLITUDE IMBALANCE CALCULATIONS

We consider the case where the Q channel is not equal in amplitude to the I channel. The two channel waveforms may then be represented as

$$\begin{aligned} I &= A\cos\theta \\ Q &= A(1+\gamma)\sin\theta\,. \end{aligned} \tag{58}$$

The amplitude of the complex vector is

$$R = A\sqrt{\cos^2\theta + (1+\gamma)^2\sin^2\theta} \tag{59}$$

$$= A\sqrt{1+(2\gamma+\gamma^2)^2\sin^2\theta}\,. \tag{60}$$

Differentiating to find the maximum and minimum amplitudes

$$\frac{dR}{d\theta} = \frac{A\left(2\gamma+\gamma^2\right)\sin\theta\cos\theta}{\sqrt{1+(2\gamma+\gamma^2)^2\sin^2\theta}} \tag{61}$$

thus $\frac{dR}{d\theta} = 0$ when $\sin\theta\cos\theta = 0$. Hence we get minimum or maximum R when

$$\theta = \frac{n\pi}{2} \quad n = 0,1,2,\ldots \tag{62}$$

As expected from figure 4 we get : $R_{max} = A(1+\gamma)$

$$R_{min} = A\,. \tag{63}$$

The signal phase error is given by : $\theta_e = \theta - \tan^{-1}\left[\frac{(1+\gamma)\sin\theta}{\cos\theta}\right]$. (64)

The maximum phase error will be obtained when $\frac{d\theta_e}{d\theta} = 0$ ie

$$\frac{d\theta_e}{d\theta} = 1 - \frac{d}{d\theta}\left(\tan^{-1}\left[(1+\gamma)\tan\theta\right]\right) \tag{65}$$

$$= 1 - \left[\frac{1}{1+(1+\gamma)^2\tan^2\theta}\right](1+\gamma)\sec^2\theta \tag{66}$$

$$= 1 - \left[\frac{1+\gamma}{1+(1+\gamma)^2\tan^2\theta\cos^2\theta}\right]\,. \tag{67}$$

Hence : $\frac{d\theta_e}{d\theta} = 0$ (68)

occurs when : $(1+\gamma) = 1+(1+\gamma)^2\tan^2\theta\cos^2\theta$ (69)

which is satisfied when : $$\gamma = (2\gamma + \gamma^2)\sin^2\theta \tag{70}$$

or : $$\sin\theta = \pm\sqrt{\frac{1}{2+\gamma}}\,. \tag{71}$$

For small amplitude imbalance γ, the maximum phase error occurs at $\theta \approx \frac{\pi}{4}$. The maximum phase error is found by substituting back into the equation for the phase error as

$$\begin{aligned} \theta_{emax} &= \frac{\pi}{4} - \tan^{-1}\left[(1+\gamma)\tan\frac{\pi}{4}\right] \\ &= \frac{\pi}{4} - \tan^{-1}(1+\gamma) \end{aligned}$$

or : $$\theta_{emax} = \tan^{-1}\left[\frac{-\gamma}{2+\gamma}\right]\,. \tag{72}$$

APPENDIX V

DC OFFSET CALCULATIONS

We consider the case where the Q and the I channel are offset from the origin. The two channel waveforms may be represented as

$$\begin{aligned} I &= A\cos\theta + \delta_i \\ Q &= A\sin\theta + \delta_q \ . \end{aligned} \tag{73}$$

The amplitude of the resultant complex vector is

$$R = \sqrt{I^2 + Q^2} \tag{74}$$

$$= \sqrt{A^2 + 2A\left[\sin\theta\delta_q + \cos\theta\delta_i\right] + \delta_q^2 + \delta_i^2} \ . \tag{75}$$

We obtain the minimum and maximum amplitudes from

$$\frac{dR}{d\theta} = \frac{2A\left[\cos\theta\delta_q - \sin\theta\delta_i\right]}{\sqrt{A^2 + 2A\left[\sin\theta\delta_q + \cos\theta\delta_i\right] + \delta_q^2 + \delta_i^2}} \ . \tag{76}$$

$\frac{dR}{d\theta} = 0$ when $\cos\theta\delta_q - \sin\theta\delta_i = 0$ and this condition is satisfied by

$$\tan\theta = \frac{\delta_q}{\delta_i} = \tan(\theta + \pi) \tag{77}$$

which we then substitute back into the equation for R to get

$$R^2 = A^2 + 2A\left[\tan\theta\delta_q + \delta_i\right]\cos\theta + \delta_i^2 + \delta_q^2 \ . \tag{78}$$

We make the observation that the maximum R occurs when the signal phase equals the angle formed by the dc offsets (see fig.V.1) therefore we may make the following substitutions

$$\cos\theta = \frac{\delta_i}{\sqrt{\delta_i^2 + \delta_q^2}} \tag{79}$$

(80)

$$\sin\theta = \frac{\delta_q}{\sqrt{\delta_i^2 + \delta_q^2}} \ . \tag{81}$$

Therefore the amplitude R, in terms of A,δ_q and δ_i is

$$R^2 = A^2 + 2A\left[\frac{\delta q^2}{\delta_i} + \delta_i\right]\left[\frac{\delta_i}{\sqrt{\delta_i^2 + \delta_q^2}}\right] + \delta_i^2 + \delta_q^2 \tag{82}$$

$$= A^2 + 2A\sqrt{\delta_i^2 + \delta_q^2} + \delta_i^2 + \delta_q^2 \tag{83}$$

$$= \left[A + \sqrt{\delta_i^2 + \delta_q^2}\right]^2 \tag{84}$$

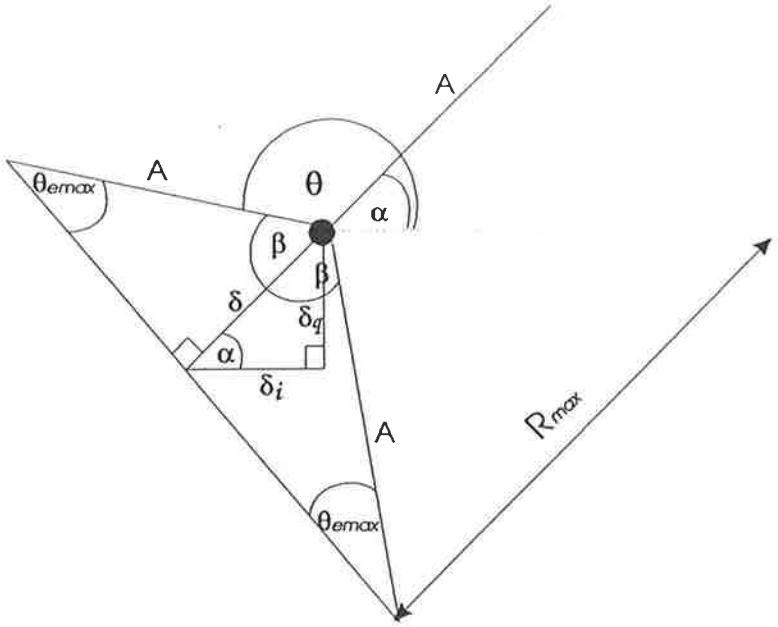

Figure V.1 DC offset construction - showing position of maximum phase error.

$$\text{Thus :} \qquad \begin{aligned} R_{max} &= A + \sqrt{\delta_i^2 + \delta_q^2} \\ R_{min} &= A - \sqrt{\delta_i^2 + \delta_q^2}\,. \end{aligned} \tag{85}$$

$$\text{The error in the signal phase is given by :} \quad \theta_e = \theta - \tan^{-1}\left[\frac{A\sin\theta + \delta_q}{A\cos\theta + \delta_i}\right]. \tag{86}$$

$$\text{Hence we shall obtain a zero phase error when :} \quad \theta = \tan^{-1}\left[\frac{A\sin\theta + \delta_q}{A\cos\theta + \delta_i}\right] \tag{87}$$

$$\tan\theta = \left[\frac{A\sin\theta + \delta_q}{A\cos\theta + \delta_i}\right] \tag{88}$$

$$\tan\theta\,[A\cos\theta + \delta_i] = A\sin\theta + \delta_q \tag{89}$$

$$\tan\theta = \frac{\delta_q}{\delta_i} \tag{90}$$

$$\text{ie zero phase error occurs when :} \qquad \theta = \tan^{-1}\left[\frac{\delta_q}{\delta_i}\right] \tag{91}$$

So we see that zero phase error occurs at the same angle as for the maximum and minimum amplitude. To find the maximum phase error we differentiate the expression for phase error and equate to zero, thus

$$\theta_e = \theta - \tan^{-1}\left[\frac{A\sin\theta + \delta_q}{A\cos\theta + \delta_i}\right] \tag{92}$$

$$\text{making the following substitutions :} \quad \begin{aligned} u &= A\sin\theta + \delta_q \\ v &= \frac{1}{A\cos\theta + \delta_i} \end{aligned} \tag{93}$$

$$\frac{du}{d\theta} = A\cos\theta \tag{94}$$
$$\frac{dv}{d\theta} = \frac{A\sin\theta}{[A\cos\theta+\delta_i]^2}$$

we obtain : $$\frac{d\theta_e}{d\theta} = 1 - \left[\frac{1}{1+\left[\frac{A\sin\theta+\delta_q}{A\cos\theta+\delta_i}\right]^2}\right]\left[\frac{[A\sin\theta+\delta_q]\,A\sin\theta}{[A\cos\theta+\delta_i]^2} + \frac{A\cos\theta}{A\cos\theta+\delta_i}\right]$$

$$= 1 - \left[\frac{[A\sin\theta+\delta_q]\,A\sin\theta + [A\cos\theta+\delta_i]\,A\cos\theta}{[A\cos\theta+\delta_i]^2 + [A\sin\theta+\delta_q]^2}\right] \tag{95}$$

and so $\frac{d\theta_e}{d\theta} = 0$ when : $$-\left(\delta_i^2+\delta_q^2\right) = A\left(\cos\theta\delta_i + \sin\theta\delta_q\right) \tag{96}$$

or : $$-\left[\frac{\delta_i^2+\delta_q^2}{A}\right] = \cos\theta\delta_i + \sin\theta\delta_q\ . \tag{97}$$

Making the following substitutions : $$\delta = \sqrt{\delta_i^2+\delta_q^2} \tag{98}$$
$$\cos\alpha = \frac{\delta_i}{\delta} \tag{99}$$
$$\sin\alpha = \frac{\delta_q}{\delta}\ . \tag{100}$$

Hence we have : $$-\left[\frac{\delta^2}{A}\right] = \delta\left[\cos\theta\frac{\delta_i}{\delta} + \sin\theta\frac{\delta_q}{\delta}\right] \tag{101}$$
$$= \delta\left[\cos\theta\cos\alpha + \sin\theta\sin\alpha\right] \tag{102}$$
$$= \delta\cos\left(\theta-\alpha\right)\ . \tag{103}$$

We now let $\cos\beta = \frac{\delta}{A}$ which gives : $$-\cos\beta = \cos\left(\theta-\alpha\right) \tag{104}$$

or : $$\cos\left(\beta+\pi\right) = \cos\left(\theta-\alpha\right) \tag{105}$$

and so the maximum phase error occurs when $\theta = \alpha+\beta+\pi$. Using the substitution which have been made we can form figure V.1 and also deduce that

$$\sin\theta_{emax} = \frac{\delta}{A} = \frac{\sqrt{\delta_i^2+\delta_q^2}}{A}\ . \tag{106}$$

APPENDIX VI

APPROXIMATE FUNCTIONAL FORMS FOR THE PHASE ERROR

In order to support the phase error calculations performed in appendices III, IV and V we can carry out small angle approximations as follows :-

$$\text{for small phase errors :} \qquad \frac{d}{dx}\tan\theta = \sec^2\theta\frac{d\theta}{dx}\,. \tag{107}$$

$$\text{Let } y = \tan\theta\text{, therefore :} \qquad \Delta y = \tan(\theta+\Delta\theta) - \tan\theta \tag{108}$$

$$\Delta\theta\sec^2\theta = \tan(\theta+\Delta\theta) - \tan\theta \tag{109}$$

$$\text{and so :} \qquad \tan(\theta+\Delta\theta) = \Delta\theta\sec^2\theta + \tan\theta \tag{110}$$

VI.1 Quadrature phase error case

For the quadrature phase error case the phase error is given by

$$-\Delta\theta = \theta_e = \theta - \tan^{-1}\left[\frac{\sin(\theta+\epsilon)}{\cos\theta}\right]\,. \tag{111}$$

Now for small ϵ, $\cos\epsilon \approx 1$ and $\sin\epsilon \approx \epsilon$ so we get

$$-\Delta\theta \approx \theta - \tan^{-1}\left[\frac{\sin\theta+\epsilon\cos\theta}{\cos\theta}\right] \tag{112}$$

$$= \theta - \tan^{-1}[\tan\theta+\epsilon] \tag{113}$$

$$\theta+\Delta\theta = \tan^{-1}[\tan\theta+\epsilon] \tag{114}$$

$$\tan[\theta+\Delta\theta] = \tan\theta+\epsilon\,. \tag{115}$$

$$\text{From above we get :} \qquad \Delta\theta\sec^2\theta + \tan\theta = \tan\theta+\epsilon \tag{116}$$

$$\text{and so :} \qquad \Delta\theta = \frac{\epsilon}{\sec^2\theta} = \epsilon\cos^2\theta \tag{117}$$

$$\text{thus the functional form of the phase error is :} \qquad \theta_e = -\epsilon\cos^2\theta\,. \tag{118}$$

We note that the maximum phase error is $-\epsilon$ as derived earlier in appendix III and this occurs when $\theta = n\pi$, in close agreement with appendix III calculations.

VI.2 Amplitude imbalance case

For the amplitude imbalance case the phase error is given by

$$-\Delta\theta = \theta_e = \theta - \tan^{-1}\left[\frac{(1+\gamma)\sin\theta}{\cos\theta}\right] \tag{119}$$

$$\theta + \Delta\theta = \tan^{-1}\left[(1+\gamma)\tan\theta\right] . \tag{120}$$

Using the small angle approximation we get

$$\Delta\theta\sec^2\theta + \tan\theta = \tan\theta + \gamma\tan\theta \tag{121}$$

$$\Delta\theta = \frac{\gamma\tan\theta}{\sec^2\theta} = \gamma\sin\theta\cos\theta \tag{122}$$

$$\text{and so :} \qquad \Delta\theta = \frac{\gamma}{2}\sin(2\theta) . \tag{123}$$

$$\text{The functional form of the phase error is :} \quad \theta_e = -\frac{\gamma}{2}\sin(2\theta) \tag{124}$$

The maximum phase error is $-\frac{\gamma}{2}$ which will occur when $\theta = \frac{n\pi}{4}$ n odd. Compare this with the exact calculation in appendix IV where the maximum phase error is given by

$$\theta_{emax} = \tan^{-1}\left[\frac{-\gamma}{2+\gamma}\right] \tag{125}$$

$$\text{at an angle given by :} \qquad \sin\theta = \pm\sqrt{\frac{1}{2+\gamma}} \tag{126}$$

VI.3 DC offset case

For the dc offset case the phase error is given by

$$-\Delta\theta = \theta_e = \theta - \tan^{-1}\left[\frac{sin\theta + \delta_q}{\cos\theta + \delta_i}\right] \tag{127}$$

$$\tan(\theta + \Delta\theta) = \left[\frac{sin\theta + \delta_q}{\cos\theta + \delta_i}\right] \tag{128}$$

and making the small angle approximation

$$\Delta\theta\sec^2\theta + \tan\theta = \left[\frac{sin\theta + \delta_q}{\cos\theta + \delta_i}\right] \tag{129}$$

$$\text{or :} \qquad \left[\Delta\theta\sec^2\theta + \tan\theta\right]\left[\cos\theta + \delta_i\right] = \sin\theta + \delta_q \tag{130}$$

$$\text{thus :} \qquad A\Delta\theta \approx \delta_q\cos\theta - \delta_i\sin\theta . \tag{131}$$

We now make the following substitutions

$$\delta = \sqrt{\delta_i^2 + \delta_q^2} \tag{132}$$

$$\begin{aligned} \cos\alpha &= \frac{\delta_i}{\delta} \\ \sin\alpha &= \frac{\delta_q}{\delta} \end{aligned} \tag{133}$$

which leads to :

$$A\Delta\theta = \delta\sin\alpha\cos\theta - \delta\cos\alpha\sin\theta \tag{134}$$

$$= \delta\sin(\alpha - \theta)\ . \tag{135}$$

The functional form for the phase error is thus

$$\theta_e = -\Delta\theta = -\frac{\delta}{A}\sin(\alpha - \theta) \tag{136}$$

$$\theta_e = -\left[\frac{\sqrt{\delta_i^2 + \delta_q^2}}{A}\right]\sin(\alpha - \theta)\ . \tag{137}$$

We see that the maximum phase error will be : $\theta_{emax} = -\dfrac{\sqrt{\delta_i^2 + \delta_q^2}}{A}$ (138)

at angles of $\theta = \frac{n\pi}{2}$. Compare this with the results from the exact calculations in appendix V where

$$\sin\theta_{emax} = -\frac{\sqrt{\delta_i^2 + \delta_q^2}}{A} \tag{139}$$

at signal phase angles of $\theta = \alpha + \beta + \pi$ where, for small phase errors, β approaches $\frac{\pi}{2}$.

APPENDIX VII

VARIANCE CALCULATIONS

We wish to calculate the statistics : mean error in phase difference and variance in the phase difference, for a single phase difference data point and also for a number, N, of consecutive phase difference data points.

VII.1 Quadrature phase error case

From appendix III we may write the functional form for the phase error as

$$\theta_e = -\theta_{emax} \cos^2\left(\theta + \frac{3\epsilon}{4}\right) = -\frac{\theta_{emax}}{2}\left[1 + \cos\left(2\theta + \frac{3\epsilon}{2}\right)\right] \tag{140}$$

and so the error in the phase difference between any two consecutive data points is

$$\Delta\theta_e = -\frac{\theta_{emax}}{2}\left[1 + \cos\left(2\theta + \frac{3\epsilon}{2}\right) - 1 - \cos\left(2\theta + \frac{3\epsilon}{2} - 2\theta_f\right)\right] \tag{141}$$

where θ_f is the angle traversed by the signal vector during one sample period.

$$\text{Thus :} \quad \Delta\theta_e = -\frac{\theta_{emax}}{2}\left[\cos\left(2\theta + \frac{3\epsilon}{2}\right) - \cos\left(2\theta + \frac{3\epsilon}{2} - 2\theta_f\right)\right] . \tag{142}$$

The phase difference error becomes

$$\Delta\theta_e = \theta_{emax}\left[\sin\left(2\theta + \frac{3\epsilon}{2} - \theta_f\right)\sin\theta_f\right] \tag{143}$$

from which we see that the mean value of phase difference is

$$mean\left\{\sin\left(2\theta + \frac{3\epsilon}{2} - \theta_f\right)\right\} = 0 \tag{144}$$

$$\text{or :} \qquad E\{\Delta\theta_e\} = 0 . \tag{145}$$

The mean square error in the phase difference, or $Var\{\Delta\theta\}$ is

$$ms\{\Delta\theta_e\} = \frac{\theta_{emax}^2}{2}\sin^2\theta_f \tag{146}$$

Consider now N consecutive data samples. The mean phase difference taken over the N samples is

$$\widehat{\Delta\theta} = \frac{1}{N}\sum_{i=1}^{i=N}\Delta\theta_i = \frac{1}{N}\sum_{i=1}^{i=N}[\theta_f + \Delta\theta_{ei}] . \tag{147}$$

Since $\widehat{\Delta\theta}$ depends only on $\Delta\theta_{e1}$ and $\Delta\theta_{eN}$ then

$$\widehat{\Delta\theta} = \frac{1}{N}\left[N\theta_f + \Delta\theta_{eN} + \Delta\theta_{e1}\right] \tag{148}$$

$$= \theta_f + \frac{1}{N}\left[\Delta\theta_{eN} + \Delta\theta_{e1}\right] . \tag{149}$$

The error in the initial phase difference is

$$\Delta\theta_{e1} = \theta_{emax}\left[\sin\left(2\theta_1 + \frac{3\epsilon}{2} - \theta_f\right)\sin\theta_f\right] \tag{150}$$

and the error in the final phase difference will be

$$\Delta\theta_{eN} = \theta_{emax}\left[\sin\left(2\theta_1 + \frac{3\epsilon}{2} + (2N-3)\,\theta_f\right)\sin\theta_f\right] . \tag{151}$$

Hence the combined phase difference error will be

$$\Delta\theta_{eN} + \Delta\theta_{e1} = \theta_{emax}\left[\sin\left(2\theta_1 + (2N-3)\,\theta_f + \frac{3\epsilon}{2}\right) + \sin\left(2\theta_1 + \frac{3\epsilon}{2} - \theta_f\right)\right]\sin\theta_f \tag{152}$$

which becomes

$$\Delta\theta_{eN} - \Delta\theta_{e1} = 2\theta_{emax}\sin\theta_f\left[\sin\left(2\theta_1 + (N-2)\,\theta_f + \frac{3\epsilon}{2}\right)\cos\left([N-1]\,\theta_f\right)\right] \tag{153}$$

and so the mean value of the combined phase difference error is

$$mean\left\{\frac{1}{N}\left(\Delta\theta_{eN} + \Delta\theta_{e1}\right)\right\} = 0 . \tag{154}$$

The mean square of the combined phase difference error or $Var\{\widehat{\Delta\theta}\}$, taken over N consecutive phase difference data points is

$$ms\left\{\frac{1}{N}\left(\Delta\theta_{eN} + \Delta\theta_{e1}\right)\right\} = \left[\frac{2\theta_{emax}^2}{N^2}\right]\sin^2\theta_f\cos^2\left([N-1]\,\theta_f\right) . \tag{155}$$

In general the variance of the phase difference error will depend on the initial and final signal phase values. However there are some special cases which we shall now consider :-

(i) If $N\theta_f = 2k\pi$ then

$$\frac{1}{N^2}ms\left\{\Delta\theta_{eN} + \Delta\theta_{e1}\right\} = \left[\frac{2\theta_{emax}^2}{N^2}\right]\sin^2\theta_f\cos^2\theta_f \tag{156}$$

$$= \left[\frac{\theta_{emax}^2}{2N^2}\right]\sin^2\left(2\theta_f\right) . \tag{157}$$

(ii) If $(N-1)\,\theta_f = \frac{k\pi}{2}$ then

$$\frac{1}{N^2}ms\left\{\Delta\theta_{eN} + \Delta\theta_{e1}\right\} = 0 . \tag{158}$$

(iii) If $(N-1)\,\theta_f = k\pi$ then

$$\frac{1}{N^2}ms\left\{\Delta\theta_{eN} + \Delta\theta_{e1}\right\} = \left[\frac{2\theta_{emax}^2}{N^2}\right]\sin^2\left(\theta_f\right) . \tag{159}$$

VII.2 Amplitude imbalance case

From appendix IV we may write the functional form for the phase error as

$$\theta_e = -\theta_{emax} \sin(2\theta) \tag{160}$$

and so the error in the phase difference between any two consecutive data points is

$$\Delta\theta_e = \theta_{emax} [\sin(2\theta) - \sin(2\theta - 2\theta_f)] \tag{161}$$

where θ_f is the angle traversed by the signal vector during one sample period.

$$\text{Thus :} \qquad \Delta\theta_e = 2\theta_{emax} [\cos(2\theta - 2\theta_f) \sin(\theta_f)] \tag{162}$$

from which we see that the mean value of phase difference is

$$mean\{\Delta\theta_e\} = 0 \tag{163}$$

$$\text{or } E\{\Delta\theta_e\} = 0 \tag{164}$$

The mean square of the phase difference error, or $Var\{\Delta\theta\}$ is

$$ms\{\Delta\theta_e\} = 2\theta_{emax}^2 \sin^2\theta_f \ . \tag{165}$$

Consider now N consecutive data samples. The mean phase difference taken over the N phase difference samples is

$$\widehat{\Delta\theta} = \frac{1}{N} \sum_{i=1}^{i=N} \Delta\theta_i = \frac{1}{N} \sum_{i=1}^{i=N} [\theta_f + \Delta\theta_{ei}] \tag{166}$$

$$\widehat{\Delta\theta} = \frac{1}{N} [N\theta_f + \Delta\theta_{eN} + \Delta\theta_{e1}] \tag{167}$$

$$= \theta_f + \frac{1}{N} [\Delta\theta_{eN} + \Delta\theta_{e1}] \tag{168}$$

since $Var\{\widehat{\Delta\theta}\}$ depends only on $\Delta\theta_{e1}$ and $\Delta\theta_{e1}$. The error in the initial phase difference is

$$\Delta\theta_{e1} = 2\theta_{emax} [\cos(2\theta_1 - \theta_f) \sin\theta_f] \tag{169}$$

and the error in the final phase difference will be

$$\Delta\theta_{eN} = 2\theta_{emax} [\cos(2\theta_1 + [2N-3]\theta_f) \sin\theta_f] \ . \tag{170}$$

Hence the combined phase difference error will be

$$\begin{aligned} \Delta\theta_{eN} + \Delta\theta_{e1} &= 2\theta_{emax} [\cos(2\theta_1 + [2N-3]\theta_f) + \cos(2\theta_1 - \theta_f)] \sin\theta_f \\ &= 2\theta_{emax} 2\cos(2\theta_1 + [N-2]\theta_f) \cos([N-1]\theta_f) \sin\theta_f \end{aligned} \tag{171}$$

and so the mean value of the combined phase difference error is

$$mean\left\{\frac{1}{N}\left(\Delta\theta_{eN}+\Delta\theta_{e1}\right)\right\}=0\,. \tag{172}$$

The mean square of the combined phase difference error, taken over N consecutive phase difference data points is

$$ms\left\{\frac{1}{N}\left(\Delta\theta_{eN}+\Delta\theta_{e1}\right)\right\}=8\left[\frac{\theta_{emax}^2}{N^2}\right]\cos^2\left(\left[N-1\right]\theta_f\right)\sin^2\theta_f\,. \tag{173}$$

In general the mean square of the phase difference error will depend on the initial and final signal phase values. However there are some special cases which we shall now consider :-

(i) If $N\theta_f=k\pi$ then

$$\begin{aligned}\frac{1}{N^2}ms\left\{\Delta\theta_{eN}+\Delta\theta_{e1}\right\} &= 8\left[\frac{\theta_{emax}^2}{N^2}\right]\cos^2\theta_f\sin^2\theta_f\\ &= 2\left[\frac{\theta_{emax}^2}{N^2}\right]\sin^2 2\theta_f\,.\end{aligned} \tag{174}$$

(ii) If $(N-1)\theta_f=\frac{k\pi}{2}$ k odd,then

$$\frac{1}{N^2}ms\left\{\Delta\theta_{eN}+\Delta\theta_{e1}\right\}=0\,. \tag{175}$$

(iii) If $(N-1)\theta_f=k\pi$ then

$$\frac{1}{N^2}ms\left\{\Delta\theta_{eN}+\Delta\theta_{e1}\right\}=8\left[\frac{\theta_{emax}^2}{N^2}\right]\sin^2\theta_f\,. \tag{176}$$

VII.3 DC offset case

We consider here the case where the dc offset in the Q channel is equal to the dc offset in the I channel. Hence we have

$$\tan^{-1}\left[\frac{\delta_q}{\delta_i}\right] = \frac{\pi}{4} \tag{177}$$

From appendix V we may write the functional form for the phase error as

$$\theta_e = -\theta_{emax}\sin(\alpha - \theta) \tag{178}$$

$$= -\theta_{emax}\sin\left(\frac{\pi}{4} - \theta\right) \tag{179}$$

and so the error in the phase difference between any two consecutive data points is

$$\Delta\theta_e = -\theta_{emax}\left[\sin\left(\frac{\pi}{4} - \theta\right) - \sin\left(\frac{\pi}{4} - \theta - \theta_f\right)\right] \tag{180}$$

$$= -2\theta_{emax}\cos\left(\frac{\pi}{4} - \theta\right)\sin\left(\frac{\theta_f}{2}\right) \tag{181}$$

where θ_f is the angle traversed by the signal vector during one sample period. We see that the mean value of phase difference is

$$mean\{\Delta\theta_e\} = 0 \tag{182}$$

$$\text{or } E\{\Delta\theta_e\} = 0\,. \tag{183}$$

The mean square error in the phase difference, or $Var\{\Delta\theta\}$ is

$$ms\{\Delta\theta_e\} = 2\theta_{emax}^2\sin^2\left(\frac{\theta_f}{2}\right) \tag{184}$$

Consider now N consecutive data samples. The mean phase difference taken over the N samples is

$$\widehat{\Delta\theta} = \frac{1}{N}\sum_{i=1}^{i=N}\Delta\theta_i = \frac{1}{N}\sum_{i=1}^{i=N}[\theta_f + \Delta\theta_{ei}] \tag{185}$$

$$\widehat{\Delta\theta} = \frac{1}{N}[N\theta_f + \Delta\theta_{eN} + \Delta\theta_{e1}] \tag{186}$$

$$= \theta_f + \frac{1}{N}[\Delta\theta_{eN} + \Delta\theta_{e1}] \tag{187}$$

since $\widehat{\Delta\theta}$ depends only on $\Delta\theta_{e1}$ and $\Delta\theta_{e1}$. The error in the initial phase difference is

$$\Delta\theta_{e1} = -2\theta_{emax}\cos\left(\frac{\pi}{4} - \theta_1 - \frac{\theta_f}{2}\right)\sin\left(\frac{\theta_f}{2}\right) \tag{188}$$

and the error in the final phase difference will be

$$\Delta\theta_{eN} = -2\theta_{emax}\cos\left(\frac{\pi}{4} - \theta_1 - [N-1]\theta_f - \frac{\theta_f}{2}\right)\sin\left(\frac{\theta_f}{2}\right)\,. \tag{189}$$

Hence the combined phase difference error will be

$$\begin{aligned}\Delta\theta_{eN} + \Delta\theta_{e1} &= -2\theta_{emax}\left[\cos\left(\frac{\pi}{4} - \theta_1 - \frac{\theta_f}{2}\right) + \cos\left(\frac{\pi}{4} - \theta_1 - \left[N - \frac{1}{2}\right]\theta_f\right)\right] sin\left(\frac{\theta_f}{2}\right) \\ &= -4\theta_{emax}\cos\left(\frac{\pi}{4} - \theta_1 - \frac{N}{2}\theta_f\right)\cos\left(\left[\frac{N-1}{2}\right]\theta_f\right)\sin\left(\frac{\theta_f}{2}\right) \qquad (190)\end{aligned}$$

and so the mean value of the combined phase difference error is

$$mean\left\{\frac{1}{N}\left(\Delta\theta_{eN} + \Delta\theta_{e1}\right)\right\} = 0\,. \qquad (191)$$

The mean square of the combined phase difference error, taken over N consecutive phase difference data points is

$$ms\left\{\frac{1}{N}\left(\Delta\theta_{eN} + \Delta\theta_{e1}\right)\right\} = 8\left[\frac{\theta_{emax}^2}{N^2}\right]\cos^2\left(\left[\frac{N-1}{2}\right]\theta_f\right)\sin^2\left(\frac{\theta_f}{2}\right)\,. \qquad (192)$$

In general the mean square of the phase difference error will depend on the initial and final signal phase values. However there are some special cases which we shall now consider :-

(i) If $N\theta_f = 2k\pi$ then

$$\frac{1}{N^2}ms\left\{\Delta\theta_{eN} + \Delta\theta_{e1}\right\} = 2\left[\frac{\theta_{emax}^2}{N^2}\right]\sin^2\theta_f\,. \qquad (193)$$

(ii) If $\left[\frac{N-1}{2}\right]\theta_f = \frac{k\pi}{2}$ k odd,then

$$\frac{1}{N^2}ms\left\{\Delta\theta_{eN} + \Delta\theta_{e1}\right\} = 0\,. \qquad (194)$$

(iii) If $\left[\frac{N-1}{2}\right]\theta_f = k\pi$ then

$$\frac{1}{N^2}ms\left\{\Delta\theta_{eN} + \Delta\theta_{e1}\right\} = 8\left[\frac{\theta_{emax}^2}{N^2}\right]\sin^2\left(\frac{\theta_f}{2}\right)\,. \qquad (195)$$

APPENDIX VIII

GAUSSIAN NOISE GENERATION

Extract from 'Numerical Recipes

Normal (Gaussian) Deviates

Transformation methods generalize to more than one dimension. If $x_1, x_2, \ldots$ are random deviates with a *joint* probability distribution $p(x_1, x_2, \ldots)\, dx_1 dx_2 \ldots$, and if $y_1, y_2, \ldots$ are each functions of all the x's (same number of y's as x's), then the joint probability distribution of the y's is

$$p(y_1, y_2, \ldots) dy_1 dy_2 \ldots = p(x_1, x_2, \ldots) \left| \frac{\partial(x_1, x_2, \ldots)}{\partial(y_1, y_2, \ldots)} \right| dy_1 dy_2 \ldots \tag{7.2.8}$$

where $|\partial(\ )/\partial(\ )|$ is the Jacobian determinant of the x's with respect to the y's (or reciprocal of the Jacobian determinant of the y's with respect to the x's).

An important example of the use of (7.2.8) is the *Box-Muller* method for generating random deviates with a normal (Gaussian) distribution,

$$p(y)dy = \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \tag{7.2.9}$$

Consider the transformation between two uniform deviates on (0,1), x_1, x_2, and two quantities y_1, y_2,

$$\begin{aligned} y_1 &= \sqrt{-2 \ln x_1} \cos 2\pi x_2 \\ y_2 &= \sqrt{-2 \ln x_1} \sin 2\pi x_2 \end{aligned} \tag{7.2.10}$$

Equivalently we can write

$$\begin{aligned} x_1 &= \exp\left[-\frac{1}{2}(y_1^2 + y_2^2)\right] \\ x_2 &= \frac{1}{2\pi} \arctan \frac{y_2}{y_1} \end{aligned} \tag{7.2.11}$$

Now the Jacobian determinant can readily be calculated (try it!)

$$\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} = -\left[\frac{1}{\sqrt{2\pi}} e^{-y_1^2/2}\right]\left[\frac{1}{\sqrt{2\pi}} e^{-y_2^2/2}\right] \tag{7.2.12}$$

Since this is the product of a function of y_2 alone and a function of y_1 alone, we see that each y is independently distributed according to the normal distribution (7.2.9).

One further trick is useful in applying (7.2.10). Suppose that, instead of picking uniform deviates x_1 and x_2 in the unit square, we instead pick v_1 and

v_2 as the ordinate and abscissa of a random point inside the unit circle around the origin. Then the sum of their squares, $R \equiv v_1^2 + v_2^2$ is a uniform deviate, which can be used for x_1, while the angle that (v_1, v_2) defines with respect to the v_1 axis can serve as the random angle $2\pi x_2$. What's the advantage? It's that the cosine and sine in (7.2.10) can now be written as $v_1/\sqrt{R}$ and $v_2/\sqrt{R}$, obviating the trigonometric function calls!

We thus have

```
      FUNCTION GASDEV(IDUM)
        Returns a normally distributed deviate with zero mean and unit variance, using RAN1(IDUM)
        as the source of uniform deviates.
      DATA ISET/0/
      IF (ISET.EQ.0) THEN              We don't have an extra deviate handy, so
1       V1=2.*RAN1(IDUM)-1.            pick two uniform numbers in the square extending from -1 to +1 in
          V2=2.*RAN1(IDUM)-1.                each direction,
          R=V1**2+V2**2                see if they are in the unit circle,
          IF(R.GE.1.)GO TO 1           and if they are not, try again.
          FAC=SQRT(-2.*LOG(R)/R)       Now make the Box-Muller transformation
          GSET=V1*FAC                  to get two normal deviates. Return one and save the other for next
          GASDEV=V2*FAC                      time.
          ISET=1                       Set flag.
      ELSE                             We have an extra deviate handy,
          GASDEV=GSET                  so return it,
          ISET=0                       and unset the flag.
      ENDIF
      RETURN
      END
```

REFERENCES AND FURTHER READING:

Knuth, Donald E. 1981, *Seminumerical Algorithms*, 2nd ed., vol. 2 of *The Art of Computer Programming* (Reading, Mass.: Addison-Wesley), pp. 116ff.

```
FUNCTION expdev(VAR idum: integer): real;
BEGIN
    expdev := -ln(ran3(idum))
END;

FUNCTION gasdev(VAR idum: integer): real;
(* Programs using GASDEV must declare the variables
VAR
    gliset: integer; glgset: real;
in the main routine and must initialize gliset to
    gliset := 0;    *)
VAR
    fac,r,v1,v2: real;
BEGIN
    IF (gliset = 0) THEN BEGIN
        REPEAT
            v1 := 2.0*ran3(idum)-1.0; v2 := 2.0*ran3(idum)-1.0;
            r := sqr(v1)+sqr(v2);
        UNTIL (r < 1.0);
        fac := sqrt(-2.0*ln(r)/r); glgset := v1*fac; gasdev := v2*fac;
        gliset := 1 END
    ELSE BEGIN
        gasdev := glgset; gliset := 0 END
END;
```

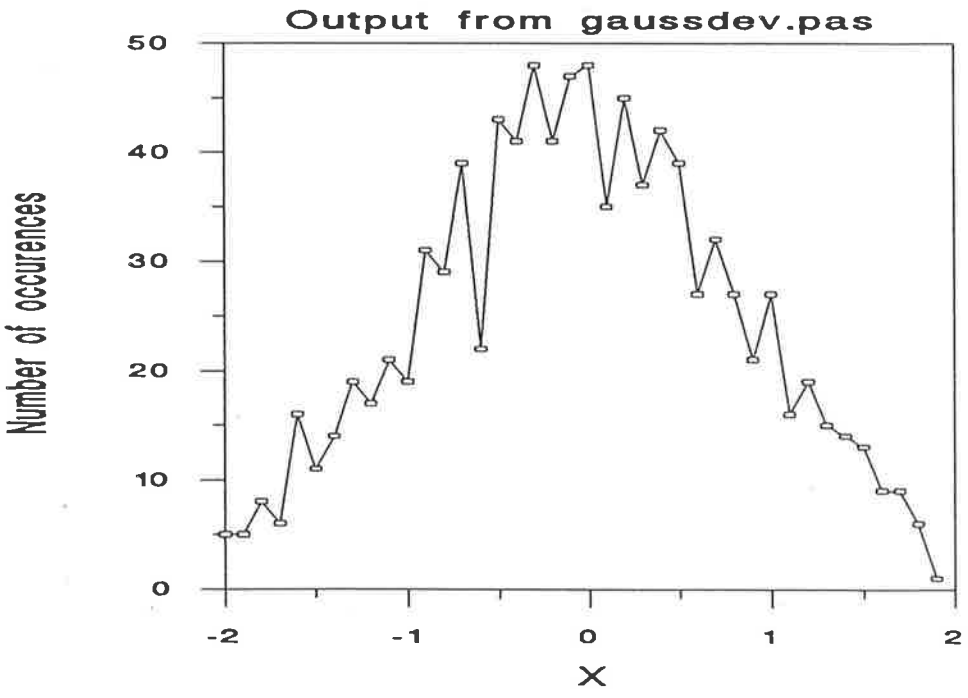

Figure VIII.1 Simulated Gaussian noise , histogram of 1000 values generated by computer program 'Gaussdev.pas'.

APPENDIX IX

THE KAY PHASE DIFFERENCE FREQUENCY ESTIMATOR

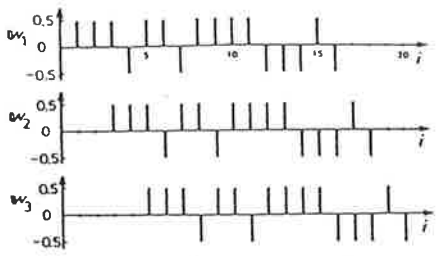

Fig. 3. An example of weighting functions for $N = 3$, $n = 16$.

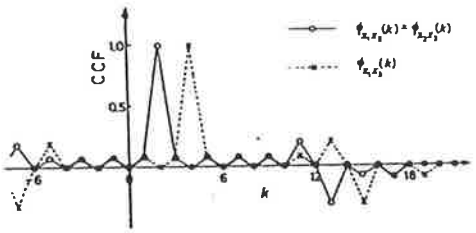

Fig. 4. CCF's for $m = 1$.

$n \times N$ weight elements are the same absolute value $2/\sqrt{n}$. And the others, $2N(N-1)$ weight elements with a value of 0, need not be implemented in the weighted adders and do not have an effect on the amplitude distribution. Therefore, the implementation of the N filters is very easy, and the amplitude distribution of every random signal $x_K(mj)$ is best approximated to a normal from the viewpoint of kurtosis.

Fig. 3 is an example of the proposed weighting functions for $N = 3$, $n = 16$. The CCF's of random signals generated by these weighting functions are given in Fig. 4 when $m = 1$. Since all CCF's are zero at $k = 6j$ ($j = 0, \pm 1, \pm 2, \cdots$), three random signals are made uncorrelated by sampling $r(j)$ every 6 shift pulses.

IV. Conclusions

A set of N weighting functions for N digital filters has been proposed based on an even-shift orthogonal sequence. The filter can generate N uncorrelated random signals from a single binary random signal. All distribution functions of the generated random signals become the same normal distribution. The generation speed is a function of N and does not depend on n. Hence, the faster generation of the random signals could be obtained. Moreover, the implementation of the filters is very easy because all $n \times N$ weight elements and input signal are binary.

References

[1] R. Zimmerman and D. L. Hunt, "Multiple-input excitation using burst random for modal testing," *Sound and Vibration*, pp. 12–21, Oct. 1985.

[2] J. L. Brown, "Generating uncorrelated random outputs by nonlinear processing of a single noise source," *IEEE Trans. Appl. Industry*, vol. 83, pp. 408–410, Nov. 1964.

[3] T. Izumi, "A method of generating multi-dimensional normally distributed random signals," *Trans. Japan Soc. Instrum. Contr. Eng.*, vol. 13, pp. 517–522, 1977.

[4] ——, "Fast generation of a white and normal random signal," *IEEE Trans. Instrum. Meas.*, vol. 37, pp. 316–318, June 1988.

A Fast and Accurate Single Frequency Estimator

STEVEN KAY

Abstract—A new frequency estimator for a single complex sinusoid in complex white Gaussian noise is proposed. The estimator is more computationally efficient than the optimal maximum likelihood estimator yet attains as good performance at moderately high signal-to-noise ratios. Also, the estimator is shown to be related to the linear prediction estimator. This relationship is exploited to reveal why the linear prediction estimator does not attain the Cramer–Rao bound even at high signal-to-noise ratios.

I. Introduction

The estimation of the frequency of a single complex sinusoid in white Gaussian noise is a problem in signal processing which has received much attention. See [1] for a summary. The optimal maximum likelihood estimator (MLE) is well known to be given by the location of the peak of a periodogram. This estimator attains the Cramer–Rao lower bound on variance for a high-enough signal-to-noise ratio (SNR). In many instances, however, the computation is prohibitive even with an FFT implementation, and so simpler methods are desirable. In this correspondence we present an approach which is strongly motivated by the recent work of Tretter [2]. It is shown that the proposed estimator is computationally much simpler than the periodogram, yet attains the Cramer–Rao bound for moderately high SNR's.

In particular, consider the received data to consist of a single complex sinusoid in complex white Gaussian noise, or

$$x_t = Ae^{j(\omega_0 t + \theta)} + z_t \qquad t = 0, 1, 2, \cdots, N-1. \tag{1}$$

The amplitude A, frequency ω_0, and phase θ are deterministic but unknown constants. It is the frequency ω_0 that we are interested in estimating. The amplitude and phase are considered to be nuisance parameters. The noise z_t is assumed to be a zero mean complex white Gaussian process with $z_t = z_{1t} + jz_{2t}$. z_{1t}, z_{2t} are each real Gaussian random variables with zero means, variances of $\sigma_z^2/2$ (σ_z^2 is the variance of z_t) and uncorrelated with each other. We now assume that the SNR, which is A^2/σ_z^2, is large, allowing the data model of (1) to be replaced by an approximate model which will form the basis for the proposed estimator. This approximate model is [2]

$$x_t = Ae^{j(\omega_0 t + \theta + u_t)} \tag{2}$$

where u_t is zero mean white Gaussian noise with variance $\sigma_z^2/2A^2$.

Denoting the phase of x_t by $\angle x_t$, we have finally

$$\angle x_t = \omega_0 t + \theta + u_t, \qquad t = 0, 1, 2, \cdots, N-1. \tag{3}$$

Having obtained (3), Tretter suggested estimating ω_0 and θ using a least squares estimator which is equivalent to an MLE. His approach provides the insightful result that frequency and phase estimation is equivalent to linear regression of the phase data. The only difficulty is that the phase needs to be unwrapped in computing these estimates. This unwrapping, besides adding to the computation, may prove to be difficult at lower SNR's. In the next section, we show how to avoid phase unwrapping but still attain the Cramer–Rao bound. Also, the proposed estimator is shown to be an improved version of a correlation or linear prediction estimator previously studied.

Manuscript received June 25, 1987; revised November 14, 1988. This work was supported by Sanders Associates, Nashua, NH, and by the Office of Naval Research under Contract N00014-87-K-0221.

The author is with the Department of Electrical Engineering, University of Rhode Island, Kingston, RI 02881.

IEEE Log Number 8931363.

II. Derivation of Estimator

Assuming that we wish to estimate only the frequency, we can avoid phase unwrapping by considering the differenced phase data

$$\Delta_t = \angle x_{t+1} - \angle x_t \tag{4}$$

for $t = 0, 1, \cdots, N-2$, which becomes, from (3),

$$\Delta_t = \omega_0 + u_{t+1} - u_t. \tag{5}$$

It is clear from (5) that the problem now is to estimate the mean, ω_0, of a colored Gaussian noise process. The process is actually a moving average with coefficients 1 and -1. The MLE of ω_0, which is equivalent to the minimum variance unbiased estimator for the linear model of (5), is found by minimizing [4]

$$J = (\Delta - \omega_0 \mathbf{1})^T C^{-1} (\Delta - \omega_0 \mathbf{1}) \tag{6}$$

where $\Delta = [\Delta_0 \;\; \Delta_1 \;\; \cdots \;\; \Delta_{N-2}]^T$, $\mathbf{1} = [1 \;\; 1 \;\; \cdots \;\; 1]^T$, and C is the $(N-1) \times (N-1)$ covariance matrix of Δ_t. The solution to this problem is well known and is

$$\hat{\omega}_0 = \frac{\mathbf{1}^T C^{-1} \Delta}{\mathbf{1}^T C^{-1} \mathbf{1}}. \tag{7}$$

Also, it can be shown that the variance of this estimator is

$$\text{Var}(\hat{\omega}_0) = \frac{1}{\mathbf{1}^T C^{-1} \mathbf{1}}. \tag{8}$$

It remains only to explicitly evaluate $\hat{\omega}_0$ and Var $(\hat{\omega}_0)$. Note that $\hat{\omega}_0$ is unbiased (let $\Delta = \omega_0 \mathbf{1} + \bar{u}$ in (7), where $[\bar{u}]_t = u_{t+1} - u_t$ for $t = 0, 1, 2, \cdots, N-2$), and that $\hat{\omega}_0$ is a Gaussian random variable, being a linear function of the data. To evaluate C^{-1}, first note that Δ_t is a real moving average process with driving noise variance $\sigma_z^2/2A^2$ and coefficients $b_0 = 1$, $b_1 = -1$. The covariance function is thus

$$c(0) = \frac{\sigma_z^2}{2A^2}(b_0^2 + b_1^2) = \frac{\sigma_z^2}{A^2}$$

$$c(1) = c(-1) = \frac{\sigma_z^2}{2A^2} b_0 b_1 = -\frac{\sigma_z^2}{2A^2}$$

$$c(k) = 0 \qquad |k| \geq 2.$$

The covariance matrix takes on the tridiagonal form

$$C = \frac{\sigma_z^2}{2A^2} \begin{pmatrix} 2 & -1 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix}.$$

The inverse is well known with the (i, j) element of the $(N-1) \times (N-1)$ matrix being given by [5]

$$[C]_{ij} = \frac{2A^2}{\sigma_z^2}\left[\min(i, j) - \frac{ij}{N}\right] \qquad 1 \leq i, j \leq N-1 \tag{9}$$

where $\min(i, j)$ denotes the minimum of i and j. After some algebra, we have that

$$\mathbf{1}^T C^{-1} \mathbf{1} = \frac{N(N^2-1)A^2}{6\sigma_z^2}$$

$$\mathbf{1}^T C^{-1} \Delta = \frac{N(N^2-1)A^2}{6\sigma_z^2} \sum_{t=0}^{N-2} w_t \Delta_t$$

where

$$w_t = \frac{3N}{N^2-1}\left\{1 - \left[\frac{t - \left(\frac{N}{2} - 1\right)}{\frac{N}{2}}\right]^2\right\}.$$

Hence, we have from (7) that

$$\hat{\omega}_0 = \sum_{t=0}^{N-2} w_t \Delta_t$$

Note that $\sum_{t=0}^{N-2} w_t = 1$ since $\hat{\omega}_0$ is an unbiased estimator (at least at high SNR). The frequency estimator may further be written by using the equivalence

$$\Delta_t = \angle x_{t+1} - \angle x_t = \angle x_t^* x_{t+1} \tag{10}$$

as

$$\hat{\omega}_0 = \sum_{t=0}^{N-2} w_t \angle x_t^* x_{t+1} \tag{11}$$

with variance which follows from (8) as

$$\text{Var}(\hat{\omega}_0) = \frac{6}{\frac{A^2}{\sigma_z^2} N(N^2-1)}. \tag{12}$$

But (12) is identical to the Cramer–Rao bound [6]. Additionally, the least squares or MLE estimator of Tretter has also been shown to attain the Cramer–Rao bound. It is clear then that $\hat{\omega}_0$ as given by (11) and Tretter's estimator must be identical. In practice, however, (11) is to be preferred since *it avoids phase unwrapping*. To verify this equivalence directly, we may rewrite (11) using (10) as

$$\hat{\omega}_0 = \frac{12}{N(N^2-1)} \sum_{t=0}^{N-1} t \angle x_t - \frac{6}{N(N+1)} \sum_{t=0}^{N-1} \angle x_t \tag{13}$$

which is identical to the linear regression estimator of Tretter. That (11) and (13) must be the same estimator is also guaranteed by the theorem that if an efficient estimator exists (i.e., it attains the Cramer–Rao bound), then it must be unique [7].

The form of the estimator given by (11) is similar to that of a previously proposed estimator as will be discussed in Section III. It is of interest to note here that w_t is a window which is symmetric about the point $t = N/2 - 1$. Some examples of this window are given in Fig. 1. As will be discussed shortly, it is this window which is responsible for $\hat{\omega}_0$ attaining the Cramer–Rao bound. If $w_t = 1/N - 1$ were chosen, for example, then $\hat{\omega}_0$ would be the sample mean of the measurements $\angle x_t^* x_{t+1}$, $t = 0, 1, \cdots, N-2$. This choice would neglect the colored noise of (5), which led to the need for C^{-1} in (6) and ultimately produced w_t in (11). In fact, for $w_t = 1/N - 1$, the estimator becomes

$$\hat{\omega}_0 = \frac{1}{N-1} \sum_{t=0}^{N-2} \angle x_t^* x_{t+1}$$

$$= \frac{1}{N-1} \sum_{t=0}^{N-2} \angle x_{t+1} - \angle x_t$$

$$= \frac{1}{N-1}(\angle x_{N-1} - \angle x_0)$$

which although unbiased [see (3)] can be expected to exhibit a large variance due to the lack of averaging. It is easily verified that for no windowing

$$\text{Var}(\hat{\omega}_0) = \frac{1}{(N-1)^2 \frac{A^2}{\sigma_z^2}} \tag{14}$$

which follows by using (3). The ratio of variances is found from (12) and (14) as

$$\frac{\text{Var}(\hat{\omega}_0)\big|_{\text{no window}}}{\text{Var}(\hat{\omega}_0)\big|_{\text{window}}} = \frac{N(N+1)}{6(N-1)} \approx \frac{N}{6} \tag{15}$$

For large data records, this loss in performance can be substantial.

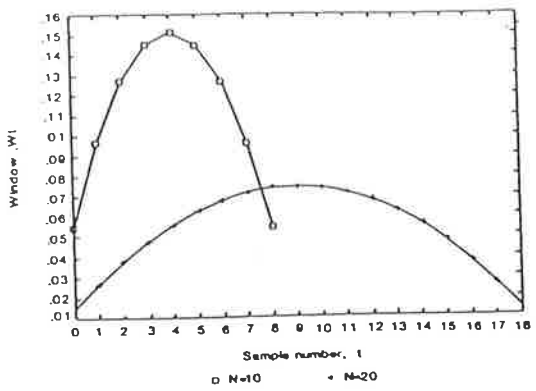

Fig. 1. Frequency estimator window.

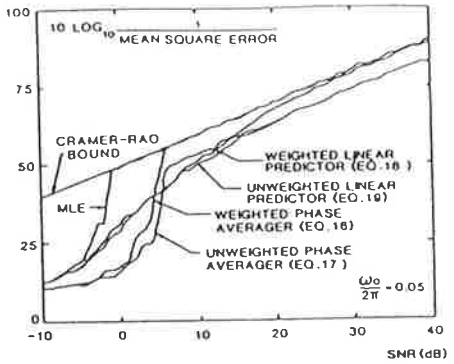

Fig. 2. Performancy of frequency estimators.

III. Relationship with Linear Prediction Estimates

The frequency estimates considered in Section II were

$$\hat{\omega}_0 = \sum_{t=0}^{N-2} w_t \angle x_t^* x_{t+1} \tag{16}$$

where

$$w_t = \frac{3N}{N^2-1}\left\{1 - \left[\frac{t - \left(\frac{N}{2}-1\right)}{\frac{N}{2}}\right]^2\right\}$$

and

$$\hat{\omega}_0 = \frac{1}{N-1}\sum_{t=0}^{N-2} \angle x_t^* x_{t+1}. \tag{17}$$

It is possible to find two additional frequency estimators which are equivalent to (16) and (17) at high SNR. These are, respectively,

$$\hat{\omega}_0 = \angle \sum_{t=0}^{N-2} w_t x_t^* x_{t+1} \tag{18}$$

and

$$\hat{\omega}_0 = \angle \frac{1}{N-1}\sum_{t=0}^{N-2} x_t^* x_{t+1}. \tag{19}$$

To form these new estimators, we have interchanged the operations of taking the angle and summation. At high SNR, the performance of (18) is identical to (16), and that of (19) is identical to (17) as shown in the Appendix. It is interesting to note that (19) has been proposed by Lank, Reed, and Pollon [3] and later studied by Jackson and Tufts [8] and by Kay [9] as a linear prediction estimator. The variance of (19) is given in [3] and is identical to (14). Our results show that a *windowed* linear prediction estimator as given by (18) is optimal in that it achieves the Cramer-Rao bound at high SNR. At lower SNR, (16) and (18) are different estimators having distinctly different performance. The same can be said about (17) and (19). Computer simulations which will be described in Section IV show, however, that (16) provides the best overall performance.

IV. Computer Simulations

A computer simulation was performed to compare the performance of the four estimators given in (16)-(19). We term (16) as the weighted phase averager, (17) as the unweighted phase averager, (18) as the weighted linear prediction estimator, and (19) as the linear prediction estimator. A data record of $N = 24$ points was used, and the mean square error versus SNR (both in decibels) determined. As a basis for comparison, the Cramer-Rao bound, which assumes an unbiased estimator, is plotted, and the performance of the MLE or location of the peak of a periodogram was also determined. The results are shown in Fig. 2 for $\omega_0/2\pi = 0.05$. As predicted by theory for high enough SNR, the MLE as well as the weighted phase averager and weighted linear prediction estimator attain the Cramer-Rao bound. The threshold for the MLE is about -1 dB, while that for the weighted phase averager is 6 dB. The weighted linear predictor does not appear to exhibit a sharp threshold but gradually deteriorates in performance with decreasing SNR. The unweighted phase averager and linear predictor exhibit the same performance at high SNR which as predicted from (15) is about $10 \log_{10} N/6 = 6$ dB below the Cramer-Rao bound. The threshold for the unweighted phase averager is about 6 dB, while for the linear predictor no threshold is apparent.

Appendix
Interchange of Angle and Summation Operators

Consider the estimators

$$\hat{\omega}_0^{(1)} = \sum_{t=0}^{N-2} w_t \angle x_t^* x_{t+1}$$

$$\hat{\omega}_0^{(2)} = \angle \sum_{t=0}^{N-2} w_t x_t^* x_{t+1}$$

where $\sum_{t=0}^{N-2} w_t = 1$.

We will show that for high SNR, these two estimators are identical. At high SNR using (4) and (5)

$$\angle x_t^* x_{t+1} = \Delta_t = \omega_0 + v_t$$

where $v_t = u_{t+1} - u_t$ so that

$$x_t^* x_{t+1} = e^{j\omega_0} e^{jv_t}$$

$$\hat{\omega}_0^{(1)} = \sum_{t=0}^{N-2} w_t(\omega_0 + v_t) = \omega_0 + \sum_{t=0}^{N-2} w_t v_t.$$

Now consider the second estimator

$$\hat{\omega}_0^{(2)} = \angle \sum_{t=0}^{N-2} w_t (e^{j\omega_0} e^{jv_t})$$

$$= \angle e^{j\omega_0} + \angle \sum_{t=0}^{N-2} w_t e^{jv_t}$$

$$= \omega_0 + \angle \sum_{t=0}^{N-2} w_t e^{jv_t}.$$

Assuming $|v_i| \ll 1$,

$$\begin{aligned}\tilde{\omega}_0^{(2)} &= \omega_0 + \angle \sum_{i=0}^{N-2} w_i(1 + jv_i) \\ &= \omega_0 + \angle\left(1 + j\sum_{i=0}^{N-2} w_i v_i\right) \\ &= \omega_0 + \arctan \sum_{i=0}^{N-2} w_i v_i \\ &\approx \omega_0 + \sum_{i=0}^{N-2} w_i v_i \\ &= \tilde{\omega}_0^{(1)}.\end{aligned}$$

Acknowledgment

The author would like to thank B. Musicus of the Massachusetts Institute of Technology for fruitful discussions of the work presented herein, and D. Sengupta of the University of California for performing the computer simulations.

References

[1] S. Kay, *Modern Spectral Estimation: Theory and Application*. Englewood Cliffs, NJ: Prentice-Hall, 1988.
[2] S. A. Tretter, "Estimating the frequency of a noisy sinusoid by linear regression," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 832–835, Nov. 1985.
[3] G. W. Lank, I. S. Reed, and G. E. Pollon, "A semicoherent detection and Doppler estimation statistic," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-9, pp. 151–165, Mar. 1973.
[4] S. Zacks, *Parametric Statistical Inference*. New York: Pergamon, 1981.
[5] M. Marcus, "Basic theorems in matrix theory," National Bureau of Standards, Appl. Math. Series, 1960.
[6] D. C. Rife and R. R. Boorstin, "Single tone parameter estimation from discrete-time observations," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 591–598, Sept. 1974.
[7] M. Kendall and A. Stuart, *The Advanced Theory of Statistics II*. New York: Macmillan, 1979.
[8] L. B. Jackson and D. W. Tufts, "Frequency estimation by linear prediction," presented at ICASSP, Tulsa, OK, 1978.
[9] S. Kay, "Comments on 'Frequency estimation by linear prediction,'" *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, Apr. 1979.

Performance Analysis of ESPRIT and TAM in Determining the Direction of Arrival of Plane Waves in Noise

BHASKAR D. RAO AND K. V. S. HARI

Abstract—In this correspondence, two subspace based methods, ESPRIT and the Toeplitz Approximation Method (TAM), for estimating the direction of arrival (DOA) of plane waves in white noise in the case of a linear equispaced sensor array are evaluated. It is shown that the least squares version of ESPRIT and TAM result in the same estimate, and are statistically equivalent. It is shown that, asymptotically, the estimates obtained using Least Squares ESPRIT and Total Least Squares ESPRIT have the same mean squared error. Expressions for the asymptotic mean squared error in the estimates of the direction of arrival are derived for the methods. Simple closed-form expressions are derived for the one and two source case to get further insight. Computer simulations are provided to substantiate the analysis.

Manuscript received May 17, 1988; revised February 28, 1989. This work was supported by the Army Research Office under Grant DAAL-03-86-K-0107.

The authors are with the AMES Department–Systems Science, University of California, San Diego, La Jolla, CA 92093.

IEEE Log Number 8931337.

I. Introduction

Subspace based methods for estimating the Direction of Arrival (DOA) of plane waves in noise have been developed and studied by a number of researchers [1]–[4]. In this correspondence, we analyze the performance of two of these methods, ESPRIT [1] and the Toeplitz Approximation Method (TAM) [3].

MUSIC was the first method that showed the benefits of using a subspace based approach [4]. Due to length and technical considerations, an evaluation of MUSIC and a comparison to the above methods is the subject of [5]. Some theoretical results comparing MUSIC and the Minimum-Norm method can be found in [6] and [7] where a characterization of the methods was done by examining the null spectrum. Our work, motivated by these papers, characterizes the error in the estimate of the direction of arrival directly. Some comparisons of MUSIC to ESPRIT can be found in [9]. ESPRIT is, like MUSIC, a general approach, and was developed to overcome some of the computation and prior information requirements of MUSIC [1], [8], [9]. Here we only consider its performance in the context of a linear equispaced sensor array. The use of TAM for DOA estimation was first suggested in [3]. A key feature of the method is that it is based on a state space model. In fact, it will be shown that for the linear equispaced sensor array case, ESPRIT can also be described using this formulation.

The organization of the correspondence is as follows. First, TAM and ESPRIT are briefly reviewed and their relationship is established. It is shown that the least squares version of ESPRIT and TAM are statistically equivalent. It is then shown that, asymptotically, the estimates obtained using Least Squares ESPRIT and Total Least Squares ESPRIT have the same mean squared error. Expressions for the asymptotic mean squared error in the estimates of the DOA are then derived. The results are specialized for the one and two source case leading to interesting insights. Simulation results are presented and they support the analysis.

II. Problem Formulation

The problem of estimating the direction of arrival of M incoherent plane waves incident on a linear equispaced array of L sensors is considered in this correspondence. For the kth observation period (snapshot), the spatial samples of the signal plus noise are given by

$$\begin{aligned}Y_k^T &= [y_1^{(k)}, y_2^{(k)}, \cdots, y_L^{(k)}] \\ &= \left[\sum_{i=1}^{M} p_i^{(k)}, \sum_{i=1}^{M} p_i^{(k)} e^{j\omega}, \cdots, \sum_{i=1}^{M} p_i^{(k)} e^{j(L-1)\omega}\right] + N_k^T, \quad (1)\end{aligned}$$

where $\omega_i = 2\pi(d/\lambda)\sin\theta_i$, d being the separation between sensors, λ the wavelength of the incident signal, and θ_i the direction of arrival. It is assumed that N_k is a mean zero Gaussian random vector with independent elements, i.e., $E[N_k N_k^H] = \sigma_n^2 I$. The noise is assumed independent of the complex signal amplitudes $p_i^{(k)}$ which are also modeled as being jointly Gaussian. The covariance matrix P of the amplitudes whose elements are P_{ij}, where $P_{ij} = E[p_i^{(k)} p_j^{*(k)}]$ is assumed to be of rank M and has distinct eigenvalues. In this correspondence, $E[\cdot]$ and the overbar "$-$" will be used interchangeably to denote the expectation operator.[1]

[1] In this correspondence, T is used to denote transpose, $*$ to denote complex conjugate, H to denote complex conjugate transpose. Also $\hat{}$ is used to denote estimates, and the subscript s to denote parameters associated with the signal alone.

APPENDIX X

MEASURING AND REMOVING QUADRATURE HYBRID ERRORS

ON MEASURING THE GAIN AND PHASE UNBALANCE AND DC OFFSETS OF QUADRATURE A-TO-D CONVERTERS WITH AN ADAPTIVE CANCELLING FILTER

fred harris

Electrical And Computer Engineering Department
College of Engineering
San Diego State University
San Diego, California 92182

ABSTRACT

In many signal processing based systems, a time sequence corresponding to the complex envelope of an input signal is formed by in-phase and quadrature (I&Q) heterodynes followed by a pair of analog-to-digital converters (ADC). In order to obtain a specified minimum attenuation of signal artifact, the gain and phase matching between these quadrature channels (as well as the DC offsets of the converters) must be held to acceptable tolerance levels. The standard method of obtaining these desired tolerances is to estimate the value of a specific mismatch by measurement of the appropriate artifacts and then perform adjustments to improve the match. We describe here a novel technique to accurately measure the phase and gain unbalance as well as the DC offsets by use of a sequence of adaptive cancelling filters.

INTRODUCTION

Many signal processing algorithms are designed to operate on the time series obtained from the complex envelope of the signal being processed. Many examples of such processing can be found in sonar, radar, and communication systems. In many of these systems, the complex envelope is formed by an analog operation consisting of an in-phase and quadrature (I&Q) heterodyne of a real signal (with an arbitrary center frequency) followed by a pair of matched lowpass filters. The resulting pair of signals are then sampled and quantized by matched analog-to-digital converters. The standard block diagram of such a process is shown in figure 1.

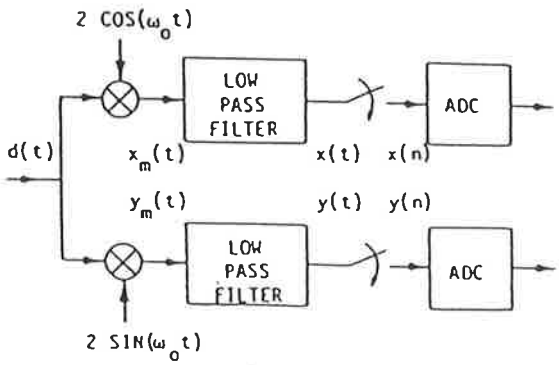

FIGURE 1. BLOCK DIAGRAM OF I-Q HETERODYNE

The operation of a quadrature heterodyne for a sinusoidal input signal is presented in (1)-(5).

$$d(t) = A\,\cos[(\omega_o+\omega_s)t] \tag{1}$$

$$x_m(t) = 2A\,\cos[(\omega_o+\omega_s)t]\,\cos[\omega_o t] \tag{2a}$$

$$= A\{\cos[\omega_s t] + \cos[(2\omega_o+\omega_s)t]\} \tag{2b}$$

$$y_m(t) = -2A\,\cos[(\omega_o+\omega_s)t]\,\sin[\omega_o t] \tag{3a}$$

$$= A\{\sin[\omega_s t] - \sin[(2\omega_o+\omega_s)t]\} \tag{3b}$$

$$x(t) = A\,\cos[\omega_s t] \tag{4a}$$

$$y(t) = A\,\sin[\omega_s t] \tag{4b}$$

$$x(t) + jy(t) = A\{\cos[\omega_s t] + j\sin[\omega_s t]\} \tag{5a}$$

$$= A\,e^{j\omega_s t} \tag{5b}$$

The spectra of $x(t)$, $y(t)$, $jy(t)$, and $x(t)+jy(t)$ are presented in figure 2. Note in particular that the addition of the real and imaginary signals (ie., $x+jy$) results in exact cancellation of their negative frequency components. The cancellation is exact because the spectral terms are of equal and opposite size.

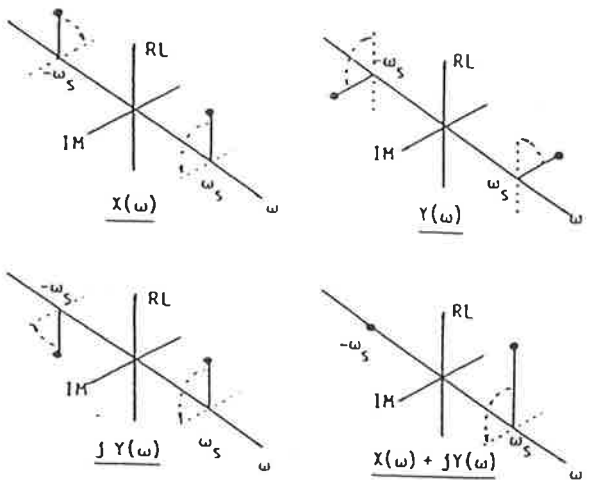

FIGURE 2. SPECTRA AT OUTPUT OF I-Q HETERODYNE

TWENTY FIRST ANNUAL ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS, AND COMPUTERS
NOVEMBER 2-4, 1987

Due to component tolerances of the analog mixers and their low pass filters, the conversion gain of the two channels are not precisely equal and the phase shift between the two channels may not be exactly 90 degrees. For ease of discussion, we arbitrarily attribute all of the mismatch terms to one of the mixers. We do so by assuming that the conversion gain of the cosine and sine heterodynes are 1.0 and $(1.0+\epsilon)$ respectively and that the phase shift between the cosine and (nominal) sine is $(90-\alpha)$ degree. We further assume that the two errors (ϵ and α) can be of either polarity. Equations (6)-(10) present the operation of the I-Q heterodyne incorporating the gain and phase errors.

$$d(t) = A\,COS[(\omega_0+\omega_s)t] \tag{6}$$

$$x_m(t) = 2A\,COS[(\omega_0+\omega_s)t]\,COS[\omega_0 t] \tag{7a}$$

$$= A\{COS[\omega_s t] + COS[(2\omega_0+\omega_s)t]\} \tag{7b}$$

$$y_m(t) = -2(1+\epsilon)A\,COS[(\omega_0+\omega_s)t]\,SIN[\omega_0 t-\alpha] \tag{8a}$$

$$= A(1+\epsilon)\{SIN[\omega_s t+\alpha] - SIN[(2\omega_0+\omega_s)t-\alpha]\} \tag{8b}$$

$$x(t) = A\,COS[\omega_s t] \tag{9a}$$

$$y(t) = A(1+\epsilon)\,SIN[\omega_s t+\alpha] \tag{9b}$$

$$x(t)+jy(t) = A\{COS[\omega_s t] + j(1+\epsilon)SIN[\omega_s t+\alpha]\} \tag{10a}$$

$$= A\{[1+(1+\epsilon)e^{j\alpha}]e^{j\omega_s t} + [1-(1+\epsilon)e^{-j\alpha}]e^{-j\omega_s t}\} \tag{10b}$$

These equations offer little insight into the effects of the gain and phase unbalance execept we note the existance of the negative frequency component which did not appear when the unbalance terms were absent. Figure 3. presents the spectra of $x(t)$, $y(t)$, $jy(t)$, and $x(t)+jy(t)$ for the unbalanced mixers. We see here that the primary effect of the gain and phase unbalance has been the failure to achieve exact cancellation of the negative frequency components. These residual negative frequency components are usually referred to as spectral images or ghosts.

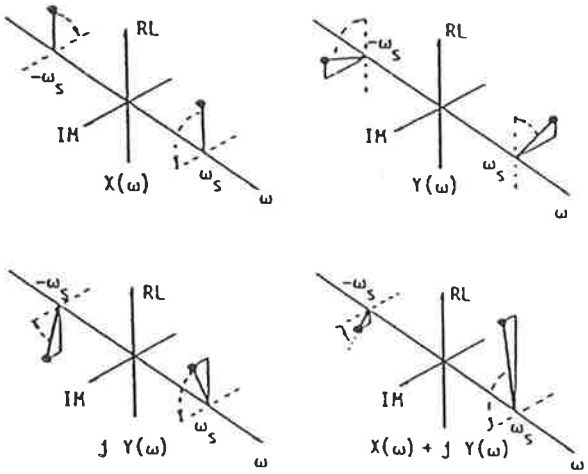

FIGURE 3. SPECTRA AT OUTPUT OF I-Q HETERODYNE EXHIBITING GAIN AND PHASE UNBALANCE

The parameters required to correct these unbalance terms are normally determined during a calibration cycle by processing the I & Q components obtained with a known sinusoidal input test signal. In the absence of other signals, the pair of sample auto correlation terms and the sample cross correlation for the I & Q channels would result in the expressions shown in (11).

$$R_{xx}(0) = A^2 \tag{11a}$$

$$R_{yy}(0) = A^2(1+\epsilon)^2 \tag{11b}$$

$$R_{xy}(0) = A^2(1+\epsilon)\,SIN(\alpha) \tag{11c}$$

We can easily compute the unbalance terms ϵ and α, as well as the signal level A, from the paramaters listed in (11). We note, however, that the measurement and processing techniques required to form the terms of (11) are affected by the ADC imperfections such as deviations from linearity and by quantization noise as well as by leakage effects related to the measurement of the low level (image and offset) signals in the presence of the high level (primary) signal.

We thus present an alternate technique to determine the desired offset terms by use of a set of adaptive, one tap, cancelling filters. These cancelling filters whiten the output signal by successively estimating and removing the known spectral components at the positive and negative frequency (as well as the ADC DC offsets). By whitening the output signal , one component at a time, each succeeding estimation and whitening task is simplified by not having to contend with the leakage terms associated with the previously removed signal components. Further, the signal remaining after the partial whitening can be used as a quality measurement of the ADC since it contains the remaining effects of the ADC non-linearities and quantizing noise.

ONE-TAP, ADAPTIVE, CANCELLING FILTER

Figure 4. presents the block diagram of a one-tap, complex weight, adaptive cancelling filter.

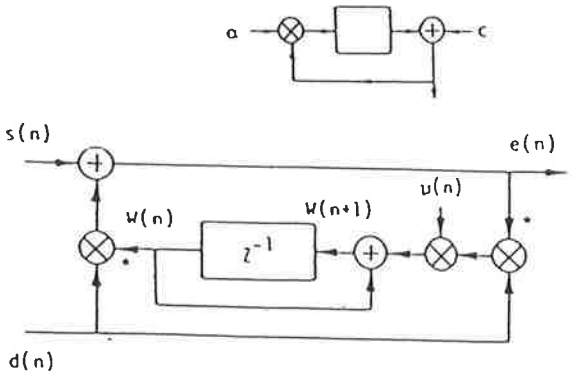

FIGURE 4. BLOCK DIAGRAM OF A ONE-TAP, COMPLEX WEIGHT, ADAPTIVE CANCELLING FILTER.

The equations describing this filter are presented in (12). Depending on the structure of the paramameter $\mu(n)$, this filter can be viewed and derived from a number of perspectives. These include the Steepest Descent (SD), Least Mean Square (LMS), and Recursive Least Square (RLS).

$$W(n+1) = W(n) + \mu(n)\cdot e^{*}(n)\cdot d(n) \qquad (12a)$$

$$e(n) = s(n) - W^{*}(n)\cdot d(n) \qquad (12b)$$

Under reasonable assumptions of first and second order statistics, the filter weight $W(n)$ converges to the steady state solution "W" which minimizes the mean square prediction error shown in (13).

$$J(n) = E\{e(n)\cdot e^{*}(n)\} \qquad (13a)$$

$$= E\{[W^{*}(n)\cdot d(n) - s(n)][W(n)\cdot d^{*}(n) - s^{*}(n)]\} \qquad (13b)$$

For the application we are describing here, the signals are ergodic so we may replace the ensemble average with a time(index) average.

The optimal weight W of course satisfies the (single weight) Wiener-Hopf equation shown in (14).

$$E\{d(n)\cdot d^{*}(n)\}\cdot W = E\{d(n)\cdot s^{*}(n)\} \qquad (14)$$

Let us assume that the desired signal $d(n)$ and the input signal $s(n)$ are of the form show in (15).

$$d(n) = e^{j\theta_0 n} \qquad (15a)$$

$$s(n) = Ae^{j\phi}e^{j\theta_0 n} + c(n) \qquad (15b)$$

We require that $d(n)$ and $c(n)$ are orthogonal as indicated in (16).

$$E\{d(n)\cdot c^{*}(n)\} = 0 \qquad (16)$$

It is easy to demonstrate (by substitution and by use of the constraint in (16)) that for the indicated $d(n)$ and for $0 < |\mu(n)| < 2$ the expected steady state solution for $W(n)$ in (12) is given in (17).

$$E\{W(n)\} = W = Ae^{j\phi} \qquad (17)$$

The transient time to achieve this expected weight is related to the size of $\mu(n)$. If $\mu(n)$ is replaced by a constant μ, the transient time constant is approximately $1/|\mu|$. We see then, that if μ is small, the transient time $(4/|\mu|)$ is large and we might be tempted to select a large μ to realize a short transient time.

If we define J_{min} as the minimum mean square error of (13), due to the signal components orthogonal to the desired signal $d(n)$, we can determine that the steady state mean square error of (13) is approximated by (18). The second term in (18) is called the excess mean square error and is due to the adaption process.

$$J(n) = J_{min} + |\Delta W(n)|^2\cdot|s(n)|^2 \qquad (18a)$$

$$= J_{min} + |\mu(n)|^2 J_{min}[A^2 + J_{min}] \qquad (18b)$$

$$= J_{min}\{1 + |\mu(n)|^2 A^2[1 + J_{min}/A^2]\} \qquad (18c)$$

$$\approx J_{min}\{1 + |\mu(n)|^2 A^2\} \qquad (18d)$$

Thus while the weight $W(n)$ converges in the mean to the optimum "W" (as in (17)) the mean square error exhibits an excess component proportional to $|\mu(n)|^2$. Here we see that in order to control the excess mean square error, we require $\mu(n)$ to very small.

We are faced with conflicting requirements on $\mu(n)$, from one consideration, we want $\mu(n)$ to be large to reduce the transient time to achieve steady state (in the mean) and from another consideration we want $\mu(n)$ to be small to control the excess mean square error. We respond to these dual requirements by scheduling $\mu(n)$ to be large early in the adaption process and then small late in the same process. As long as the scheduled changes proceed at rates slower than the instantaneous adaption time constant, both a rapid transient and a small excess mean square error can be achieved.

One easily generated scheduling technique is to use the exponential decay of a single pole filter as the $\mu(n)$ function. Speed of the adaption process then becomes dependent upon the initial $\mu(0)$ and its exponential rate of decay while the final excess mean square error becomes dependent on the input to the filter. This scheduling technique is the one shown in figure 4.

APPLYING THE ONE-TAP ADAPTIVE CANCELLING FILTER

An input signal with a known frequency offset (from the quadrature mixing frequency) is applied to the I-Q heterodyne, filters, and ADC chain. The output time series will contain a spectral component at the expected offset frequency as well as an image component due to the gain and phase unbalances of the heterodyne process. A cascade pair of adaptive filters can estimate and cancel both components by using a single complex reference signal of unit amplitude at the known offset frequency. Relative phasing between the reference and the components being cancelled is not important.

The filters are staged so that the primary signal, being the largest, is processed and cancelled first. The output of the first canceller is passed to the second filter which estimates and cancels the image component. The reference signal to the image canceller is the conjugate of the reference signal used by the first stage.

The DC offsets of the ADC are also spectral components in the output time series which can be estimated and cancelled by a third adaptive filter following the fixed tone cancellers. This filter uses a (real) unit amplitude constant as its reference signal. Figure 5. shows the form of the three cascade cancellers.

The three stage canceller can be operated in two distinct modes. In the first mode, the three filters adapt and cancel concurrently (or simultaneously). The error signal from each filter is

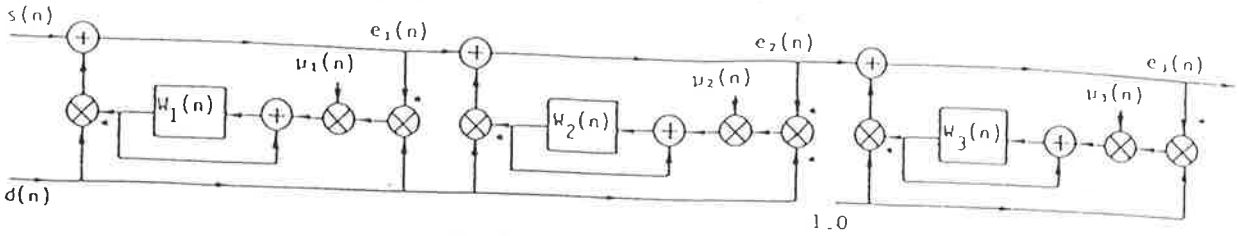

FIGURE 5. THREE STAGE CANCELLING FILTER

processed by the following filter even during its adaption and convergence process. In the second mode, the filters adapt and cancel then freeze and cancel one stage at a time. After each filter achieves its steady state response, its weight is frozen at the value obtained by averaging over at least a period of the input signal. The filter then operates as a canceller (with a frozen weight) while the next stage performs its adaption and cancellation task in preparation for its freeze and cancel mode.

The steady state output time series response is the same for the two modes of operation. This is reasonable and, in view of the robustness of adaptive systems, is expected. What is surprizing, at first glance, is that the steady state value of the filter weights obtained from the two modes are not the same.

The difference in weights can be easily explained by examining the performance of the first stage filter. We note that during adaption the filter weight is not a constant. In fact, by virtue of the additive correction of that weight, the weight is modulated with a scaled version of the input reference signal. The product of this time varying weight with the reference signal results in a DC component (the difference frequency) which is added to the input's DC component. Consequently, when the DC canceller operates on this error signal and successfully cancels the DC delivered to it, it is not the same level of DC brought in with the input signal. Hence the cancelling weights will differ.

The weight modulation is easily seen in Figures 6. This figure shows the estimates of gain and phase unbalance obtained when operating in the concurrent mode. The true gain and phase unbalance terms for this example were 0.909 and 10 degrees respectively as indicated by the "truth" lines on the curves. The oscillation of the estimates about the true level is due to the modulation of the weights during simultaneous adaption.

For the unbalance described in (10) and for the reference (input signal) leading the primary signal component in s(n) by ϕ, the weights of the first two stages converge to the terms in (19).

$$W1 = A[1+(1+\varepsilon)]\, e^{j\alpha}\, e^{j\phi} \qquad (19a)$$

$$W2 = A[1-(1+\varepsilon)]\, e^{-j\alpha}\, e^{-j\phi} \qquad (19b)$$

As shown in (20), by simple manipulation of (19) we can determine the magnitude and phase unbalance terms.

$$W1 - W2^{*} = 2A\, e^{j\phi} \qquad (20a)$$

$$W1 + W2^{*} = 2A\, e^{j\phi}\,(1+\varepsilon)e^{j\alpha} \qquad (20b)$$

$$\frac{W1 + W2^{*}}{W1 - W2^{*}} = (1+\varepsilon)e^{j\alpha} \qquad (20c)$$

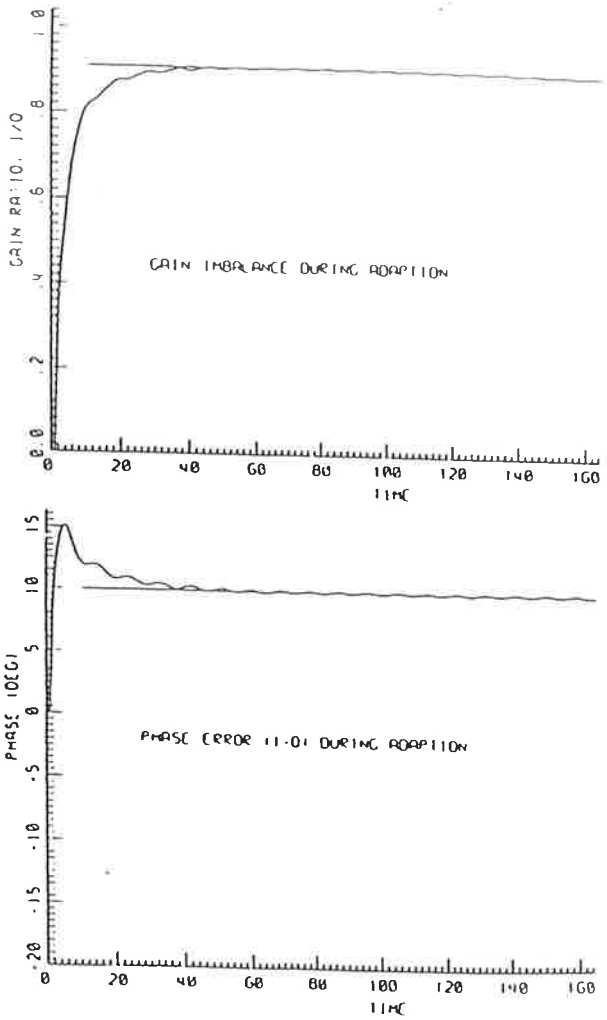

FIGURE 6. ESTIMATES OF GAIN AND PHASE UNBALANCE WEIGHTS FORMED DURING SIMULTANEOUS CANCELLING

EXAMPLE OF CANCELLING TECHNIQUE

To demonstrate the effectiveness of this technique we have simulated a signal equivalent to that obtained from an I - Q heterodyne exhibiting a +10% gain unbalance, a +10 degree phase mismatch, a -40 dB third harmonic distortion, a +10% and +5% DC additive bias for the I&Q channels respectively, and a -60 dB additive noise to simulate ADC quantizing noise. This signal was processed in a three stage converge and freeze mode cancelling filter.

The unbalance terms estimated by the canceller for the conditions decsribed above are presented in Table 1. for two levels of quantizing noise.

	ACTUAL	σ_q^2=0.000	σ_q^2=0.001
I/Q GAIN RATIO	0.90909	0.90909	0.91553
I-Q PHASE ERROR	10.0000	9.9998	9.9394
I-CHAN DC BIAS	0.1000	0.1000	0.0989
Q-CHAN DC BIAS	0.0500	0.0500	0.0520

TABLE 1. PERFORMANCE COMPARISON FOR UNBALANCE TERMS OBTAINED BY ADAPTIVE CANCELLERS

Figure 7. presents the instantaneous power for this composite signal (top curve) as well as the instantaneous power for the signals at the output of each successive canceller. The second curve presents the power after cancelling the primary signal, the third curve is the power after cancelling the image signal, and the last curve is the remaining power after cancelling the DC.

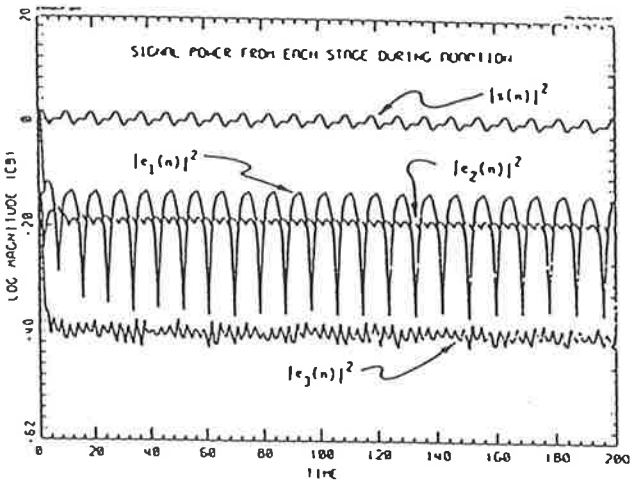

FIGURE 7. INSTANTANEOUS POWER AT INPUT AND OUTPUT OF EACH STAGE OF A CONVERGE AND FREEZE CANCELLER

Figure 8. graphs the windowed power spectrum of the composite input signal (dashed line) and the windowed power spectrum of the steady state output signal from the three stage canceller. Note that the three selected signal components have been cancelled down to the additive levels and that these additive terms consisting of third harmonic and the additive noise are essentially untouched.

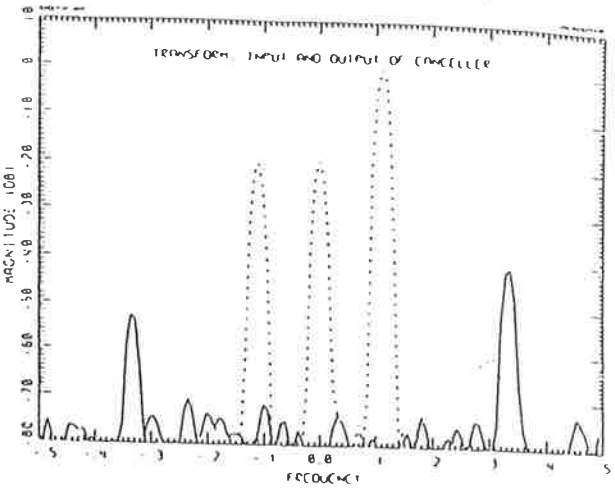

FIGURE 8. POWER SPECTRA FOR INPUT AND OUTPUT OF A THREE STAGE CONVERGE AND FREEZE CANCELLER

CONCLUSIONS

We have reviewed the requirements for balanced gain and phase in the quadrature channels of an I - Q mixer. We then described a novel technique to measure the artifacts caused by the unbalance terms. Knowing the complex amplitude of the artifacts, we are able to compute the unbalnce terms. This technique uses a tandem set of single tap adaptive cancelling filters to obtain measures of the artifacts as side information during the cancelling process. This technique is not sensitive to the presence of secondary signals such as harmonic distortion, additive noise, or DC biases.

REFERENCES

1. Simon Haykin, "Adaptive Filter Theory", (1986), Prentice-Hall, Englewood Cliffs, N.J.

2. Bernie Widrow and S. D. Stearns, "Adaptive Signal Processing",(1985), Prentice-Hall, Englewood Cliffs, N.J.

APPENDIX XI

ERROR CORRECTION CALCULATIONS

The following calculations are listed to show how the corrections for dc offset, amplitude imbalance and phase error have been derived for implementation in the computer programs.

We shall consider the amplitude and phase errors to both be in the Q channel and a different dc offset in each of the I and Q channels as follows

$$\begin{aligned} I_t &= A\cos(\omega t) + dc_i \\ Q_t &= (1+\gamma)\,A\sin(\omega t + \epsilon) + dc_q \,. \end{aligned} \tag{196}$$

XI.1 DC offset correction

The mean dc offsets in each channel can be calculated independently as

$$\begin{aligned} mean\,\{dc_i\} &= \frac{1}{N}\sum_{i=1}^{i=N} I_i \\ mean\,\{dc_q\} &= \frac{1}{N}\sum_{i=1}^{i=N} Q_i \,. \end{aligned} \tag{197}$$

To remove the dc offsets then it is simply a matter of subtracting dc_i and dc_q from the I signal vector and the Q signal vector respectively.

XI.2 Amplitude imbalance correction

If we perform an autocorrelation, with zero shift, on each channel we obtain

$$R_{II}(0) = m.s.\,\{A\cos(\omega t)\} = \frac{A^2}{2} \tag{198}$$

and

$$R_{QQ}(0) = m.s.\,\{(1+\gamma)\,A\sin(\omega t + \epsilon)\} \tag{199}$$

$$= \frac{A^2\,(1+\gamma^2)}{2} \tag{200}$$

and the cross correlation between the I and Q channels gives

$$R_{IQ}(0) = m.s.\left\{(1+\gamma)A^2 \sin(\omega t+\epsilon)\cos(\omega t)\right\} \quad (201)$$

$$= \frac{A^2(1+\gamma)}{2}\sin\epsilon\,. \quad (202)$$

Therefore we may estimate the amplitude imbalance from

$$\gamma = \sqrt{\frac{R_{QQ}}{R_{II}}} - 1 \quad (203)$$

and we may estimate the quadrature phase error from

$$\sin\epsilon = R_{IQ}\left[\frac{2}{A^2(1+\gamma)}\right]\,. \quad (204)$$

XI.3 Phase error correction

We require

$$I = A\cos\theta \quad (205)$$

$$Q = A\sin\theta\,.$$

we have

$$i = A\cos\theta = I \quad (206)$$

$$q = A\sin(\theta+\epsilon)\,.$$

If we expand out q we obtain

$$q = A\sin\theta\cos\epsilon + A\cos\theta\sin\epsilon \quad (207)$$

$$= Q\cos\epsilon + I\sin\epsilon\,.$$

Hence

$$\begin{bmatrix} i \\ q \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \sin\epsilon & \cos\epsilon \end{bmatrix}\begin{bmatrix} I \\ Q \end{bmatrix} \quad (208)$$

$$\begin{bmatrix} I \\ Q \end{bmatrix} = \frac{1}{\cos\epsilon}\begin{bmatrix} \cos\epsilon & 0 \\ -\sin\epsilon & 1 \end{bmatrix}\begin{bmatrix} i \\ q \end{bmatrix}\,. \quad (209)$$

APPENDIX XII

TABLES OF RESULTS

The variables used in tables XII.1 to XII.5 are defined as follows :-

$\widehat{\Delta\theta}$ = expected value of phase difference difference taken over the N phase difference data points

θ_f = angle through which the complex signal vector rotates
ie phase change of signal between sample instants

δ_i = dc offset in the I channel

δ_q = dc offset in the Q channel

γ = relative amplitude gain error

ϵ = quadrature phase error

N = the number of phase difference samples

$\rho = \frac{A^2}{\sigma^2}$ = the signal-to-noise power ratio

Table XII.1 Summary of results - signal plus noise case.

$Var\{\Delta\theta\}$	$\frac{1}{\rho}$
$Var\left\{\widehat{\Delta\theta}\right\}$	$\frac{1}{\rho N^2}$
$E\{\Delta\theta_e\}$	0

Table XII.2 Summary of results - quantisation case.

$Var\{\Delta\theta\}$	$\frac{1}{3A^2}$
$Var\left\{\widehat{\Delta\theta}\right\}$	$\frac{1}{3A^2N^2}$
$E\{\Delta\theta_e\}$	0

Table XII.3 Summary of results - quadrature phase error case.

min and max R when	$\theta = \pm\left(\frac{n\pi}{4} - \frac{\epsilon}{2}\right) \quad n\ odd$
R_{min}	$A\sqrt{1-\sin(\epsilon)}$
R_{max}	$A\sqrt{1+\sin(\epsilon)}$
max phase error when	$\theta = \pm n\pi - \left[\frac{3\epsilon}{4}\right] \quad n = 0, 1, 2, \ldots$
max phase error	$\theta_{emax} = -\epsilon$
zero phase error when	$\theta = -\epsilon\cos^2\theta$
form of phase error	$\theta_e = -\epsilon\cos^2\theta$
$E\{\Delta\theta_e\}$	0
$Var\{\Delta\theta\}$	$\left[\frac{\theta^2_{emax}}{2}\right]\sin^2\theta_f$
$Var\left\{\widehat{\Delta\theta}\right\}$	$\left[\frac{2\theta^2_{emax}}{N^2}\right]\sin^2\theta_f\cos^2((N-1)\theta_f)$

Table XII.4 Summary of results - amplitude imbalance case.

min and max R when	$\theta = \frac{n\pi}{2} \quad n\ odd$
R_{min}	A
R_{max}	$A(1+\gamma)$
max phase error when	$\sin\theta = \pm\sqrt{\frac{1}{2+\gamma}}$
max phase error	$\theta_{emax} \approx -\tan^{-1}\left[\frac{-\gamma}{2+\gamma}\right]$
zero phase error when	$\theta = \frac{n\pi}{2} \quad n = 0, 1, 2, \ldots$
form of phase error	$\theta_e = -\left[\frac{\gamma}{2}\right]\sin(2\theta)$
$E\{\Delta\theta_e\}$	0
$Var\{\Delta\theta\}$	$2\theta_{emax}^2 \sin^2\theta_f$
$Var\left\{\widehat{\Delta\theta}\right\}$	$8\left[\frac{\theta_{emax}^2}{N^2}\right]\cos^2((N-1)\theta_f)\sin^2\theta_f$

Table XII.5 Summary of results - dc offset case.

min and max R when	$\tan\theta = \frac{\delta_q}{\delta_i} = \tan(\theta + \pi)$
R_{min}	$A - \sqrt{\delta_i^2 + \delta_q^2}$
R_{max}	$A + \sqrt{\delta_i^2 + \delta_q^2}$
max phase error when	$\theta = \alpha \pm \beta \pm \pi$ where $\cos\beta = \frac{\sqrt{\delta_i^2+\delta_q^2}}{A}$ and $\cos\alpha = \frac{\delta_i}{\sqrt{\delta_i^2+\delta_q^2}}$
max phase error	$\sin\theta_{emax} = \left(\frac{\delta}{A}\right) = \frac{\sqrt{\delta_i^2+\delta_q^2}}{A}$
zero phase error when	$\theta = \tan^{-1}\left[\frac{\delta_q}{\delta_i}\right]$
form of phase error	$\theta_e = -\left[\frac{\sqrt{\delta_i^2+\delta_q^2}}{A}\right]\sin(\alpha - \theta)$
$E\{\Delta\theta_e\}$	0
$Var\{\Delta\theta\}$	$2\theta_{emax}^2 \sin^2\left(\frac{\theta_f}{2}\right)$
$Var\left\{\widehat{\Delta\theta}\right\}$	$8\left[\frac{\theta_{emax}^2}{N^2}\right]\cos^2\left[(N-1)\frac{\theta_f}{2}\right]\sin^2\left(\frac{\theta_f}{2}\right)$

APPENDIX XIII

TRIGONOMETRIC IDENTITIES

(210) $\sin(-\theta) = -\sin(\theta)$

(211) $\cos(-\theta) = \cos(\theta)$

(212) $\tan(-\theta) = -\tan(\theta)$

(213) $\sin(\frac{\pi}{2} + \theta) = \cos(\theta)$

(214) $\cos(\frac{\pi}{2} + \theta) = -\sin(\theta)$

(215) $\tan(\frac{\pi}{2} + \theta) = -\cot(\theta)$

(216) $\sin(\pi + \theta) = -\sin(\theta)$

(217) $\cos(\pi + \theta) = -\cos(\theta)$

(218) $\tan(\pi + \theta) = \tan(\theta)$

(219) $\cos(A + B) = \cos A \cos B - \sin A \sin B$

(220) $\cos(A - B) = \cos A \cos B + \sin A \sin B$

(221) $\sin(A + B) = \sin A \cos B + \cos A \sin B$

(222) $\sin(A - B) = \sin A \cos B - \cos A \sin B$

(223) $\cos A + \cos B = 2\cos(\frac{A+B}{2})\cos(\frac{A-B}{2})$

(224) $\cos A - \cos B = 2\sin(\frac{A+B}{2})\sin(\frac{B-A}{2})$

(225) $\sin A + \sin B = 2\sin(\frac{A+B}{2})\cos(\frac{A-B}{2})$

(226) $\sin A - \sin B = 2\cos(\frac{A+B}{2})\sin(\frac{A-B}{2})$

(227) $2\cos A \cos B = \cos(A + B) + \cos(A - B)$

(228) $2\cos A \sin B = \sin(A + B) - \sin(A - B)$

(229) $2\sin A \cos B = \sin(A + B) + \sin(A - B)$

(230) $2\sin A \sin B = \cos(A - B) - \cos(A + B)$

(231) $\cos 2\theta = 2\cos^2\theta - 1 = 1 - 2\sin^2\theta = \cos^2\theta - \sin^2\theta$

(232) $\sin 2\theta = 2\sin\theta\cos\theta$

(233) $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$

(234) $\sin 3\theta = 3\sin\theta - 4\sin^3\theta$

(235) $A\cos\theta + B\sin\theta = \sqrt{A^2 + B^2}\cos(\theta - \tan^{-1}\frac{B}{A})$

REFERENCES

[1] L. Lamport. LaTeX *A Document Preparation System.* Addison-Wesley, 1986.

[2] S.A. Hovanessian. *Radar System Design and Analysis.* Artech House, 1984.

[3] J.B. Tsui. *Microwave Receivers with Electronic Warfare applications.* Wiley, 1986.

[4] M. Bellanger. *Digital Processing of Signals.* Wiley, 1989.

[5] A. Papoulis. *Probability, random variables and stochastic processes.* McGraw-Hill, 1989.

[6] A.H. Naegeli and E.D. McHenry. Simulation and analysis of I*Q vector modulation for testing of modern microwave systems. In *RF and microwave symposium and exhibition.* Hewlett-Packard, April 1988.

[7] S. Kay. A fast and accurate single frequency estimator. *IEEE Trans. ASSP*, 37(12):1987–1990, December 1989.

[8] A.I. Sinsky and P.C.P. Wang. Error analysis of a quadrature coherent detector processor. *IEEE Trans. AES*, pages 880–883, November 1974.

[9] C.E. Persons. Quadrature sampling error formula. *Acoust Soc. Amer.*, 57(2):511–512, February 1975.

[10] J.L. Brown. On quadrature sampling of bandpass signals. *IEEE Trans. AES*, 17(1):366–370, January 1981.

[11] F.E. Chuchill, G.W. Ogar, and B.J. Thompson. The correction of I and Q errors in a coherent processor. *IEEE Trans. AES*, 17(1):131–137, January 1981.

[12] D.W. Rice and K.H. Wu. Quadrature sampling with high dynamic range. *IEEE Trans. AES*, 18(4):736–739, November 1982.

[13] CA-Cricket graph, 1990.

[14] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numerical recipes.* Cambridge University Press, 1986.

[15] J.M. Mendel. *Lessons in digital estimation theory.* Prentice-Hall, 1987.

[16] D.C. Rife and R.R. Boorstyn. Single tone parameter estimation from discrete-time observations. *IEEE Trans. Inform. Theory*, 20(5):591–598, September 1974.

[17] F. Harris. On measuring the gain and phase unbalance and dc offsets of quadrature a-to-d converters with an adaptive cancelling filter. In *21st annual ASILOMAR Conf. on signals, systems and computers*, November 1987.

[18] S.D. Stearns and R.A. David. *Signal processing algorithms.* Prentice-Hall, 1988.