A WIDEBAND ADAPTIVE FEEDFORWARD AMPLIFIER LINEARISER

by

Alison M. Smith

B.E., University of Canterbury, 1992

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF APPLIED SCIENCE
in the School
of
Engineering Science

© Alison M. Smith 1997
SIMON FRASER UNIVERSITY
August 1997

All rights reserved. This work may not be
reproduced in whole or in part, by photocopy
or other means, without the permission of the author.

National Library of Canada
Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services
Acquisitions et services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-24241-2

Canada

APPROVAL

Name: Alison M. Smith

Degree: Master of Applied Science

Title of thesis: A WIDEBAND ADAPTIVE FEEDFORWARD AMPLIFIER LINEARISER

Examining Committee:

Dr. Jamal Deen
Chair

Dr. Jim Cavers
Senior Supervisor

Dr. Mehrdad Saif
Supervisor

Dr. Andrew Wright
Examiner
Datum Telegraphic Inc.

Date Approved: 1997 08 20

Abstract

Amplifier linearisation plays a key role in improving the performance of wireless communications systems. It allows for the operation of the power amplifier in its nonlinear but power efficient region of saturation while utilising spectrally efficient modulation schemes. Of the available linearisation techniques, feedforward is presently the only technique capable of wide bandwidth linearisation. The maximum linearisation bandwidth is limited largely by the delay mismatch within the circuit and the unwanted frequency dependence of the circuit elements.

As an extension to earlier SFU research carried out on the standard circuit configuration, this thesis presents a novel feedforward configuration to enhance the wide bandwidth capability of the circuit. Solving the two key problems, the design incorporates adaptive delay matching as well as adaptive frequency compensation for the 0.2 dB frequency variations across the bandwidth of interest. In addition, it presents a method to speed the convergence time of the adaptation coefficients. By selecting appropriate frequency subbands to perform the correlations for gradient based adaptation, the problem of masking identified in the previous standard configuration is also avoided.

Linearisation bandwidths obtained are in the order of 40 MHz with 40 dB reduction of the third order intermodulation products.

Acknowledgements

I would like to express my deepest gratitude to my Senior Supervisor, Jim Cavers, who has shared a wealth of information and opened up so many windows of opportunity for me. This project has been an incredibly rewarding and challenging experience from beginning to end and I am grateful to have had experience.

I would also like to thank the members of my examining committee who gratiously accepted their positions at such late notice. Much thanks must go out to Steve Grant who put up with my banana sandwiches and helped me to take over from his project. I would especially like to thank Olly for all his support and encouragement.

With any project along came its share of problems and I'd like to thank Ian Treleaven and Kirsti Racine for helping to "dig me out of the mud" with my Windows C++ programming. In addition, I'd like to thank most of the suppliers of my circuit components for their advice and professionalism in delivery. A special thanks goes out to Tira Sakata of Cain-Sweet and Alex Bettenhausen of Hewlett Packard.

Finally I'd like to thank the suppliers of my funding; CITR and the N.Z. Guardian Trust.

Dedication

I would like to dedicate this thesis to my parents, Don and Jàn, who have always been supportive of the paths that I choose to embark on.

List of Abbreviations

The convention followed in this thesis identifies the voltages, v, and the corresponding instantaneous powers (squared voltages), x, of the individual complex baseband signals in the feedforward circuit by their associated subscripts:

v_m, x_m	modulated signal into the lineariser
v_a, x_a	main amplifier output
v_d, x_d	intermodulation distortion in main amplifier output
v_e, x_e	error signal from signal cancellation loop
v_o, x_o	feedforward amplifier output after distortion cancellation loop

Bandpass signals are represented using a tilde such that $v(t)$ is the complex envelope of the bandpass signal $\widetilde{v}(t)$

$\widetilde{v}(t) = Re[v(t)e^{j2\pi ft}]$

$x(t) = |v(t)|^2$ = instantaneous power of $v(t)$

$P = E[x(t)]$ = average power of $v(t)$

$R(\tau) = E[x(t)x^*(t-\tau)]$ = autocorrelation function

The complex coefficients used to express the adaptive attenuation and phase shift in the circuit are given below with associated relative errors and gradient signals for adaptation

α	Signal cancellation coefficient defined in original single coefficient feedforward amplifier
β	Distortion cancellation coefficient defined in single coefficient feedforward amplifier
$\epsilon_\alpha, \epsilon_\beta$	Relative error in coefficients α and β respectively
D_α	Gradient signal for adaptation of α
D_β	Gradient signal for adaptation of β
α_0, α_1	Signal cancellation coefficient's used in delay matching feedforward amplifier
β_0, β_1	Distortion cancellation coefficient's used in delay matching feedforward amplifier
D_{α_0}	Gradient signal for adaptation of α_0
D_{α_1}	Gradient signal for adaptation of α_1
D_{β_0}	Gradient signal for adaptation of β_0
D_{β_1}	Gradient signal for adaptation of β_1

Delay mismatch between the upper and lower branches is represented with the following notation

τ_α, τ_β	delay mismatch in the signal and distortion cancellation loops for single coefficient case
τ_a	delay in the main amplifier path of the delay matching signal cancellation loop
τ_1	added delay in the $\alpha_1$path of the delay matching signal cancellation loop
τ_b	delay in the upper branch of the delay matching distortion cancellation loop
τ_2	added delay in the $\beta_1$path of the delay matching distortion cancellation loop

Abbreviations used to improve the readability of the thesis are listed below

ADC	Analog to Digital Converter
DAC	Digital to Analog Converter
LUT	Look up tables
VM	Vector Modulator
PA	Power Amplifier
VCA	Voltage Controlled Attenuator
QPSK	Quadrature Phase Shift Keying
IMD	Intermodulation Distortion
IMP	Intermodulation Product
SNR	Signal to Noise Ratio

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Due to the increasing demands being placed on radio communications services and the limited channel bandwidth available, techniques to enhance the system capacity have become of great importance. Linear modulation techniques allow for greater spectral usage due to their varying amplitude and phase characteristics but require highly linear power amplifiers to avoid generation of intermodulation distortion (IMD) in adjacent channels.

It is also desirable to use power efficient RF power amplifiers in the mobile transmitter to reduce the frequency of battery recharging. Unfortunately, with power amplifier design, there is a trade-off between linearity and power efficiency. The highest efficiency is achieved when the power amplifier operates in its nonlinear region of saturation.

Constant envelope modulations, such as FM, have traditionally been used because of the ability to use nonlinear amplifiers at saturation with excellent power efficiency. They are not, in general, as spectrally efficient as linear modulation schemes such as QPSK or 16QAM.

The problem lies in how to preserve the narrow band characteristics of the spectrally efficient modulated signal while utilising a power efficient amplifier. The solution is achieved by using a nonlinear but highly power efficient Class AB amplifier and designing external linearisation circuitry to reduce the amount of IMD generated at the amplifier output. Several methods have been developed to linearise the power amplifier output, such as cartesian feedback, predistortion and the method chosen for this thesis, feedforward. All have associated advantages and drawbacks and these will be discussed in the following sections.

1.1 The Nonlinear Power Amplifier

In a nonlinear RF amplifier the input signal undergoes a level dependent gain and phase shift due to the AM/AM and AM/PM characteristics of the amplifier. The levels of IMD generated will depend on the choice of modulation scheme. Signals that pass through the cutoff region of the amplifier, like QPSK, or signal that have large varying envelopes (enhanced by pulse shaping), will exhibit higher levels of IMD in the output spectrum.

Phase distortion in the amplified signal is due to frequency dependent characteristics of the nonlinear amplifier, and the amplifier is described as possessing memory. A truly memoryless nonlinear amplifier will only exhibit amplitude distortion. For narrow bandwidth signals, where the memory time constants are significantly smaller than the reciprocal value of the bandwidth, these effects can be neglected and the power amplifier can be modelled as quasi-memoryless [6]. However, as we increase the bandwidth of our input signal the memory effect of power amplifiers becomes more noticeable. Asymmetric IM skirts clearly indicate a power amplifier with memory; however, symmetric skirts could represent either case. Modelling the effects of memory is difficult as it requires long polynomial representations to accurately describe the output as a function of the present input signal and all the previous inputs spanning the memory of the amplifier. The Volterra series [18] is practical only for polynomial nonlinearities of low degree. Although the feedforward lineariser can handle the asymmetry of the IM skirts, the power amplifier has been assumed to be quasi-memoryless, for ease in simulation and analysis.

Two-tone tests are useful for providing the memory information of the amplifier and provide a worst case approach to analysing the linearisation circuitry's ability to reduce the IM skirts. However it does not directly provide the AM/AM and AM/PM characteristics of the amplifier. Details of methods used to characterise the nonlinear amplifier can be found in [6].

For simulation and analysis purposes, the quasi-memoryless nonlinear amplifier is represented by a complex baseband function, $G(x)$, composed of the voltage gain and phase shift as a function of the instantaneous input power, x.

$$G(x) = g(x)e^{j\Phi(x)} \tag{1.1}$$

Throughout this text $G(x)$ will be used to specify the nonlinear characteristics of the

main amplifier in the feedforward lineariser. The amplifier output $v_a(t)$ is found by multiplying the input signal voltage $v_m(t)$ with the amplifier complex gain, $G(x_m(t))$, and can also be expressed in linear estimation terms where

$$\begin{aligned} v_a(t) &= v_m(t)G(x_m(t)) \\ &= \gamma_o v_m(t) + v_d(t) \end{aligned} \tag{1.2}$$

which represents the sum of the linearly amplified input signal and the intermodulation distortion $v_d(t)$. Thus γ_o is the linear gain of the amplifier.

1.2 Predistortion

Predistortion uses a predistorter inserted prior to the nonlinear power amplifier to achieve amplifier linearisation. Roughly speaking, the predistorter generates the IMD products equal in magnitude but in antiphase with those generated by the memoryless nonlinear power amplifier so that the distortion is cancelled from the amplifier output and the result is a linearly amplified version of the input to the predistorter. The predistorter complex gain, $F(x)$, is found by characterising the inverse amplitude and phase nonlinearities of the power amplifier. From the analysis in [7],the amplifier output is given as

$$v_a(t) = v_m(t)F\left[x_m(t)\right]G\left[x_m(t)|F(x_m(t)|^2\right] \tag{1.3}$$

so for complete IMD suppression the multiplication of the predistorter complex gain and the power amplifier complex gain must be equal to some constant K

$$K = F\left[x_m(t)\right]G\left[x_m(t)|F(x_m(t)|^2\right] \tag{1.4}$$

The predistorter is usually realised using a digital signal processor (DSP) which provides look up tables (LUT) containing the information required to distort the input signal depending on the magnitude of the input power. Cavers [10] proposed a table comprised of complex-valued gain factors in cartesian form which significantly reduces the number of operations needed to predistort the signal. This configuration is shown in Figure 1.1. The magnitude of the input power of the digital baseband inputs selects the appropriate complex gain factor, represented in its cartesian form, and a single complex multiplication

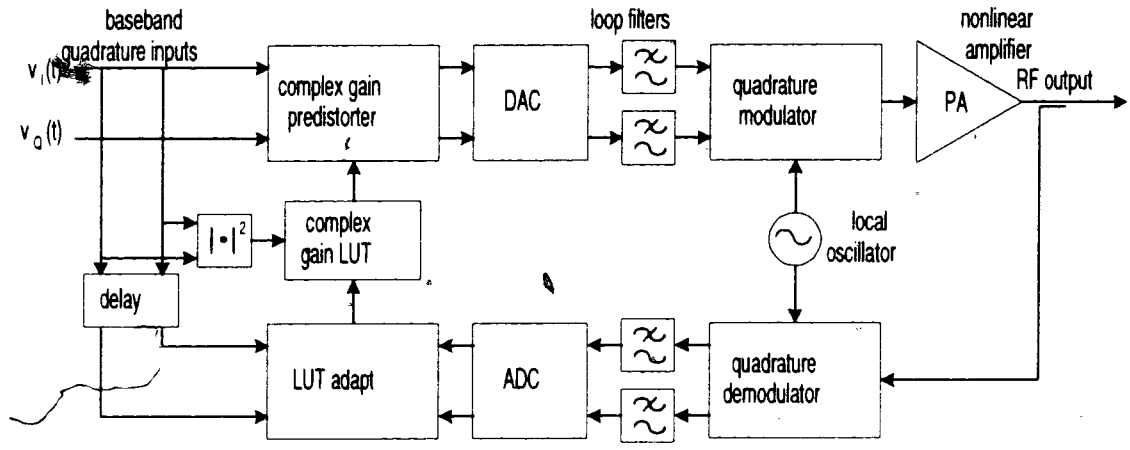

Figure 1.1: Model of Complex Gain Predistortion Lineariser

is then performed with the input signals. After D/A conversion and filtering, the predistorted signal is upconverted to the desired carrier frequency then amplified to produce a linearised output signal. Feedback is used for adaptation of the predistorter coefficients in the LUT to compensate for variations in the amplifiers nonlinear characteristics. Wright [9] showed a reduction in IMD in excess of 20 dB can be achieved using the adaptive complex gain predistorter. Complex gain predistortion has the advantage over other predistortion techniques in that it is unrestricted by modulation format and requires less table memory.

An increase in the requirement of suppression of IMD requires an increase to the size of the lookup tables, which will result in long adaptation times. A general rule is that a doubling in table size will only provide an increase of -6 dB suppression of the IMD.

Adaptive predistortion techniques using baseband DSP are very useful for low bandwidth systems [11] due to limitations of the DSP computation rate and its resulting high power consumption at high sampling rates. The predistortion lineariser cannot handle memory effects in the power amplifier but is insensitive to loop delays and large AM and PM components which cause instability problems in feedback linearisation techniques.

1.3 Cartesian Feedback

Cartesian feedback, a modulation feedback technique, is presently the most widely accepted technique for amplifier linearisation. Because it is a closed loop configuration it has the

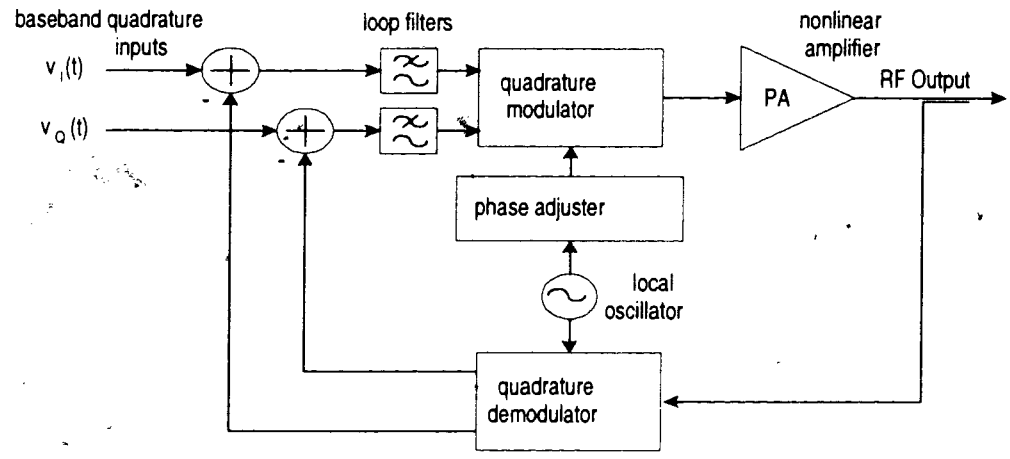

Figure 1.2: Model of Cartesian Feedback Lineariser

disadvantage of being only conditionally stable. Its operation can be explained with reference to Figure 1.2. The feedback loop demodulates the RF output signal to produce the complex signal expressed in cartesian coordinate form. These real and imaginary components are subtracted from the corresponding I and Q input signals to produce a loop error signal, which is passed through the baseband loop filters to drive the quadrature modulator and the nonlinear RF power amplifier. To preserve stability around the loop the cartesian feedback system uses a phase adjuster to adjust the phase in the quadrature modulator and demodulator and this changes for the various frequencies of operation. This added complexity results in an increased loop delay which is one of the primary limiting factors of modulation feedback. The performance of cartesian feedback is highly dependent on correct setting of the phase adjuster.

The advantage of cartesian feedback is that it can easily compensate for any changes in the operating conditions of the circuit. It compensates for any nonlinearities in the forward path devices which includes the loop filters, the baseband amplifiers, the quadrature modulator and the RF amplifier. Cartesian feedback has the advantage of simplicity and is an excellent linearisation technique for modulation bandwidths up to 500 kHz.

If the gain in the forward path is represented by A and the feedback gain is given by β, the output of the cartesian feedback lineariser as given in [13] is expressed as

$$v_{o(}t) \approx \frac{v_m(t)}{\beta} + \frac{v_d(t)}{A\beta} + v_f(t) \tag{1.5}$$

where $v_f(t)$ represents error and distortion products introduced in the feedback path. Clearly, the intermodulation distortion is reduced by an amount equal to the loop gain $A\beta$. Unfortunately the gain is also reduced by the same amount and a trade-off is made between loop gain and distortion suppression. Any additional distortion products in the feedback path introduce an uncorrectable error in the output spectrum and care must be taken in the design to keep the contribution as small as possible. The amplifier gain should be sufficiently high to allow for high gain and high distortion suppression. Increasing the gain in the feedback path results in a less stable system. As the loop gain at high frequencies is moderate the application of cartesian feedback is limited.

Multiple cartesian feedback modules in parallel can be used for broadband multicarrier linearisation [12]. Each channel has its own cartesian feedback module and local oscillator for upconversion to the unique carrier frequency. The outputs of every module are combined before ampliflication in a common amplifier. A sample of the output from the multicarrier power amplifier is fed back and demodulated by each cartesian feedback module. The experimental results showed an improvement in the IMD levels of 10-30 dB over a linearisation bandwidth of approximately 100 kHz per channel.

In contrast to predistortion [6]-[11], cartesian feedback [12]-[17] has received very little published analysis. In view of this, a more indepth analysis is presented here.

1.3.1 An Analysis of the Gain-Bandwidth-Delay Product

The primary limiting factor of the cartesian feedback lineariser is the magnitude of the propagation delay which dominates the phase characteristic at high frequencies. The gain-bandwidth-delay product provides an upper bound to the value of the product of the loop gain and the linearising bandwidth when the loop delay is known, to ensure a stability in the feedback loop. Reducing loop delay in the cartesian feedback system will allow higher loop gain or operating bandwidth.

The phase shift caused by the loop time delay at specific frequency can be calculated as

$$\Phi = 2\pi f \tau_d \tag{1.6}$$

where τ_d is the loop delay and f is the frequency offset from centre frequency. This phase

shift adds to the phase shift produced by the misadjustment of the phase adjuster represented by ϕ. A stable feedback system requires that the loop gain is less than unity when the phase around the feedback loop reaches $\pm 180^o$ and it is desirable to include some margin so that the system is not operating near the point of instability. The phase margin, φ_m, is the phase difference from $\pm 180^o$ when there is unity loop gain, and the gain margin, g_m, is the difference in gain from unity when the phase is $\pm 180^o$. The phase margin places a limit on the maximum acceptable loop delay given by

$$\tau_{d_{max}} = \frac{\pi/2 - \varphi_m - \phi}{\pi \Delta f_{cp}}, \tag{1.7}$$

where Δf_{cp} is the crossover bandwidth specified by the crossover frequencies, f_{cp+} and f_{cp-}. Phase crossover frequencies are the frequencies where the phase is equal to $\pm(180^o - \varphi_m)$. This bandwidth needs to accommodate all the significant intermodulation distortion products of the amplifer as only these products will be suppressed.

The crossover frequencies can be found in terms of the phase misadjustment and loop delay

$$\left\{ \begin{array}{c} f_{cp-} \\ f_{cp+} \end{array} \right\} = \frac{\phi}{2\pi\tau_d} \mp \left(\frac{1}{4\tau_d} \mp \frac{\varphi_m}{2\pi\tau_d} \right) \tag{1.8}$$

This shows that for imperfect adjustment of the phase adjuster the "normal" crossover frequencies are shifted by a factor dependent on the residual phase misadjustment around the loop, $\frac{\phi}{2\pi\tau_d}$. It was stated in [13] that the phase ϕ has to be adjusted to within $\pm 15^o$ to give more than 22 dB suppression of the third order intermodulation products. This means the phase has to be set to one of twelve discrete phase setting 0, 30, 60... 360. A reduction in the open loop gain leads to less stringent requirements on the phase adjustment. Experimental results obtained in [13], using a 500 kHz modulated sinewave two tone test with an amplifier biased in Class C, showed a 20 dB reduction of the third order intermodulation products at 1.5 MHz away from the carrier.

Figure 1.3 shows the simulated results of the intermodulation distortion suppression as a function of loop delay and phase misadjustment using a single pole filter at 50 kHz. The phase margin has been set to 60^o with a loop gain of 0 dB at the smallest phase crossover frequency. The loop delay is set to 60 ns and the phase misadjustment is at 15^o. A maximum loop gain for the amplifier of 11 dB is used to maintain a gain-bandwidth-delay product within the limits of stability.

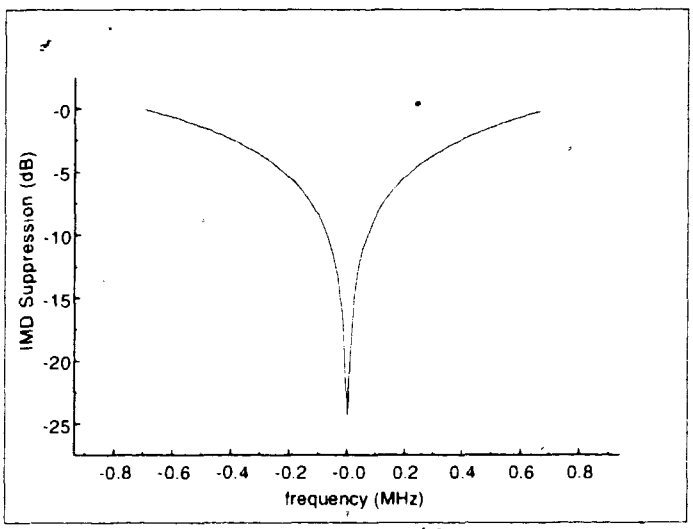

Figure 1.3: Plot of IMD suppression vs frequency of cartesian feedback lineariser with 60 ns loop delay

Due to the stringent requirements placed on the loop delay to maintain stability around the loop, DSP cannot be used to control the feedback loop as it introduces a minimum of one sample delay between the input and output.

1.4 Feedforward

Feedforward is a linearisation technique developed four years prior to the invention of negative feedback by the same inventor, H.S. Black of Bell Telephone Laboratories, but only after the published experimental results carried out by Seidel [5] did it begin to gain recognition as an effective linearisation method for wide bandwidth systems. The simple concept of an adaptive feedforward amplifier can be described with the aid of Figure 1.4. The lineariser consists of two circuits, the signal cancellation circuit followed by the distortion cancellation circuit. The signal cancellation circuit contains the nonlinear RF power amplifier (PA), which is also referred to as the main amplifier. The input signal is split into two parts to drive the PA in the upper branch and to provide a reference signal in the lower branch of the circuit. A sample of the PA output is fed down through the fixed attenuation so that it approximately matches the level of the reference signal. The reference signal is also delayed

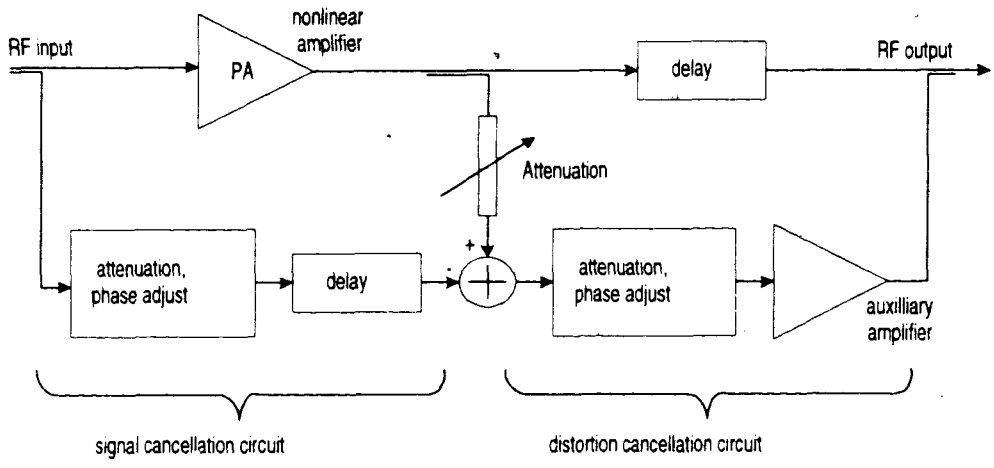

Figure 1.4: RF circuit model of an adaptive gain/phase Feedforward lineariser

by an amount equal to the group delay of the PA, to enhance the wide bandwidth capability of the lineariser. The reference signal and the PA output are then added in antiphase to produce an error signal at the output of the signal cancellation circuit.

The attenuation and phase adjustment is required to achieve complete cancellation of the reference signal from the amplifier output so that the error signal $v_e(t)$ is comprised solely of the intermodulation products expressed by $v_d(t)$. This signal is then fed into the lower branch of the distortion cancellation circuit where it again undergoes attenuation and phase adjustment. An auxiliary Class A amplifier, also referred to as the error amplifier, linearly amplifies the error signal so that the level matches that of PA intermodulation distortion, and subtracts this from the PA output to produce the final lineariser output signal which, ideally, is a linearly amplified version of the input.

The attenuation and phase adjuster in the signal cancellation circuit can be placed ahead of the PA so that any additional distortion that it may introduce can be lumped into the distortion produced by the nonlinear PA. The analysis of the feedforward amplifier with the adjuster placed in the lower branch is somewhat simpler and the behaviour of the two configurations is similar in the neighbourhood of optimum adjustment. For this reason it is preferable to model the circuit using this configuration.

The adaptive behaviour of the feedforward amplifier was first analysed by Cavers [2]. The results of this analysis are reproduced here and are explained with reference to the

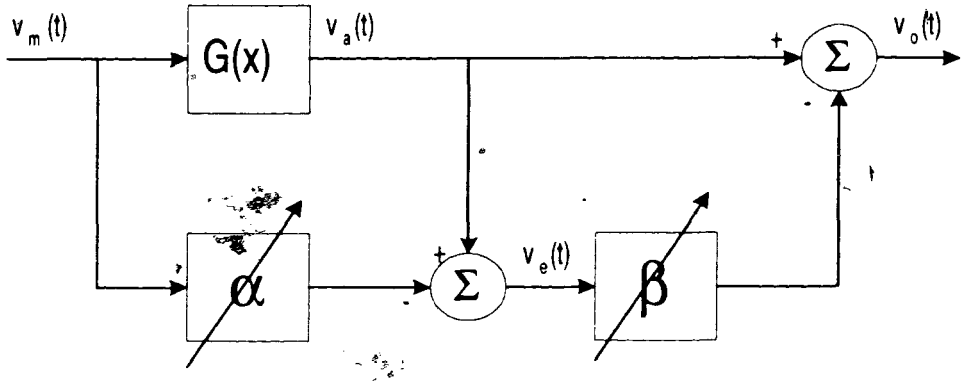

Figure 1.5: Complex baseband model of feedforward lineariser

complex baseband equivalent model of the feedforward lineariser shown in Figure 1.5. The gain and phase adjusters in the signal cancellation and distortion cancellation circuits are represented by the variable complex coefficients α and β respectively, and the nonlinear PA is represented by its complex voltage gain, $G(x)$. The error amplifier is modelled as linear with unity gain. Accurate delay matching is also assumed and no attenuation is provided by the sampling couplers or the fixed attenuation. All the components excluding the PA are considered to be lossless and have a flat response across the frequency band of interest and a linear phase response.

Since the output of the PA can be represented as the sum of a linearly amplified component plus intermodulation distortion,

$$v_a(t) = \gamma_o v_m(t) + v_d(t) \tag{1.9}$$

and the error signal is given as

$$v_\epsilon(t) = v_a(t) - \alpha v_m(t) \tag{1.10}$$

then clearly for complete signal suppression from the amplifier output, $\alpha_{opt} = \gamma_o$.

Similarly, for the distortion cancellation circuit the output is given by

$$v_o(t) = v_a(t) - \beta v_e(t) \tag{1.11}$$

To achieve complete suppression of the intermodulation distortion from the PA, then $\beta_{opt} = 1$, and the final output will be a linearly amplified version of the input.

$$v_o(t) = \gamma_o v_m(t) \tag{1.12}$$

This method has the advantage over feedback techniques because it is unconditionally stable. Moreover it is able to handle linearisation over much wider bandwidths. Like feedback, it can also adapt to changes in operating conditions due to the adaptive nature of the coefficients. It does, however, have the disadvantage of requiring two power amplifiers and, to achieve high IMD suppression, high linearity requirements are placed on the error amplifier since any error introduced in the lower branches cannot be cancelled from the final output spectrum. It is therefore important to operate the error amplifier well below its output 1dB compression point in its linear region of operation. To relieve the error amplifier of high levels of the error signal a hybrid feedforward lineariser can be designed using a predistorter ahead of the main amplifier thus reducing the level of IMD generated.

To minimise the effect of any distortion produced by the cancellation coefficients, α can alternatively be placed in the upper branch ahead of the main amplifier so that any additional distortion generated can be lumped with the distortion generated in the amplifier and subsequently cancelled. Since all distortion in the upper branch is treated in the same manner, no matter how it is generated, the feedforward lineariser is also able to cope with the effects of RF amplifiers with memory, thus has a clear advantage over the predistortion technique discussed.

Feedforward linearisation relies on the subtraction of nearly equal quantities and as a consequence is sensitive to gain, phase and delay mismatches. Stringent requirements are placed on the accuracy of the coefficients to achieve reasonable IMD suppression at the output. However, while cartesian feedback and predistortion have a finite limit to the amount of IMD suppression, feedforward can in theory provide complete cancellation of all intermodulation products.

1.5 Research Goals

1.5.1 Previous Feedforward Research at SFU

A gradient driven feedforward amplifier lineariser using a 5 watt, Class AB amplifier, with adaptive amplitude and phase was developed by Grant [1] which operated at 815 MHz providing 40 dB of distortion suppression over 7 MHz bandwidth.

Analysis was performed with the signal cancellation coefficient, α, in the upper branch ahead of the main amplifier, so that any distortion generated in the attenuation/phase adjuster could be treated as part of the PA output distortion and subsequently cancelled from the lineariser output.

Grant developed a novel use of DSP to avoid mixer DC offsets in the baseband correlations which would bias the gradient adaptation of the cancellation coefficients. It was also demonstrated by analysis that downconversion errors such as filter mismatches between the chains, amplitude ripple and phase distortion in the filters, frequency and phase offsets between the recovered complex envelopes, and incomplete image suppression of the sampled signal, do not bias the correlations, although they do affect speed of convergence.

The primary limitation on linearisation bandwidth was delay mismatch between the upper and lower branches in the signal and distortion cancellation branches. Grant found that the amount of distortion suppression was dependent on the bandwidth-delay mismatch product in the error cancellation circuit. An adaptive delay matching scheme was proposed by Cavers [4] and its analysis and implementation was the basis of the research presented in this thesis.

1.5.2 Project Goals

This project is an extension of the work carried out by Grant. The new research presented in this thesis will include an analysis of the ability of the delay matching design to compensate for amplitude variation with frequency across the linearisation bandwidth. A novel way to increase the speed of convergence of the adaptive coefficients will also be described. The main project goal was to build a working implementation of the adaptive feedforward lineariser with delay matching and compare the results with the original single coefficient case.

It was proposed to implement the feedforward lineariser at a centre frequency 1.8 GHz so that the desired linearisation bandwidth would be an appropriately small percentage of the transmit frequency (approximately 1~2%). A class AB amplifier operating at 1 W output power was chosen as the main amplifier.

The signal modulation was chosen to be QPSK so that the power amplifier output will have contributions from the IMD generated both in the cutoff region and at saturation, thus presenting a more pessimistic result for analysis. The signal is filtered with a root raised-cosine pulse with 35% rolloff. The desired linearisation bandwidth is 15 MHz to accomodate

PCS applications, however the implementation obtained a linearisation bandwidth of 40 MHz.

Delay matching is to be performed by adapting the coefficients using partial correlations in selected narrow subbands across the band. This will be explained in greater detail in Chapter 4.

Due to the difficulty of analysing the feedforward amplifier circuit with the signal cancellation coefficients α_0 and α_1 placed in the main branch, all analysis performed in this thesis will relate to the coefficients placed in the lower branch. Both methods of analysis exhibit very similar behaviour in the neighbourhood of the optimum. However in the circuit implementation it is preferable to place these coefficients in the main branch so that any distortion introduced can be lumped as part of the distortion in the main amplifier.

The goal of this research is to prove the concept of delay matching in the feedforward amplifier. No attempt was made to optimise the power efficiency of the lineariser.

Chapter 2

Background

2.1 Chapter Overview

The purpose of this chapter is to present the adaptive feedforward analysis that preceded the current delay matching research. This introduces the main building blocks for the single coefficient adaptation which can then be easily extended to the case with two coefficients adapting simultaneously.

2.2 Gradient Based Adaptation

Calculations of the linear gain in the main amplifier and the gradient adaptation of the coefficients are performed using linear estimation theory. Linear estimation minimises the second order mean-square error function of the difference between the signal $v(t)$ and the basis for the estimate $\widehat{v}(t)$, given as $|v(t) - \widehat{v}(t)|^2$. Minimising the metric $|v(t) - \widehat{v}(t)|^2$ is equivalent to the orthogonal projection of the vector $v(t)$ onto the basis $\widehat{v}(t)$. The quadratic error metric results in a familiar paraboloid surface with a minimum point equal to the minimum power in the error.

Linear estimation uses two equivalent criteria to minimise the error metric. The first is minimum power in the estimation error vector $v(t) - \widehat{v}(t)$ and the second criteria is zero correlation between the basis vector and the error vector, indicating orthogonality. The criteria of zero correlation is also equivalent to maximum correlation between the signal being estimated and the basis for the estimate.

The covariance of the basis and the estimation error is equivalent to the gradient of the

error surface; thus when the minimum point on the error surface is reached the gradient will become zero and the basis and estimation error will be uncorrelated.

Adaptation based on the criterion of power minimisation requires periodically misadjusting the coefficient to assess whether the minimum point is still achieved. This increases the amount of IMD at the output and leads to long convergence times due to noisy power measurements. Using the criteria of decorrelation of the error signal and the basis leads to the gradient algorithm for adaptation, which continually computes gradient signals in search of the optimum operating point of the system. Consequently, deliberate misadjustment of the coefficient is not required. For this reason, gradient based adaptation is the method chosen for this thesis.

Analysis is performed using the single coefficient feedforward amplifier that was shown in Figure 1.5.

2.2.1 Linear Gain Estimation

Estimation of the linear component of $v_a(t)$ relies on the knowledge that $v_m(t)$ and $v_d(t)$ are uncorrelated, or mutually orthogonal. The output from the main amplifier is expressed as

$$v_a(t) = v_m(t)G(x_m(t)) \tag{2.1}$$

and is also given in linear estimation terms as

$$v_a(t) = \gamma_o v_m(t) + v_d(t) \tag{2.2}$$

The criterion for the linear gain γ_o is the value which results in the greatest correlation between the amplifier output $v_a(t)$ and input signal $v_m(t)$, or zero correlation between $v_m(t)$ and the distortion component $v_d(t)$. Equivalently minimum power in $v_d(t)$ can be used. Multiplying (2.1) and (2.2) by the complex conjugate (represented with an asterisk) of $v_m(t)$ and taking the expectation results in

$$\begin{aligned} E[v_a(t)v_m^*(t)] &= E[x_m(t)G(x_m(t))] \\ &= \gamma_o P_m + E[v_d(t)v_m^*(t)] \end{aligned} \tag{2.3}$$

where E[.] denotes expectation with respect to the probability density function (pdf) of the instantaneous power of the input signal $v_m(t)$. Using the criterion of zero correlation between $v_m(t)$ and $v_d(t)$ sets $E[v_d(t)v_m^*(t)] = 0$ and by rearranging (2.3) this leads to the solution for γ_o

$$\gamma_o = \frac{1}{P_m} E[x_m(t)G(x_m(t))] \tag{2.4}$$

To show the equivalence of the two criteria the minimum power in $v_d(t)$ will now be calculated. Again employing (2.1) and (2.2) the power in $v_d(t)$ is

$$\begin{aligned} P_d &= E[x_d(t)] = E[|v_m(t)G(x_m(t)) - \gamma_o v_m(t)|^2] \\ &= P_a - 2Re[\gamma_o^* E[x_m(t)G(x_m(t))]] + |\gamma_o|^2 P_m \end{aligned} \tag{2.5}$$

which represents a quadratic surface in γ_o. Finding the minimum point on this surface is equivalent to minimising the power of P_d. Thus by differentiating with respect to γ_o and setting the result to zero, we obtain the same result shown in (2.4).

2.2.2 Adaptation of the Signal Cancellation Coefficient

The adaptation coefficient α adapts the reference signal $v_m(t)$ in the lower branch of the signal cancellation circuit to form an estimate of the desired component $\gamma_o v_m(t)$ in the upper branch. The error signal at the output of the signal cancellation circuit is

$$v_\epsilon(t) = (\gamma_o - \alpha)v_m(t) + v_d(t) \tag{2.6}$$

Clearly when $\alpha = \gamma_o$ the error signal $v_\epsilon(t)$ will be equal to the amplifier distortion $v_d(t)$. The power in $v_\epsilon(t)$ is

$$P_\epsilon(\alpha) = E[x_\epsilon(t)] = [|\gamma_o|^2 - 2Re[\gamma_o \alpha^*] + |\alpha|^2]P_m + P_d \tag{2.7}$$

since $v_d(t)$ and $v_m(t)$ are uncorrelated. Differentiating with respect to α yields the gradient of the quadratic surface

$$\frac{\partial P_\epsilon(\alpha)}{\partial \alpha} = (\gamma_o - \alpha)P_m \tag{2.8}$$

The equivalent gradient signal is derived using the covariance of $v_m(t)$ and $v_\epsilon(t)$ and is given in terms of the relative error in α, $\epsilon_\alpha = (\alpha - \alpha_{opt})/\alpha_{opt}$

$$E[v_\epsilon(t)v_m^*(t)] = -\gamma_o\epsilon_\alpha P_m \tag{2.9}$$

The method of steepest descent provides an iterative approach to finding the minimum point on the error surface corresponding to the optimum value of the coefficient $\alpha_{opt} = \gamma_o$. This involves calculating the gradient at an arbitrary point on the error surface and correcting the coefficient using small steps in the direction opposing the gradient signal. The minimum point is achieved when $v_\epsilon(t)$ and $v_m(t)$ are uncorrelated, corresponding to the zero gradient.

In a practical implementation of the steepest descent algorithm, an estimate of the gradient signal is used. This is an instantaneous value of the gradient of the mean-square error known as the stochastic gradient signal and provides a noisy, but unbiased estimate to the true gradient signal

$$D_\alpha(t) = v_\epsilon(t)v_m^*(t) \tag{2.10}$$

This leads to the stochastic gradient algorithm for adaptation given by

$$\alpha(t) = K_\alpha \int_0^t D_\alpha(\tau)d\tau \tag{2.11}$$

where K_α is the step size parameter chosen to provide a compromise between fast convergence speed and the excess mean-squared error known as jitter in the coefficient. The integrator helps to remove the self noise generated by the gradient estimate. Updates to α are made in small increments determined by K_α until the gradient estimate becomes zero on average, when $v_\epsilon(t)$ and $v_m(t)$ are uncorrelated and the optimum value of α is achieved. If there is any uncorrelated phase shift in (2.10) then the gradient estimate does not follow the true path of steepest descent, but instead follows a spiral-like path toward the optimum value on the surface. This results in slightly longer adaptation times.

2.2.3 Adaptation of the Distortion Cancellation Coefficient

The adaptation of β follows in a similar fashion. The coefficient β is used to estimate the distortion component in the output of the main amplifier. The lineariser output signal is given as

$$v_o(t) = v_a(t) - \beta v_d(t) \tag{2.12}$$

The basis of the estimate is $v_d(t)$ and an orthogonal projection of $v_o(t)$ onto the basis provides the minimum error between the output signal and its estimate. Since the error surface P_o is quadratic in β both the criteria of zero correlation between $v_o(t)$ and $v_d(t)$ or the criteria of minimum power in $v_o(t)$ can be used. Assuming α is optimised such that $v_\epsilon(t) = v_d(t)$,

$$P_o(\beta) = |\gamma_o|^2 P_m + (1 - 2Re[\beta] + |\beta|^2)P_d \tag{2.13}$$

When $\beta = \beta_{opt}$ then $v_o(t)$ and $v_d(t)$ will be uncorrelated and the lineariser output will consist only of the component $\gamma_o v_m(t)$. Differentiating the error surface with respect to β and setting the result to zero yields the optimum value for β

$$\frac{\partial P_o(\beta)}{\partial \beta} = (2\beta - 2)P_d \tag{2.14}$$

Thus $\beta_{opt} = 1$ at the minimum point on the parabolic surface.

The desired gradient signal used for the method of steepest descent is the covariance of the lineariser output $v_o(t)$ and the distortion component $v_d(t)$, and is given in terms of the relative error in β, $\epsilon_\beta = (\beta - \beta_{opt})/\beta_{opt}$

$$E[v_o(t)v_d^*(t)] = -\epsilon_\beta P_d \tag{2.15}$$

However, since the distortion signal $v_d(t)$ is not available unless the signal is completely cancelled from $v_\epsilon(t)$, then the gradient is approximated using $v_e(t)$ as the basis. Since $v_\epsilon(t)$ is a function of α, then the convergence of α and β will be coupled. Avoiding large misadjustment of β places stringent requirements on the accuracy of the coefficient α, which depends on the amount of IM reduction required at the output [2].

The stochastic gradient for the adaptation of β is

$$D_\beta(t) = v_o(t)v_e^*(t) \tag{2.16}$$

A problem arises with this gradient due to the large signal component in $v_o(t)$ (2.12), masking the distortion signal $v_d(t)$ which the adaptation is trying to estimate. This leads to excessive self noise in the gradient estimate which hinders the convergence of β caused by residual misadjustment of α. The masking problem is solved by designing a bandstop filter in DSP to filter out the desired signal component leaving the distortion signal for correlation. Not only does this result in faster convergence speeds but also reduces the effect of bias in

the β coefficient. With this approach the filtered lineariser output signal is represented by $v_o'(t)$ and the new gradient estimate becomes

$$D_\beta'(t) = v_o'(t)v_e^*(t) \tag{2.17}$$

The stochastic gradient algorithm analogous to (2.11) is

$$\beta(t) = K_\beta \int_0^t D_\beta'(\tau)d\tau \tag{2.18}$$

where K_β is the step size parameter controlling the speed of adaptation and jitter in the coefficient.

2.2.4 Analysis of Alternative Placement of α

The analysis of the feedforward lineariser can now be extended to the case with α placed ahead of the main amplifier in the upper branch as analysed in [1]. This alternative placement affects the amplifier output signal such that

$$v_a(t) = \alpha v_m(t)G(|\alpha|^2 x_m(t)) \tag{2.19}$$

Again the criterion of greatest correlation between $v_a(t)$ and $v_m(t)$ can be applied to find the linear gain γ_o of the amplifier

$$\gamma_o = \frac{1}{P_m}\alpha E[x_m(t)G(|\alpha|^2 x_m(t))] \tag{2.20}$$

In this case, however, the criterion of minimum power in $v_d(t)$ does not apply as P_d is no longer quadratic in α. Using the criteria of decorrelation of the signals $v_e(t)$ and $v_m(t)$ results in complete cancellation of the desired signal in $v_\epsilon(t)$, but the power of P_e at this optimum value is at a higher level than the derived value using the minimisation criteria. These two criteria will form equivalent optimal values of α if the input signal is backed off more than 3 dB.

For α placed in the main branch the error signal is given as

$$v_e(t) = (\gamma_o - 1)v_m(t) + v_d(t) \tag{2.21}$$

Clearly $\gamma_o = 1$ for $\alpha = \alpha_{opt}$ with $v_e(t)$ and $v_m(t)$ uncorrelated. The covariance of $v_m(t)$ and $v_e(t)$ and thus gradient of the error surface is derived using (2.20) and (2.21)

$$E[v_\epsilon(t)v_m^*(t)] = \alpha E[x_m(t)G(|\alpha|^2 x_m(t)] - P_m \tag{2.22}$$

The optimal value of α is found when the covariance is equal to zero.

The effect of the placement of α on the distortion cancellation circuit is found by substituting (2.19) with optimal values of α into (2.12) to find the output of the feedforward lineariser

$$v_o(t) = \alpha_{opt} v_m(t) G[|\alpha_{opt}|^2 x_m(t)] - \beta v_d(t) \tag{2.23}$$

The error surface is now becomes

$$P_o(\beta) = P_a(\alpha_{opt}) - 2Re\left[\beta^* \alpha_{opt} E[v_m(t) G\left(|\alpha_{opt}|^2 x_m(t)\right) v_d^*(t)]\right] + |\beta|^2 P_d \tag{2.24}$$

and the gradient of the surface with respect to β is given as

$$\frac{\partial P_o(\beta)}{\partial \beta} = -2\alpha_{opt} E[v_m(t) G\left(|\alpha_{opt}|^2 x_m(t)\right) v_d^*(t)] + 2\beta P_d \tag{2.25}$$

Setting this result to zero leads to the optimal solution for β

$$\beta_{opt} = \frac{1}{P_d} \alpha_{opt} E[v_m(t) G\left(|\alpha_{opt}|^2 x_m(t)\right) v_d^*(t)] \tag{2.26}$$

Noting that the numerator of this expression is equal to the covariance of the amplifier output and distortion signal $E[v_a(t) v_d^*(t)]$ at $\alpha = \alpha_{opt}$ then the equation reduces to

$$\beta_{opt} = \frac{1}{P_d} E[v_m(t) + v_d(t)) v_d^*(t)] = 1 \tag{2.27}$$

which is the same result previously derived for the placement of α in the lower branch.

Computing the covariance of $v_o(t)$ and $v_e(t)$ gives the approximation to the true gradient signal in (2.15) and is shown in terms of the relative errors in γ_o and β

$$E[v_o(t) v_e^*(t)] = \epsilon_{\gamma_o}^* (1 - \epsilon_{\gamma_o} \epsilon_\beta) P_m - \epsilon_\beta P_d \tag{2.28}$$

where $\epsilon_{\gamma_o} = \gamma_o - 1$, and is equal to zero when α is perfectly adjusted. This demonstrates the bias effect due to incomplete signal suppression in the error signal $v_e(t)$.

As can be seen from the above equations the analysis has become more complicated with the placement of α in the upper branch. The addition of a second coefficient ahead of the main amplifier makes the analysis intractable and therefore all subsequent analysis is performed with the signal cancellation coefficients placed in the lower branch. Although it is a different configuration, its adaptation behaviour in the neighbourhood of the convergence point is very similar to that of the circuit with α placed in the upper branch.

Chapter 3

Adaptive Delay Matching

3.1 Chapter Overview

This chapter analyses the effect of delay mismatch in the feedforward amplifier and proposes a method of providing adaptive delay matching to the original circuit. The basis for the adaptation of the coefficients is presented which adapts the coefficients independently. As will be shown, this leads to long adaptation times, and a more useful approach to adaptation will be presented in Chapter 4.

3.2 Effect of Delay Mismatch

Delay mismatch between the upper and lower branches of the feedforward signal and distortion cancellation circuits results in incomplete suppression of the respective signals thus reducing the wide band capability of the circuit. In the original single coefficient case shown in Figure 3.1 the delay mismatch between the branches in the signal and distortion cancellation circuits is represented by τ_α and τ_β respectively.

It was shown in [1] that for optimum values of the coefficients α and β ($\alpha_{opt} = \gamma_o$, $\beta_{opt} = 1$), the feedforward lineariser output can be expressed in the following form[1]

$$V_o(f) \approx V_m(f) - j2\pi f \tau_\beta V_d(f) \tag{3.1}$$

implying that the delay mismatch in the distortion cancellation branch is the most critical

[1]This approximation holds for small bandwidth-delay mismatch product, $f\tau_\alpha$ and $f\tau_\beta << 1$

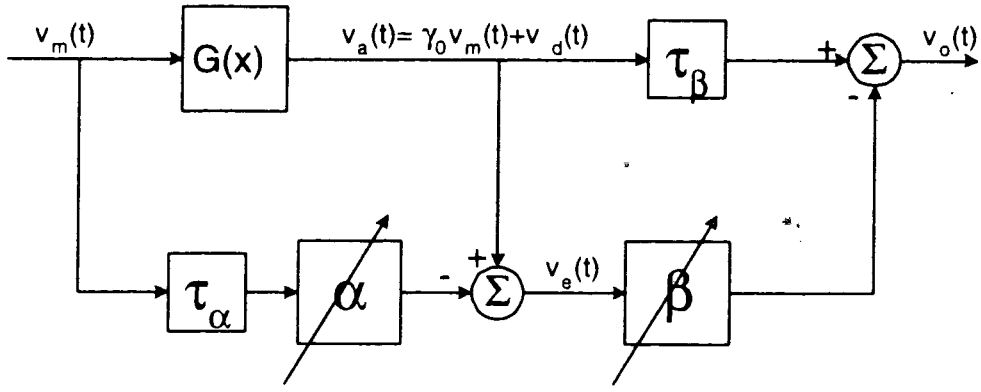

Figure 3.1: Complex baseband model of single coefficient case with delay mismatch

factor in the suppression of the intermodulation products. To achieve a linearisation bandwidth of 30 MHz with a minimum of 30 dB distortion suppression, the product $2\pi f\tau_\beta$ must not exceed 0.032 across the band of distortion. This places a stringent requirement on the maximum value of the delay mismatch in the distortion cancellation loop, where τ_β must be less than 0.3 ns.

3.2.1 Misadjustment of the Signal Cancellation Coefficient

Delay mismatch between the branches in the lineariser circuit also causes misadjustment of the coefficients. An analysis of the effect of delay mismatch in the signal cancellation branch on the coefficient α is given in [2] and the result for the relative error in α is reproduced here

$$\epsilon_\alpha \approx -\frac{\pi^2}{6}(W_m\tau_\alpha) \tag{3.2}$$

where W_m is the bandwidth of the input signal. Since convergence of α and β are coupled, any misadjustment in α will bias the convergence of β in shown in (2.28). High accuracy constraints are placed on α to avoid significant misadjustment in β and to keep the ratio of output power to signal power, $IMSR_o$, low at the lineariser output. To achieve a coefficient accuracy of 10^{-4} the delay mismatch must be be less than 0.8% of the reciprocal bandwidth [2].

3.2.2 Misadjustment of the Distortion Cancellation Coefficient

The distortion cancellation coefficient, β, is affected by both the misadjustment in α and the delay mismatch, τ_β, in the distortion cancellation circuit. An analysis of the effect of the delay mismatch is given here which assumes negligible delay mismatch in the signal cancellation circuit. With reference to Figure 3.1, the output of the feedforward amplifier is given as

$$v_o(t) = \gamma_o v_m(t - \tau_\beta) + v_d(t - \tau_\beta) - \beta v_\epsilon(t) \tag{3.3}$$

For α adjusted to its optimum value, $v_\epsilon(t) = v_d(t)$, and the mean gradient signal $E[D_\beta(t)]$ can be calculated as

$$\begin{aligned} E[D_\beta(t)] &= E[v_o(t)v_d^*(t)] \\ &= R_d(-\tau_\beta) - \beta R_d(0) \end{aligned} \tag{3.4}$$

where R_d is the autocorrelation function of the distortion signal. Ideally if $\tau_\beta = 0$, then $E[D_\beta(t)] = 0$ when $\beta = \beta opt = 1$. However, when $\tau_\beta \neq 0$, β will converge to the following value

$$\beta = \frac{R_d(-\tau_\beta)}{R_d(0)} \leq 1 \tag{3.5}$$

Approximating the distortion power spectral density $S_d(f)$ as a flat rectangular function across the third order intermodulation bandwidth such that $S_d(f) = rect_{3W_m}(f)$, then the autocorrelation function $R_d(\tau)$ can be found by taking the inverse Fourier transform of $S_d(f)$.

$$R_d(\tau) \approx sinc(3W_m\tau) \tag{3.6}$$

For small values of τ, the *sinc* function can be approximated to

$$sinc(x) \approx 1 - \frac{(\pi x)^2}{6} \tag{3.7}$$

This leads to the solution for $R_d(-\tau_\beta)$

$$R_d(-\tau_\beta) = 1 - \frac{3}{2}\pi^2(W_m\tau_\beta)^2 \tag{3.8}$$

The relative error in β is represented by ε_β in (3.9), and is found by substituting the results from (3.8) into (3.5).

$$\begin{aligned} \varepsilon_\beta &= \frac{(\beta - \beta_{opt})}{\beta_{opt}} \\ &= -\frac{3}{2}\pi^2(W_m\tau_\beta)^2 \end{aligned} \tag{3.9}$$

This misadjustment error in β represents one contribution to the amount of residual distortion remaining in the output, $v_o(t)$, of the feedforward amplifier. The total distortion power in the output signal is the sum of the squared magnitudes of ε_β and $j2\pi f\tau_\beta$ (from (3.1)) ,

$$P_d = (2\pi f\tau_\beta)^2 + \frac{9}{4}\pi^4(W_m\tau_\beta)^4 \tag{3.10}$$

A plot of the distortion power P_d in the output signal for various values of bandwidth-delay mismatch product $W_m\tau_\beta$ is given in Figure 3.2. The delay mismatch τ_β is set at 0.3 ns so that the graphs for 3, 10 and 30 MHz channel bandwidths refer to a bandwidth-delay mismatch product of 0.09%, 0.3% and 0.9% respectively. The frequency is normalised to $f = \frac{1}{2}(3W_m)$ to include the third order IM products. Clearly to achieve a minimum of 40 dB reduction in IMD products at the output of the feedforward amplifier the product $W_m\tau_\beta$ must be held within 0.09%. These stringent requirements emphasise the need for adaptive delay matching techniques.

3.3 Delay Matching Circuit

The proposed delay matching circuitry [4] for each cancellation circuit in the feedforward amplifier can be described as a two tap equalizer which adapts two complex coefficients based on linear estimation of the signal in the upper branch. The complex baseband model of the adaptive feedforward amplifier with delay matching circuitry is shown in Figure 3.3. Note that in the implementation the α coefficients are placed in the upper branch ahead of the main amplifier. This model again assumes that no attenuation is provided by the sampling couplers or the fixed attenuation and the auxilliary amplifier is modelled as linear with unity gain. The circuit components, however, do not exhibit a flat response across the band, but this variation will later be lumped as ripple in the adaptation coefficients and will vary linearly with frequency.

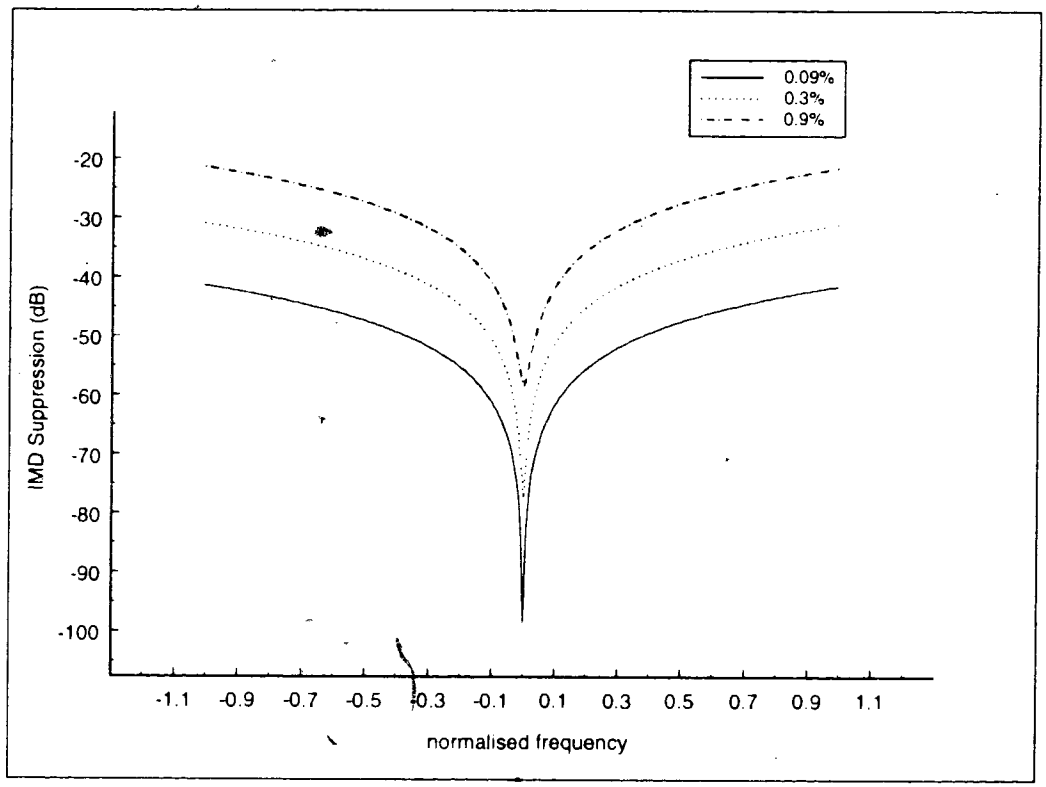

Figure 3.2: IMD power at feedforward output for different bandwidth-delay mismatch products $W_m\tau_g$.

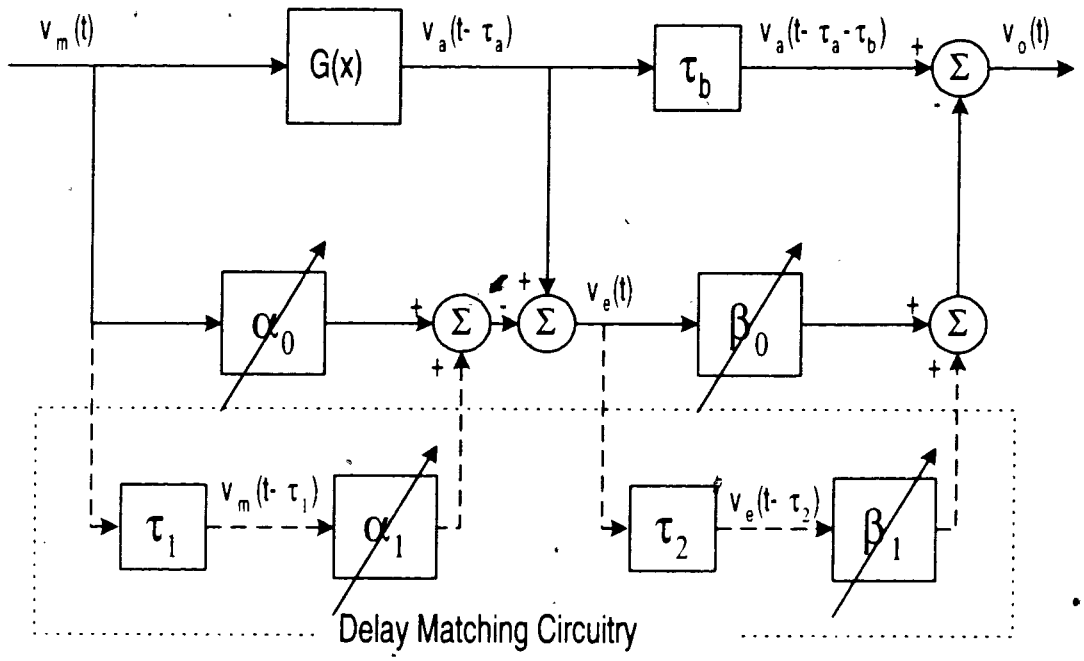

Figure 3.3: Complex baseband model of adaptive feedforward amplifier with delay matching

For the signal cancellation circuit, the relative group delay τ_a through the main amplifier (modelled by $G(x)$) is measured relative to the α_0 adaptation branch represented with zero group delay. The second signal cancellation adaptation coefficient α_1 provides an additional fixed delay line of delay τ_1 chosen such that the delay in the main amplifier lies within the two coefficient taps, $0 \leq \tau_a \leq \tau_1$. The outputs of the signal cancellation coefficients branches are summed together then subtracted from the output of the main amplifier to derive the output to the signal cancellation circuit. The coefficients adaptively adjust their attenuation and phase so that the sum of their outputs provides the best estimate of the linear component $\gamma_o v_m(t-\tau_a)$ in the upper branch. This process can also be described as an estimation of the signal in the upper branch by interpolation of the outputs of the two adaptation branches, α_0 and α_1.

Similarly for the distortion cancellation circuit, the delay in the upper branch τ_b is represented relative to the delay through the adaptation coefficient β_0 branch. In this case, however, it is a fixed delay line which falls within the two taps provided by the distortion cancellation coefficients β_0 and β_1. The fixed delay line τ_2 added to the β_1 path is chosen such that $0 \leq \tau_b \leq \tau_2$. The coefficients make attenuation and phase adjustments so that

the sum of their outputs produces a linear estimate of the distortion component in the delayed upper branch. Subtracting the estimated distortion component from the delayed main amplifier output results in the linearised feedforward amplifier output $v_o(t)$.

Power amplifiers exhibit varying delay due to frequency over wide bandwidths. The taps must therefore be designed so that they incorporate this delay variation across the band. Ideally the smaller the delay between the taps, the more accurate the coefficients will estimate the signal and better cancellation of the relevant signals will be achieved.

3.4 Adaptation of the Complex Coefficients

3.4.1 Adaptation in the Signal Cancellation Circuit

Analysis of the signal cancellation circuit is made with reference to Figure 3.3. As discussed previously, the output of the main amplifier is expressed in linear estimation terms as the sum of a linearly amplified component and the intermodulation distortion, with a relative group delay of τ_a. The signal cancellation coefficients α_0 and α_1 adjust the attenuation and phase of the input signal $v_m(t)$and its delayed version $v_m(t-\tau_1)$ respectively. The output from the subtracter in the signal cancellation circuit, $v_e(t)$, representing the error signal, is expressed as

$$\begin{aligned} v_e(t) &= v_a(t) - \alpha_0 v_m(t) - \alpha_1 v_m(t-\tau_1) \\ &= \gamma_o v_m(t-\tau_a) + v_d(t-\tau_a) - \alpha_0 v_m(t) - \alpha_1 v_m(t-\tau_1) \end{aligned} \tag{3.11}$$

Complete signal cancellation is achieved when $v_e(t) = v_d(t-\tau_a)$, which is the case when

$$\gamma_o v_m(t-\tau_a) = \alpha_{0opt} v_m(t) + \alpha_{1opt} v_m(t-\tau_1) \tag{3.12}$$

For small values of τ_a and τ_1, it can be stated that $\gamma_o \approx \alpha_{0opt} + \alpha_{1opt}$.

The criterion used to obtain the optimum values for α_0 and α_1 is zero correlation between the error signal $v_e(t)$ and the input signal $v_m(t)$. This occurs when $v_e(t)$ is comprised only of amplifier intermodulation distortion products, $v_d(t-\tau_a)$. Similar to the configuration presented in Chapter 2, the covariance of the basis $v_m(t)$ and the error signal $v_e(t)$ is used as the gradient signal to drive the adaptation of α_0 and, as shown in (3.13), this is expressed in terms of the average power and autocorrelation functions of the reference signal $v_m(t)$.

Similarly, the covariance of the basis $v_m(t-\tau_1)$ and the error signal $v_e(t)$ is used as the gradient signal to drive the adaptation of α_1 and is given in (3.14). When the correlation of $v_e(t)$ with the two basis signals $v_m(t)$ and $v_m(t-\tau_1)$ is zero, the optimum values for α_0 and α_1 have been reached.

$$E\left[v_\epsilon(t)v_m^*(t)\right] = \gamma_o R_m(-\tau_a) - \alpha_0 P_m - \alpha_1 R_m(-\tau_1) \tag{3.13}$$

$$E\left[v_\epsilon(t)v_m^*(t-\tau_1)\right] = \gamma_o R_m(\tau - \tau_a) - \alpha_1 P_m - \alpha_0 R_m(\tau_1) \tag{3.14}$$

An equivalent criterion for optimising the coefficients is to find the minimum power[2] in the error signal $v_\epsilon(t)$. $P_\epsilon(\alpha_0,\alpha_1)$ is the average power of $v_e(t)$ and is a quadratic function of both coefficients α_0 and α_1. If it were possible to plot this error surface it would have a minimum when $\alpha_0 = \alpha_{0opt}$ and $\alpha_1 = \alpha_{1opt}$.

$$\begin{aligned} P_\epsilon(\alpha_0,\alpha_1) &= E[|v_\epsilon(t)|]^2 \\ &= [|\gamma_o|^2 + |\alpha_0|^2 + |\alpha_1|^2]P_m - 2Re[\alpha_0^*\gamma_o R_m(-\tau_a)] \\ &\quad -2Re[\alpha_1^*\gamma_o R_m(\tau_1 - \tau_a)] - 2Re[\alpha_0^*\alpha_1 R_m(-\tau_1)] \end{aligned} \tag{3.15}$$

The fact that the two criteria are equivalent can be shown by finding the gradient of the error surface $P_\epsilon(\alpha_0,\alpha_1)$ with respect to the coefficients and comparing the result with (3.13) and (3.14)

$$\begin{aligned} \frac{\partial P_\epsilon(\alpha_0,\alpha_1)}{\partial \alpha_0} &= -2\gamma_o R_m(-\tau_a) + 2\alpha_0 P_m + 2\alpha_1 R_m(-\tau_1) \\ &= -2E\left[v_\epsilon(t)v_m^*(t)\right] \end{aligned} \tag{3.16}$$

$$\begin{aligned} \frac{\partial P_\epsilon(\alpha_0,\alpha_1)}{\partial \alpha_1} &= -2\gamma_o R_m(\tau_1 - \tau_a) + 2\alpha_1 P_m + 2\alpha_0 R_m(\tau_1) \\ &= -2E\left[v_\epsilon(t)v_m^*(t-\tau_1)\right] \end{aligned} \tag{3.17}$$

The iterative adaptation of the coefficients is based on the method of steepest descent. This procedure involves taking an arbitrary point on the error surface defined by initial

[2]The minimum power criterion applies to quadratic error surfaces and does not apply if the coefficients are placed in the main branch [1]

values of α_0 and α_1, then calculating the gradients for each coefficient. The corresponding coefficient is then incremented by an amount opposing the direction of the computed gradient until the optimum value is reached. However, it is evident that any changes made to one coefficient will affect the gradient adaptation of the other coefficient, causing large changes to the position on the error surface and resulting in a much altered path to the true path of steepest descent. This strong dependence will lead to long convergence times. A solution to this problem is proposed in Section 4.5.

As we saw in Chapter 2, it is not possible to calculate the true gradient shown in the equations above in a practical system. Instead, a stochastic gradient signal is used which provides a noisy but unbiased estimate of the true gradient of the quadratic error surface.

$$D_{\alpha_0}(t) = v_e(t)v_m^*(t) \tag{3.18}$$

$$D_{\alpha_1}(t) = v_e(t)v_m^*(t-\tau_1) \tag{3.19}$$

The algorithm for adjusting the coefficients α_0 and α_1 is a simple first order adaptation loop, where the integrator helps to remove some of the self noise in the gradient estimate and the step size parameters, K_{α_0} and K_{α_1}, control the speed of adaptation.

$$\alpha_0(t) = K_{\alpha_0}\int_0^t D_{\alpha_0}(\tau)d\tau \tag{3.20}$$

$$\alpha_1(t) = K_{\alpha_1}\int_0^t D_{\alpha_1}(\tau)d\tau \tag{3.21}$$

When the stochastic gradients are reduced to zero on average, the input and error signals will be decorrelated and the optimum values of the signal cancellation coefficients will have been reached. The values of $\alpha_0(t)$ and $\alpha_1(t)$ will be held at their final integrated values.

3.4.2 Adaptation in the Distortion Cancellation Circuit

The distortion cancellation circuit is described with reference again to Figure 3.3. The two adaptation coefficients β_0 and β_1 are used to adjust the attenuation and phase of the error signal $v_e(t)$ and its delayed version $v_e(t-\tau_2)$ respectively in such a way as to best estimate the IMD at the main amplifier output. Assuming a linear unity gain error amplifier with no other distortion introduced, the feedforward amplifier output $v_o(t)$ can be expressed by

$$\begin{aligned} v_o(t) &= v_a(t-\tau_a-\tau_b) - \beta_0 v_e(t) - \beta_1 v_e(t-\tau_2) \\ &= \gamma_o v_m(t-\tau_a-\tau_b) + v_d(t-\tau_a-\tau_b) - \beta_0 v_e(t) - \beta_1 v_e(t-\tau_2) \end{aligned} \tag{3.22}$$

When α_0 and α_1 are optimised as given in (3.12), $v_e(t) \approx v_d(t-\tau_a)$. Complete distortion cancellation is then achieved when $v_o(t) = \gamma_o v_m(t-\tau_a-\tau_b)$, which is true for

$$v_d(t-\tau_a-\tau_b) = \beta_{0opt} v_d(t-\tau_a) + \beta_{1opt} v_d(t-\tau_2-\tau_a) \tag{3.23}$$

For small values of the delays τ_b and τ_2, it can be stated that $\beta_{0opt} + \beta_{1opt} \approx 1$.

Optimum values for β_0 and β_1 are obtained by adapting the coefficients until the feedforward amplifier output signal $v_o(t)$ and the error signal $v_e(t)$ are uncorrelated. This occurs when $v_o(t)$ is comprised only of the linearly amplified input signal, $\gamma_o v_m(t-\tau_a-\tau_b)$, when the signal cancellation coefficients are at their optimal values. The covariance of the error signal $v_e(t)$ and the feedforward output signal $v_o(t)$ as shown in (3.24), provides an appropriate gradient signal to drive the adaptation of β_0:

$$E\left[v_o(t)v_e^*(t)\right] = R_d(-\tau_b) - \beta_0 P_d - \beta_1 R_d(-\tau_2) \tag{3.24}$$

Similarly, the covariance of the error signal $v_e(t-\tau_1)$ and the output signal $v_o(t)$ is used as the gradient signal to drive the adaptation of β_1.

$$E\left[v_o(t)v_e^*(t-\tau_1)\right] = R_d(\tau_2-\tau_b) - \beta_1 P_d - \alpha_0 R_d(\tau_2) \tag{3.25}$$

When these gradient values are equal to zero then the two signals have been succesfully decorrelated and the optimum values for α_0 and α_1 have been reached.

The average power in the output signal is quadratic in β_0 and β_1

$$\begin{aligned} P_o(\beta_0,\beta_1) &= E[|v_o(t)|]^2 \\ &= |\gamma_o|^2 P_m + [|\beta_0|^2 + |\beta_1|^2 + 1]P_d - 2Re[\beta_0^* R_d(-\tau_b)] \\ &\quad -2Re[\beta_1^* R_d(\tau_2-\tau_b)] + 2Re[\beta_0\beta_1^* R_d(\tau_2)] \end{aligned} \tag{3.26}$$

Minimising the power in the output $v_o(t)$ by adjusting the values of β_0 and β_1 is equivalent to decorrelating the signals as described above. For α_0 and α_1 adjusted to their optimum values

the gradients of the output surface $P_o(\beta_0, \beta_1)$ can be found with respect to the individual coefficients and are expressed in the following form

$$\begin{aligned} \frac{\partial P_o(\beta_0,\beta_1)}{\partial \beta_0} &= -2R_d(-\tau_b) + 2\beta_0 P_d + 2\beta_1 R_d(-\tau_2) \\ &= -2E\left[v_o(t)v_\epsilon^*(t)\right] \end{aligned} \tag{3.27}$$

$$\begin{aligned} \frac{\partial P_o(\beta_0,\beta_1)}{\partial \beta_1} &= -2R_d(\tau_2 - \tau_b) + 2\beta_1 P_d + 2\beta_0 R_d(\tau_2) \\ &= -2E\left[v_o(t)v_\epsilon^*(t-\tau_2)\right] \end{aligned} \tag{3.28}$$

Analogous to (3.18) and (3.19), the stochastic gradient signals used to provide an estimate of the true gradient of the quadratic error surface for β_0 and β_1 are

$$D_{\beta_0}(t) = v_o(t)v_\epsilon^*(t) \tag{3.29}$$

$$D_{\beta_1}(t) = v_o(t)v_\epsilon^*(t-\tau_2) \tag{3.30}$$

Consequently the adaptation algorithms for adjusting the coefficients β_0 and β_1 with the corresponding step size parameters, K_{β_0} and K_{β_1}, are given as

$$\beta_0(t) = K_{\beta_0}\int_0^t D_{\beta_0}(\tau)d\tau \tag{3.31}$$

$$\beta_1(t) = K_{\beta_1}\int_0^t D_{\beta_1}(\tau)d\tau \tag{3.32}$$

When the coefficients are at their optimal values $\beta_0 = \beta_{0opt}$ and $\beta_1 = \beta_{1opt}$, then the stochastic gradients will be zero on average and the coefficients will be held at their final integrated state. Changes in operating conditions will cause the coefficients to readapt to new optimum values.

3.4.3 Biasing and Masking

The above analysis showed how the coefficients adapt when the signal cancellation coefficients α_0 and α_1 are completely converged and the basis signal is completely cancelled from

the signal $v_e(t)$. However, misadjustment of the coefficients α_0 and α_1 will bias the adaptation of the distortion cancellation coefficients. An analysis of the accuracy requirements on the signal cancellation coefficient is given in [2].

The problem of long convergence times caused by the strong signal component in the output $v_o(t)$ masking the weak distortion component in the correlation was resolved in [1] by filtering the desired signal component from the β adaptation path. This is not necessary in this application as the adaptation over the wide bandwidth is achieved by taking partial correlations over small subbands which can be selected to avoid the bands where the signal is strong as described in [4].

Chapter 4

Delay Matching Analysis

4.1 Chapter Overview

This chapter presents the new analysis that was performed as part of the current feedforward research. Results showing the delay matching circuit's additional ability to handle linear variations with frequency across the band of interest are presented. The wide band adaptation method is discussed followed by an analysis of the effect of aliasing due to incomplete suppression of the subbands outside the Nyquist frequency of the DSP. The chapter is concluded with a detailed analysis of a novel approach to speed the convergence of the adaptation coefficients.

4.2 Effect of Frequency Dependence in the Vector Modulators

Amplitude variation with frequency in the circuit components results in coefficients which are not optimised across the band of interest and therefore reduces the circuits ability to suppress unwanted distortion at wide bandwidths. The original single coefficient feedforward amplifier [1] required that the accuracy in the coefficient β must be less than 1% to ensure 40 dB distortion cancellation. Any ripple in the distortion cancellation circuit components must therefore be held to within 0.1 dB. These restrictions do not apply to the signal cancellation coefficient α if it is placed ahead of the main amplifier since any distortion introduced in the main branch is cancelled. However, any additional distortion generated by the α vector

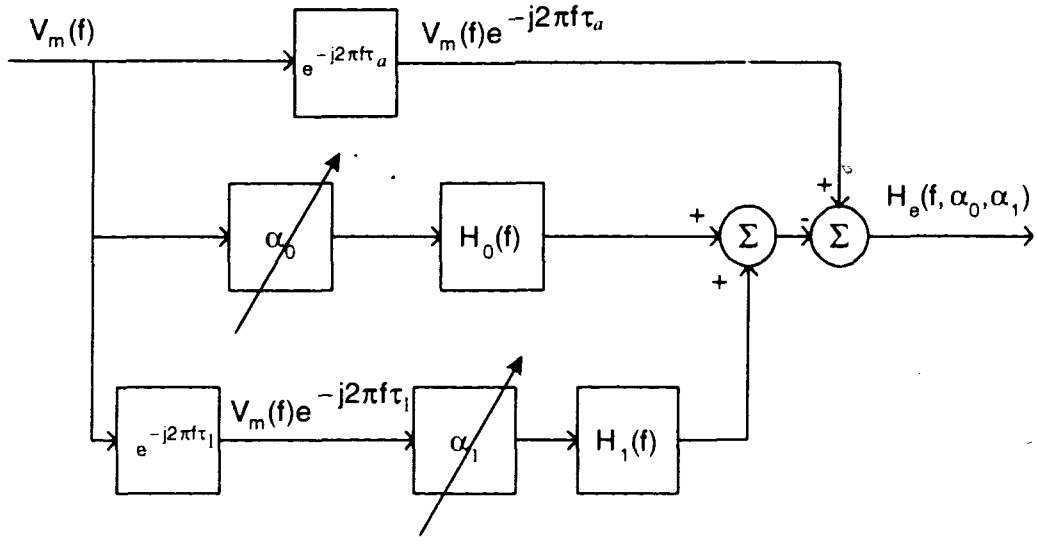

Figure 4.1: Complex baseband model of frequency response in signal cancellation circuit

modulators increases the power level that the error amplifier must handle.

Since the current design is operating at much wider bandwidths than the original system it was important to measure the variations in the components and analyse their effect on the convergence behaviour of the coefficients in the delay matching circuitry.

Measurements were made on the vector modulators, used for the attenuation and phase adjustments, to find the amplitude variation with respect to frequency across the bandwidth of interest. They were found to have a linear variation of 0.2 dB across 30 MHz (1.785-1.815 GHz), but outside this band the variation was nonlinear and the ripple increased to 0.3 dB.

An analysis was performed to determine the effect of the linear variation on the suppression of the signals in each loop. It is sufficient to analyse the behaviour of the signal cancellation loop as these results can clearly be related to the case of the distortion cancellation circuit.

The signal cancellation circuit is modelled in the frequency domain as shown in Figure 4.1. The ripple introduced by the vector modulators in the α_0 and α_1 branches is represented by the transfer functions $H_0(f)$ and $H_1(f)$ respectively, where $H_0(f) = 1 + g_0 f$ and $H_1(f) = 1 + g_1 f$, and g_0 and g_1 and complex valued coefficients representing approximately linear variation in amplitude. The signal suppression at the output of the circuit is measured by

the error transfer function

$$H_\epsilon(f, \alpha_0, \alpha_1) = \gamma_o e^{-j2\pi f \tau_a} - \alpha_0 H_0(f) - \alpha_1 H_1(f) e^{-j2\pi f \tau_1} \tag{4.1}$$

The output of the signal cancellation circuit is then represented by

$$V_\epsilon(f) = V_m(f) H_\epsilon(f, \alpha_0, \alpha_1) + V_d(f) e^{-j2\pi f \tau_a} \tag{4.2}$$

If $H_\epsilon(f, \alpha_0, \alpha_1) = 0$, then the desired output $V_\epsilon(f) = V_d(f) e^{-j2\pi f \tau_a}$ is obtained. Therefore the goal is to minimise $H_\epsilon(f, \alpha_0, \alpha_1)$ to achieve

$$\gamma_o e^{-j2\pi f \tau_a} \approx \alpha_0 H_0(f) + \alpha_1 H_1(f) e^{-j2\pi f \tau_1} \tag{4.3}$$

From (4.3), it is clear that ripple in the vector modulators controlling the adaptation of the coefficients may cause undesirable effects in the estimation of the linear component of the power amplifier. This effect will now be modelled and analysed.

Representing $|V_m(f)|^2$ with a uniform spectrum normalised to unit power over the band $2B$, the average energy contribution from the residual component $V_m(f) H_\epsilon(f)$ at the signal cancellation circuit output is given as

$$E = \frac{1}{2B} \int_{-B}^{B} |H_\epsilon(f, \alpha_0, \alpha_1)|^2 df \tag{4.4}$$

Optimal values for α_0 and α_1 can be found by differentiating E in (4.4) with respect to α_0 and α_1 respectively and equating the results to zero to obtain two equations in two unknowns. It is then "easy" to solve for optimum values of α_0 and α_1.

$$\frac{\partial E}{\partial \alpha_0} = \frac{1}{2B} \int_{-B}^{B} \frac{\partial H_\epsilon(f, \alpha_0, \alpha_1)}{\partial \alpha_0} H_\epsilon^*(f, \alpha_0, \alpha_1) df \tag{4.5}$$

$$\frac{\partial E}{\partial \alpha_1} = \frac{1}{2B} \int_{-B}^{B} \frac{\partial H_\epsilon(f, \alpha_0, \alpha_1)}{\partial \alpha_1} H_\epsilon^*(f, \alpha_0, \alpha_1) df \tag{4.6}$$

4.3 Delay Compensation

4.3.1 Vector Modulators without Frequency Dependence

For the case without frequency variation in the passband of the vector modulators, $H_0(f)$ and $H_1(f)$ are set to 1, and the optimal values for α_0 and α_1 are derived using (4.5) and (4.6) and are given by

$$\alpha_{0opt} = \frac{\gamma_o[sinc(2B\tau_a) - sinc(2B\tau_1)sinc(2B(\tau_a - \tau_1))]}{1 - sinc^2(2B\tau_1)} \tag{4.7}$$

$$\alpha_{1opt} = \gamma_o sinc[2B(\tau_a - \tau_1)] - \alpha_0 sinc(2B\tau_1) \tag{4.8}$$

where $sinc(x)$ is defined in (4.11).

The optimum values of the adaptation coefficients are obtained by setting the values of γ_o, τ_a, and τ_1. A $\pi/4$ DQPSK input signal with 35% root raised cosine filtering and input backoff of 6 dB from unity power is used as the input signal to the simulation of the signal cancellation circuit. $G(x)$ is represented by fourth order gain and phase polynomials which are fitted to the corresponding curves over the range of input powers. The linear gain is calculated using (2.4) and results in a value of $\gamma_o = 0.947 + 0.635j$. Using this value and close approximations to the delays within the actual circuit implementation, $\tau_a = 0.3$ ns and $\tau_1 = 1$ ns, a plot of $|H_e(f, \alpha_{0opt}, \alpha_{1opt})|^2$ is made as given in Figure 4.2. The optimum values of α_0 and α_1 are calculated to be $\alpha_{0opt} = 0.663 + 0.445j$ and $\alpha_{1opt} = 0.284 + 0.191j$ using (4.7) and (4.8) above. The wide bandwidth nature of the circuit is shown by the two nulls that are present compared with only one in the original circuit (see Figure 3.2). Notice the sum of the coefficients is equal to the value of linear gain γ_o.

4.3.2 Vector Modulators with Frequency Dependence

Taking into account the frequency dependence introduced by the transfer functions $H_0(f)$ and $H_1(f)$ in the α_0 and α_1 paths, new optimal values for α_0 and α_1 can be calculated by substituting (4.1) into (4.5) and (4.6) and equating the result to zero to obtain two equations in two unknowns.

$$\alpha_0 = \gamma_o \cdot \frac{\begin{bmatrix} [sinc(2B\tau_a) - \frac{g_0^*}{j2\pi}2Bsinc'(2B\tau_a)].[1 + \frac{B^2|g_1|^2}{3}] - \\ [sinc(2B\tau_1) - (\frac{g_0^* + g_1}{j2\pi})2Bsinc'(2B\tau_1) - \frac{g_0^* g_1}{(2\pi)^2}(2B)^2 sinc''(2B\tau_1)]. \\ [sinc[2B(\tau_a - \tau_1)] - \frac{g_1^*}{j2\pi}(2B)sinc'[2B(\tau_a - \tau_1)]] \end{bmatrix}}{\begin{bmatrix} [1 + \frac{B^2|g_0|^2}{3}].[1 + \frac{B^2|g_1|^2}{3}] - \\ [sinc(2B\tau_1) - (\frac{g_0^* + g_1}{j2\pi})2Bsinc'(2B\tau_1) - \frac{g_0^* g_1}{(2\pi)^2}(2B)^2 sinc''(2B\tau_1)]. \\ [sinc(2B\tau_1) + (\frac{g_0 + g_1^*}{j2\pi})2Bsinc'(2B\tau_1) - \frac{g_0 g_1^*}{(2\pi)^2}(2B)^2 sinc''(2B\tau_1)] \end{bmatrix}} \tag{4.9}$$

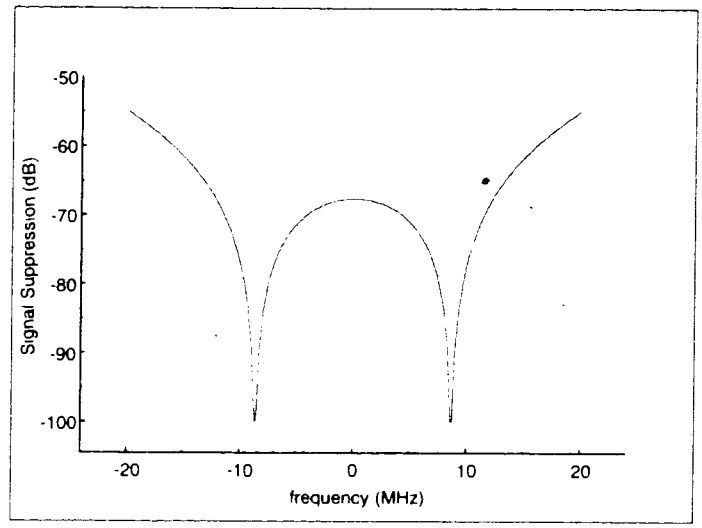

Figure 4.2: Suppression of $V_m(f)$ over 30 MHz with no amplitude variation in the passband

$$\alpha_1 = \frac{\begin{bmatrix} \gamma_o sinc[2B(\tau_a - \tau_1)] - \frac{g_1^*}{j2\pi}(2B)sinc'[2B(\tau_a - \tau_1)] \\ -\alpha_0[sinc(2B\tau_1) + (\frac{g_0 + g_1^*}{j2\pi})2Bsinc'(2B\tau_1) - \frac{g_0 g_1^*}{(2\pi)^2}(2B)^2 sinc''(2B\tau_1)] \end{bmatrix}}{[1 + \frac{B^2|g_1|^2}{3}]} \tag{4.10}$$

where

$$sinc(x) = \begin{cases} 1, & x = 0 \\ \frac{sin(\pi x)}{\pi x}, & otherwise \end{cases} \tag{4.11}$$

$$sinc'(x) = \begin{cases} 0, & x = 0 \\ \frac{cos(\pi x) - sinc(\pi x)}{x}, & otherwise \end{cases} \tag{4.12}$$

$$sinc''(x) = \begin{cases} -\pi/3, & x = 0 \\ -(\frac{\pi sin(\pi x)}{x} + \frac{2cos(\pi x) - 2sinc(\pi x)}{x^2}), & otherwise \end{cases} \tag{4.13}$$

In both cases, when α_0 and α_1 are at their optimum values, their sum is equal to the linear gain of the amplifier γ_o.

$$\gamma_o = \alpha_{0opt} + \alpha_{1opt} \tag{4.14}$$

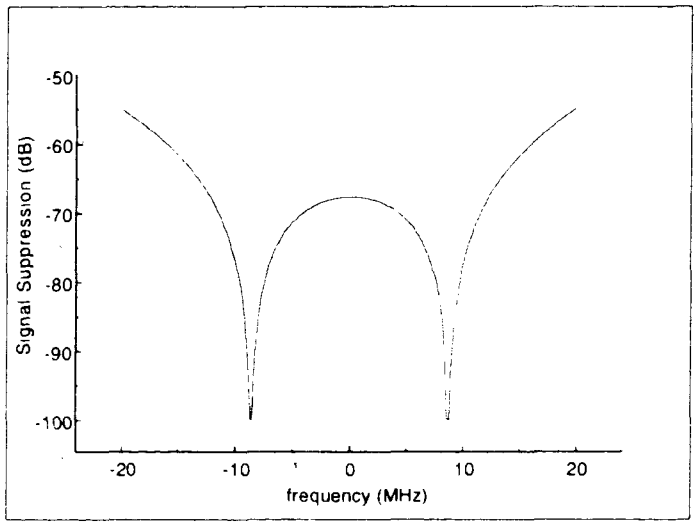

Figure 4.3: Suppression of $V_m(f)$ for delay matching circuitry with 0.2 dB linear variation in the vector modulators

Figure 4.3 shows the effect of the amplitude variation in both vector modulators on the suppression of the signal $V_m(f)$. The coefficients representing the amplitude variation in the passband are modelled as real for simplicity and are set to $g_0 = g_1 = 7.72 \times 10^{-4}$ to achieve ± 0.1 dB variation at ± 15 MHz away from centre frequency. Again the parameters $\tau_a = 0.3$ ns, $\tau_1 = 1$ ns and $\gamma_o = 0.947 + 0.635j$ are used along with g_0 and g_1 to calculate the optimal values for α_0 and α_1. A plot of the signal suppression $|H_e(f, \alpha_{0opt}, \alpha_{1opt})|^2$ can then be obtained. Comparing the result with that in Figure 4.2 shows no noticeable difference between the two plots. This demonstrates the ability of the delay matching circuitry to adapt to linear variations of amplitude ripple with frequency across the bandwidth of interest. The coefficients compensate for the amplitude variations by converging to new optimal values of $\alpha_{0opt} = 0.585 + 0.561j$ and $\alpha_{1opt} = 0.362 + 0.074j$.

4.3.3 Comparison with Single Coefficient Case

Using the same technique, a comparison can be made between the delay matching circuitry and that of the original circuit without delay compensation. The error transfer function for the single coefficient case with no delay error is given by

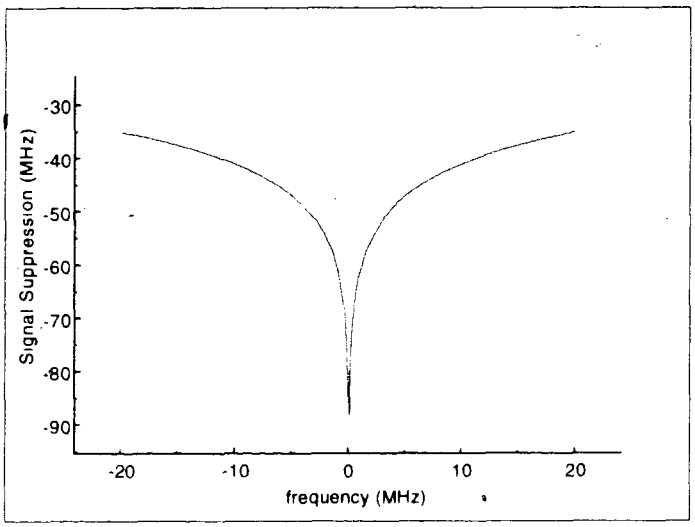

Figure 4.4: Suppression of $V_m(f)$ for single coefficient case with 0.2 dB amplitude variation and no delay mismatch

$$H_e(f,\alpha) = \gamma_o - \alpha H_0(f) \tag{4.15}$$

and from the gradient equation

$$\alpha = \frac{\gamma_o}{1 + \frac{|g_0|^2 B^2}{3}} \tag{4.16}$$

The amplitude variation has very little effect on the coefficient α, due to the product $(g_o B)^2 \ll 1$, thus $\alpha = \gamma_o$. A plot of the error transfer function $H_e(f,\alpha)$ is shown in Figure 4.4. Note that this is an optimistic result as it does not take into account the effect of delay mismatch in the loop.

Introducing the delay mismatch into the analysis the optimal value of α as derived from (4.17) can be found using (4.18)

$$H_e(f,\alpha) = \gamma_o e^{-j2\pi f \tau_a} - \alpha H_0(f) \tag{4.17}$$

$$\alpha = \frac{\gamma_o[1 + \frac{g_0'}{j2\pi} 2B sinc'(2B\tau_a)]}{1 + \frac{|g_0|^2 B^2}{3}} \tag{4.18}$$

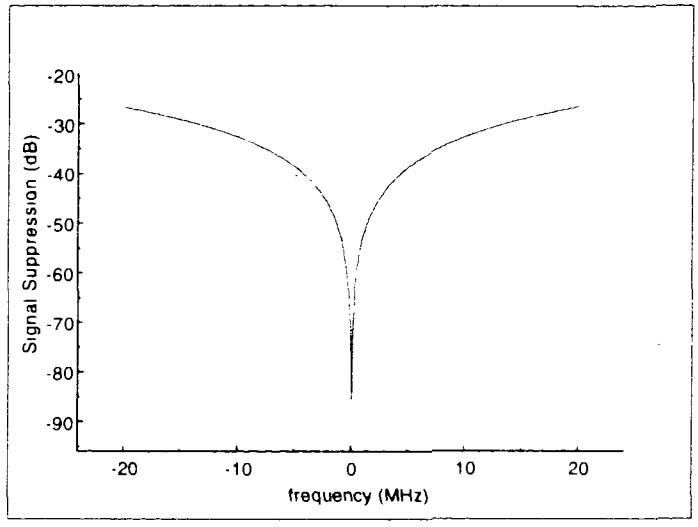

Figure 4.5: Suppression of $V_m(f)$ for single coefficient case with 0.2 dB amplitude variation and delay mismatch of 0.3 ns

Again $\alpha = \gamma_o$, but the delay mismatch of $\tau_\alpha = 0.3$ ns will result in significant misadjustment from the optimal value as the frequency moves further away from the centre band. The plot of $|H_\epsilon(f, \alpha)|^2$ in Figure 4.5 shows this swift degradation clearly.

A comparison of Figures 4.3 and 4.5 shows us that the delay matching circuitry is able to improve the suppression across the band by approximately 30 dB. This figure will vary depending on the spacing between the delay taps α_0 and α_1.

4.4 Wideband Adaptation

The correlations of the RF signals $\tilde{v}_m(t)$, $\tilde{v}_m(t-\tau_1)$, $\tilde{v}_\epsilon(t)$, $\tilde{v}_\epsilon(t-\tau_2)$ and $\tilde{v}_o(t)$ required for gradient adaptation of the coefficients are performed in DSP by downconverting the signals at appropriate points in the feedforward circuit. To suppress third order IMD the correlation must be performed over a bandwidth of $3W_m$, where W_m is the bandwidth of the input signal $\tilde{v}_m(t)$. However, the maximum sampling rate of the DSP limits the linearisation bandwidth to less than the Nyquist sampling frequency $f_s/2$. Moreover, a bandstop filter is required to selectively suppress the signal component in $\tilde{v}_o(t)$ to avoid masking of the distortion component necessary for the correlation.

Wide bandwidth adaptation is achieved by performing partial correlations in selectable subbands and averaging the values of the partial correlations across the band. The size of the subband is restricted by the maximum sampling rate of the DSP used to perform the correlations (see Chapter 5) and also the selectivity of the available narrowband filter placed in the downconversion chain of the implementation shown in Section 5.1.3. The advantage of this method is the ability to select subbands in the output spectrum $V_o(f)$ which include only the distortion component for correlation with $V_e(f)$.

To avoid DC offsets caused by the mixers in the downconversion chains which bias the correlation, a novel use of DSP was developed in [1] to perform the baseband correlation. The RF signal subband is selected and downconverted to a frequency of one quarter the value of the DSP sampling rate. The bandwidth of the subband is chosen to be less than $f_s/4$ to ensure that the highest frequency component is less than the Nyquist sampling frequency and that the lowest spectral component is above DC.

In a standard implementation the subband would be downconverted to baseband using quadrature demodulation in DSP, then low pass filtered to suppress images generated at multiples of the sampling frequency. For this case the output of the filter would be

$$y(n) = \tilde{v}(n)e^{-j\frac{\pi}{2}n} \otimes h(n) \tag{4.19}$$

where $h(n)$ is the real low pass filter designed to filter out the left shifted DC offset.

To simplify design the quadrature demodulation step is combined with the low pass filtering by spectrally shifting the original low pass filter by $f_s/4$ to produce a complex bandpass FIR filter $\tilde{h}(n) = e^{j\frac{\pi}{2}n}h(n)$ as shown in Figure 4.6. In this way bandpass filtering is performed on the subband to filter out any DC offset and higher frequency components before baseband correlation is performed. The equivalent interpretation of the filter output is given as

$$y(n) = e^{-j\frac{\pi}{2}n}[\tilde{v}(n) \otimes \tilde{h}(n)] \tag{4.20}$$

The exponential factor preceding the convolution results in a rotating signal constellation at the frequency $f_s/4$. When the gradient correlations are performed the complex conjugate multiplication counter-rotates the spin equivalent to the desired baseband correlation.

The gradient estimates for the signal cancellation coefficients are thus

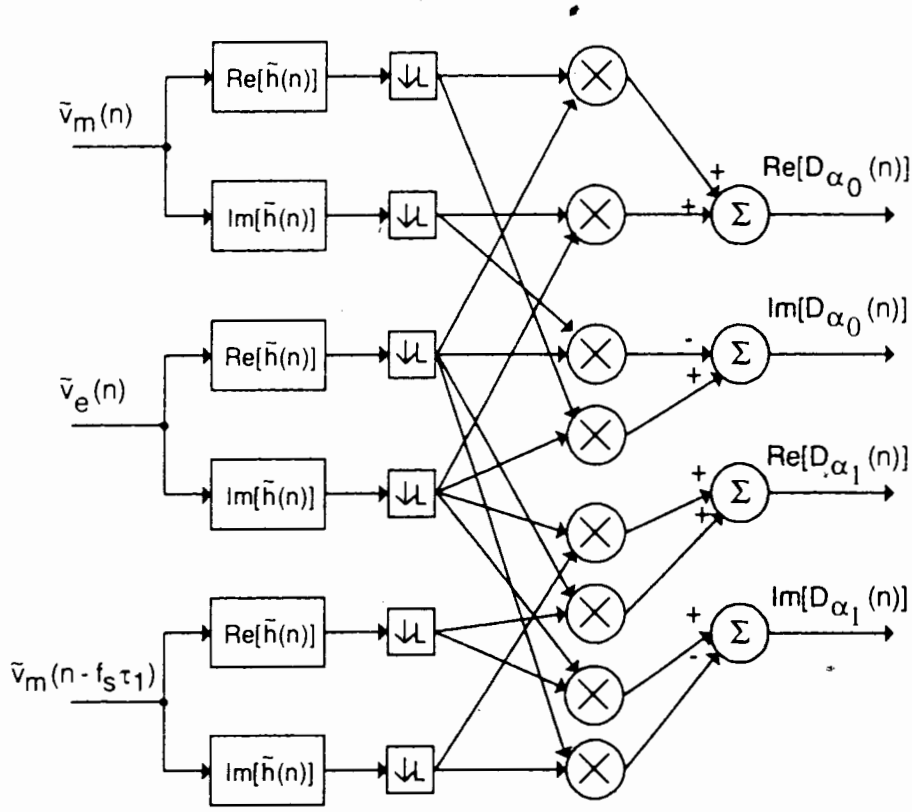

Figure 4.6: Complex filtering and multiplications performed to produce gradient estimates for α_0 and α_1.

$$D_{\alpha_0}(n) = y_e(n)y_m^*(n) = [\widetilde{v}_e(n) \otimes \widetilde{h}(n)][\widetilde{v}_m(n) \otimes \widetilde{h}(n)]^* \quad (4.21)$$

$$D_{\alpha_1}(n) = y_e(n)y_m^*(n-\tau_1) = [\widetilde{v}_e(n) \otimes \widetilde{h}(n)][\widetilde{v}_m(n-\tau_1) \otimes \widetilde{h}(n)]^* \quad (4.22)$$

$$D_{\beta_0}(n) = y_o(n)y_e^*(n) = [v_o(n) \otimes \widetilde{h}(n)][\widetilde{v}_e(n) \otimes \widetilde{h}(n)]^* \quad (4.23)$$

$$D_{\beta_1}(n) = y_o(n)y_e^*(n-\tau_2) = [\widetilde{v}_o(n) \otimes \widetilde{h}(n)][\widetilde{v}_e(n-\tau_2) \otimes \widetilde{h}(n)]^* \quad (4.24)$$

4.4.1 Effect of Downconversion Errors

As shown in [1], nonideal filter properties such as amplitude ripple in the passband and nonlinear phase response do not bias the correlations. Also since the local oscillators are shared by each of the downconversion chains any frequency offset which would cause a rotation of the signal constellation will be cancelled when the complex multiplication is performed in the gradient equation. Phase offset does not bias the correlation however it will effect the direction of the gradient step and the sign of the LMS step parameter will vary accordingly.

4.4.2 Effect of Aliasing

The bandwidth of the subbands selected by the narrowband filter in the downconversion chain is limited by the maximum sampling rate of the DSP. To avoid aliasing the maximum frequency of the subband must be less than one half the DSP sampling rate. Due to the slow rolloff of the 45 MHz Crystal filter in the downconversion chain this is difficult to achieve, and consequently aliasing will occur. An analysis of the effect of this aliasing follows.

Assume that the filter has a rectangular lowpass filter response

$$H(f) = \begin{cases} 1, & |f| < W \\ a \ll 1, & elsewhere \end{cases}$$

where f_c is the centre frequency of the filter, $2W$ is the maximum subband bandwidth equal to the Nyquist frequency $f_s/2$, and a is the attenuation factor, $a \ll 1$. The signal spectrum is attenuated at the Nyquist frequency by a factor a.

Noting that the inverse Fourier transform of the attenuated images at the sampling rate is given by

$$aV(f - f_s) \leftrightharpoons av(t)e^{j2\pi f_s t} \tag{4.25}$$

thus we can make an approximation to gradient estimate from the contributions due to aliasing from the adjacent frequency shifted spectrum at $\pm f_s$

$$\begin{aligned} D_{\alpha_0}(t) &= v_\epsilon(t)[1 + ae^{j2\pi f_s t} + ae^{-j2\pi f_s t}]v_m^*(t)[1 + ae^{-j2\pi f_s t} + ae^{j2\pi f_s t}] & (4.26) \\ &= v_\epsilon(t)v_m^*(t)[1 + 4acos(2\pi f_s t) + 2a^2 cos(4\pi f_s t)] & (4.27) \end{aligned}$$

Taking the expectation to find the mean gradient results in

$$E\left[D_{\alpha_0}(t)\right] = E\left[v_\epsilon(t)v_m^*(t)\right] E\left[1 + 4acos(2\pi f_s t) + 2a^2 cos(4\pi f_s t)\right] \tag{4.28}$$

Clearly from (4.28), the mean gradient will be zero for $v_m(t)$ and $v_\epsilon(t)$ uncorrelated, and therefore aliasing does not produce a biased result. The additional factor due to high frequency foldback will increase the self noise of the gradient estimate and can be expected to increase the convergence time of the coefficients.

The complex filter used in the DSP code for the final downconversion to baseband helps to reduce the unwanted frequencies above $f_s/2$ that have been folded back into the spectrum and any DC offset that may bias the result.

4.5 Speed of Convergence

The analysis that follows applies to matched branches of the adaptation coefficients. If there are differencs in the gain and phase between the two branches then a complex constant can be included in the gradient equation to correct for these variations.

4.5.1 Decorrelation of Signals for Gradient Adaptation

As stated in Chapter 3, the gradient estimate for each coefficient is dependent on both coefficients in the associated cancellation loop. Coupling of the adaptations is clearly shown in the gradient equations below. This leads to long convergence times since each gradient is affected by changes made by the other gradient step.

$$\begin{aligned} D_{\alpha_0}(t) &= v_\epsilon(t)v_m^*(t) \\ &= \gamma_o v_m(t-\tau_a)v_m^*(t) - \alpha_0 v_m(t)v_m^*(t) - \alpha_1 v_m(t-\tau_1)v_m^*(t) \end{aligned} \tag{4.29}$$

$$\begin{aligned} D_{\alpha_1}(t) &= v_\epsilon(t)v_m^*(t-\tau_1) \\ &= \gamma_o v_m(t-\tau_a)v_m^*(t-\tau_1) - \alpha_0 v_m(t)v_m^*(t-\tau_1) - \alpha_1 v_m(t-\tau_1)v_m^*(t-\tau_1) \end{aligned} \tag{4.30}$$

$$\begin{aligned} D_{\beta_0}(t) &= v_o(t)v_\epsilon^*(t) \\ &= \gamma_o v_a(t-\tau_b)v_\epsilon^*(t) - \beta_0 v_\epsilon(t)v_\epsilon^*(t) - \beta_1 v_\epsilon(t-\tau_2)v_\epsilon^*(t) \end{aligned} \tag{4.31}$$

$$\begin{aligned} D_{\beta_1}(t) &= v_o(t)v_\epsilon^*(t-\tau_2) \\ &= \gamma_o v_a(t-\tau_b)v_\epsilon^*(t-\tau_2) - \beta_0 v_\epsilon(t)v_\epsilon^*(t-\tau_2) - \beta_1 v_\epsilon(t-\tau_2)v_\epsilon^*(t-\tau_2) \end{aligned} \tag{4.32}$$

Due to the fast and slow modes in the eigenvalue spread, large step sizes toward the optimum value will cause the slow mode to become unstable and oscillate toward the point of convergence. If this step is too large convergence may never be reached. It would be hoped that the separate control of the two modes would allow manipulation of the time constants and noise contributions to the circuit coefficients.

The proposed method to reduce the effects of this problem with the aim of increasing the convergence speed of the coefficients is to decorrelate the signals $v_m(t)$ and $v_m(t-\tau_1)$ and thereby the gradient estimates so that the fast and slow modes can be independently controlled by changing the magnitude of the individual step parameters. This method will be analysed using the signal cancellation circuit, but it can easily be applied to the distortion cancellation circuit by decorrelating the signal $v_\epsilon(t)$ and $v_\epsilon(t-\tau_2)$.

Decorrelation of the signals is performed using the *Karhunen-Loève expansion* [19]. This provides a method of calculating the orthogonal basis coefficients c_1 and c_2 which represent the decorrelation of the components of the vector $\mathbf{v}_m$ given by

$$\mathbf{v}_m = \begin{bmatrix} v_m(t) \\ v_m(t-\tau_1) \end{bmatrix} \tag{4.33}$$

The vector $\mathbf{v}_m$ has a unique representation based on an orthonormal basis vector $\mathbf{y}_k$ with $k = 1, 2$, given by

$$\mathbf{v}_m = \sum_{k=1}^{2} c_k \mathbf{y}_k \tag{4.34}$$

where

$$\begin{aligned} c_k &= < \mathbf{v}_m, \mathbf{y}_k > \\ &= v_m(t)y(1)_k + v_m(t-\tau_1)y(2)_k \end{aligned} \tag{4.35}$$

The coefficients of (4.35) must satisfy $\mathbf{E}[c_k c_j^*] = 0$ if $k \neq j$ such that they are statistically orthogonal or uncorrelated and when $k = j, \mathbf{E}[c_k c_j^*] = \lambda_k$ where λ_k are the eigenvalues of the autocorrelation matrix of $\mathbf{v}_m$. The autocorrelation matrix $\mathbf{R}$ is given by

$$\mathbf{R} = \begin{bmatrix} P_m & R_m(\tau_1) \\ R_m^*(\tau_1) & P_m \end{bmatrix} \tag{4.36}$$

The orthonormal basis functions $\mathbf{y}_k$ can be found using

$$\mathbf{R}\mathbf{y}_k = \mathbf{E}[c_k^2]\mathbf{y}_k \tag{4.37}$$

The matrix of $\mathbf{E}[c_k^2]$ is equivalent to a diagonal matrix of the eigenvalues of $\mathbf{R}$.

$$\mathbf{E}[c_k^2] = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \tag{4.38}$$

The eigenvalues of $\mathbf{R}$ are found using the $\det(\mathbf{R} - \lambda\mathbf{I}) = \mathbf{0}$ which leads to

$$\lambda = P_m \pm \sqrt{R_m(\tau_1)R_m^*(\tau_1)} \tag{4.39}$$

The evaluation of the square root gives $\lambda_1 = P_m + |R_m(\tau_1)|$ and $\lambda_2 = P_m - |R_m(\tau_1)|$. By rearranging (4.37), the orthogonal basis functions $\mathbf{y}_k$ can be calculated using

$$(\mathbf{R} - \lambda_k)\mathbf{y}_k = 0$$

Corresponding to each eigenvalue the orthonormal eigenvectors are given as $\mathbf{y}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ \frac{R_m^*(\tau_1)}{\sqrt{R_m(\tau_1)^2}} \end{bmatrix}$ $\mathbf{y}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ \frac{-R_m^*(\tau_1)}{\sqrt{R_m(\tau_1)^2}} \end{bmatrix}$. The value of $\frac{R_m^*(\tau_1)}{\sqrt{R_m(\tau_1)^2}}$ is very close to unity if τ_1 is small so the eigenvectors are approximated as $\mathbf{y}_1 \approx \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{y}_2 \approx \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

With the knowledge of the basis function the orthogonal coefficients c_1 and c_2 can now be derived. Substituting $\mathbf{y}_1$ and $\mathbf{y}_2$ into (4.35) gives $c_1 = \frac{1}{\sqrt{2}}[v_m(t) + v_m(t - \tau_1)]$ and $c_2 = \frac{1}{\sqrt{2}}[v_m(t) - v_m(t - \tau_1)]$. This shows that the decorrelation of the signals $v_m(t)$ and $v_m(t - \tau_1)$ is equal to their sum given by c_1 and their difference given by c_2. Substituting these coefficients into (4.38), the nature of the orthogonality of the coefficients can be shown

$$\mathbf{E}[c_j c_k^*] = \begin{bmatrix} (P_m + \frac{1}{2}(R_m(\tau_1) + R_m^*(\tau_1))) & (\frac{1}{2}(R_m^*(\tau_1) - R_m(\tau_1))) \\ (\frac{1}{2}(R_m(\tau_1) - R_m^*(\tau_1))) & (P_m - \frac{1}{2}(R_m(\tau_1) + R_m^*(\tau_1))) \end{bmatrix} \quad (4.40)$$

For $R_m(\tau_1) \approx R_m^*(\tau_1)$ then matrix collapses into the form given by

$$\mathbf{\Lambda} = \begin{bmatrix} P_m + R_m(\tau_1) & 0 \\ 0 & P_m - R_m(\tau_1) \end{bmatrix} \quad (4.41)$$

Placing the orthonormal eigenvectors in a matrix known as the modal matrix $\mathbf{S}$ gives

$$\mathbf{S} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (4.42)$$

which conveniently is its own inverse $\mathbf{S} = \mathbf{S}^{-1}$. Note that the autocorrelation matrix is transformed to the form given in (4.41) using

$$\mathbf{\Lambda} = \mathbf{S}^{-1}\mathbf{R}\mathbf{S} \quad (4.43)$$

It has therefore been shown that the decoupled equivalents of the two input signals is in fact equal to the sum and the difference of the inputs. The sum is the fast mode of the adaptation and the difference represents the slow mode.

4.5.2 Adaptation using Decorrelated Gradients

In this section the associated decoupled stochastic gradients using the decoupled signals c_1 and c_2 are calculated. It is then shown how these gradients which represent the fast and slow modes of the adaptation can be individually controlled to increase the speed of convergence. Large step parameters can be used to increase the convergence time of the fast mode while using taking smaller steps in the slow mode to fine tune the coefficients to their optimal values.

The transformation of the signals can be expressed in the form

$$\begin{bmatrix} c_1(t) \\ c_2(t) \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} v_m(t) \\ v_m(t-\tau_1) \end{bmatrix} \tag{4.44}$$

The orthogonal true gradients are defined as

$$E[v_e(t)c_1^*(t)] = \frac{1}{\sqrt{2}} \left[\gamma_o \left(R_m^*(\tau_a) + R_m(\tau_1 - \tau_a)\right) - \alpha_0 \left(P_m + R_m(\tau_1)\right) - \alpha_1 \left(R_m^*(\tau_1) + P_m\right)\right] \tag{4.45}$$

$$E[v_e(t)c_2^*(t)] = \frac{1}{\sqrt{2}} \left[\gamma_o \left(R_m^*(\tau_a) - R_m(\tau_1 - \tau_a)\right) - \alpha_0 \left(P_m - R_m(\tau_1)\right) - \alpha_1 \left(R_m^*(\tau_1) - P_m\right)\right] \tag{4.46}$$

The stochastic gradients are given by

$$\begin{aligned} D_{\eta_1}(t) &= v_e(t)c_1^*(t) \\ &= \frac{1}{\sqrt{2}} \left[v_e(t)v_m^*(t) + v_e(t)v_m^*(t-\tau_1)\right] \end{aligned} \tag{4.47}$$

$$\begin{aligned} D_{\eta_2}(t) &= v_e(t)c_2^*(t) \\ &= \frac{1}{\sqrt{2}} \left[v_e(t)v_m^*(t) - v_e(t)v_m^*(t-\tau_1)\right] \end{aligned} \tag{4.48}$$

$D_{\eta_1}(t)$ is referred to as the gradient sum and $D_{\eta_2}(t)$ is the difference gradient. Multiplying these gradients by the 2 x 2 step size parameter matrix $\mathbf{K}$ allows for better control of the updates to the adaptation coefficients.

$$\mathbf{K} = \begin{bmatrix} K_1 & 0 \\ 0 & K_2 \end{bmatrix} \tag{4.49}$$

This leads to the scaled gradient estimates given by

$$\mathbf{KD}_\eta = \begin{bmatrix} K_1 D_{\eta_1}(t) \\ K_2 D_{\eta_2}(t) \end{bmatrix} \tag{4.50}$$

Due to the higher SNR of the gradient sum, the step parameter K_1 has a more significant effect on the speed of convergence of the coefficients. The smaller increments steps of the difference gradient controlled by K_2 are orthogonal to the direction of the steps taken by $K_1 D_{\eta_1}(t)$ and will act to fine tune the adaptation as it closes in on the optimum value.

Transforming this matrix by the modal matrix $\mathbf{S}^{-1}$ to return the scaled gradient estimates for the alpha coefficients gives

$$\begin{bmatrix} K_{\alpha_0} D_{\alpha_0}(t) \\ K_{\alpha_1} D_{\alpha_1}(t) \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} K_1 D_{\eta_1}(t) + K_2 D_{\eta_2}(t) \\ K_1 D_{\eta_1}(t) - K_2 D_{\eta 2}(t) \end{bmatrix} \tag{4.51}$$

which is equivalent to

$$\begin{bmatrix} K_{\alpha_0} D_{\alpha_0}(t) \\ K_{\alpha_1} D_{\alpha_1}(t) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} K_1 \left(v_e(t)v_m^*(t) + v_e(t)v_m^*(t-\tau_1)\right) + K_2 \left[v_e(t)v_m^*(t) - v_e(t)v_m^*(t-\tau_1)\right] \\ K_1 \left(v_e(t)v_m^*(t) + v_e(t)v_m^*(t-\tau_1)\right) - K_2 \left[v_e(t)v_m^*(t) - v_e(t)v_m^*(t-\tau_1)\right] \end{bmatrix} \tag{4.52}$$

Notice that if $K_1 = K_2$ then the solution reduces to the case with the original gradients

$$\begin{bmatrix} K_{\alpha_0} D_{\alpha_0}(t) \\ K_{\alpha_1} D_{\alpha_1}(t) \end{bmatrix} = \begin{bmatrix} K_1 \left(v_e(t)v_m^*(t)\right) \\ K_2 \left(v_e(t)v_m^*(t-\tau_1)\right) \end{bmatrix} \tag{4.53}$$

where $K_{\alpha_0} = K_1$ and $K_{\alpha_1} = K_2$. Since there is no reason to adapt α_0 and α_1 using different step parameters in the standard adaptation, the parameters will be set to $K_{\alpha_0} = K_{\alpha_1} = K$ for clarity in the proceeding analysis.

4.6 Analysis of Variable Delay Tracker

Decorrelation of the gradient signals for adaptation is used to increase the speed of convergence. This section derives equations to analyse the effect of this method on adaptation time and jitter in the coefficients and compare with the original method of adaptation.

The adaptation circuit shown in Figure 4.7 operates as a first order tracking loop. The transformation matrix $\mathbf{M} = \mathbf{M}^{-1}$ is substituted with the 2 x 2 identity matrix $\mathbf{I}$ for the case

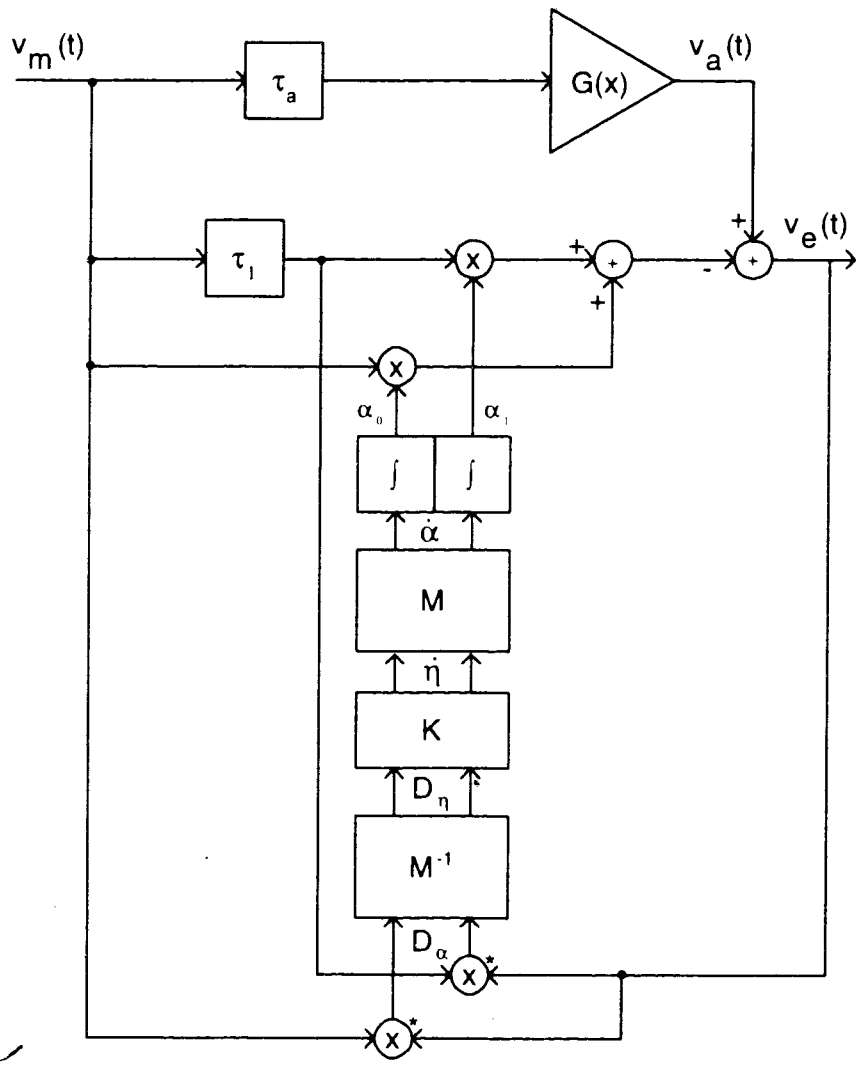

Figure 4.7: Complex baseband adaptation circuit for signal cancellation circuit

when the coefficients adapt together ($K_1 = K_2 = K$), and is equal to the modal matrix when adapting with decorrelated gradients. The corresponding linear model which illustrates the following equations is shown in Figure 4.8

The true gradients of the coefficients are expressed in matrix form as

$$\begin{bmatrix} E[D_{\alpha_0}(t)] \\ E[D_{\alpha_1}(t)] \end{bmatrix} = \begin{bmatrix} \gamma_o R^*_m(\tau_a) \\ \gamma_o R_m(\tau_1 - \tau_a) \end{bmatrix} - \begin{bmatrix} P_m & R^*_m(\tau_1) \\ R_m(\tau_1) & P_m \end{bmatrix} \begin{bmatrix} \alpha_0(t) \\ \alpha_1(t) \end{bmatrix} \tag{4.54}$$

By introducing the self noise of the gradient estimates in terms of a noise vector $\mathbf{n}(t)$, then from (4.54) the first order vector differential equation for the stochastic gradients can be expressed as

$$\mathbf{D}_\alpha(t) = \mathbf{w} - \mathbf{R}\boldsymbol{\alpha}(t) + \mathbf{n}(t) \tag{4.55}$$

where $\mathbf{D}_\alpha(t) = \begin{bmatrix} D_{\alpha_0}(t) \\ D_{\alpha_1}(t) \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} \gamma_o R^*_m(\tau_a) \\ \gamma_o R_m(\tau_1 - \tau_a) \end{bmatrix}$, $\mathbf{R} = \begin{bmatrix} P_m & R^*_m(\tau_1) \\ R_m(\tau_1) & P_m \end{bmatrix}$, $\boldsymbol{\alpha}(t) = \begin{bmatrix} \alpha_0(t) \\ \alpha_1(t) \end{bmatrix}$ and $\mathbf{n}(t) = \begin{bmatrix} n_0(t) \\ n_1(t) \end{bmatrix}$. Multiplying (4.55) by the matrix $\mathbf{K}$ of step-size parameters yields the gradient equation for the adaptation coefficients

$$\dot{\boldsymbol{\alpha}}(t) = \mathbf{K}\mathbf{w} - \mathbf{K}\mathbf{R}\boldsymbol{\alpha}(t) + \mathbf{K}\mathbf{n}(t) \tag{4.56}$$

where $\dot{\boldsymbol{\alpha}}(t) = \mathbf{K}\mathbf{D}_\alpha(t)$ and the state vector has initial conditions $\boldsymbol{\alpha}(0) = \begin{bmatrix} \alpha_0(0) \\ \alpha_1(0) \end{bmatrix}$. The choice of the initial conditions for $\boldsymbol{\alpha}(0)$ is arbitrary.

The solution to (4.56) can be found by considering the two inputs $\mathbf{K}\mathbf{w}$ and $\mathbf{K}\mathbf{n}(t)$ separately, thereby decomposing it to the zero input response and two zero state responses.

4.6.1 Zero Input Response

With no inputs, (4.56) becomes

$$\dot{\boldsymbol{\alpha}}(t) + \mathbf{K}\mathbf{R}\boldsymbol{\alpha}(t) = \mathbf{0} \tag{4.57}$$

The solution to this equation is well known, and can be verified by substitution into (4.57) to be

$$\boldsymbol{\alpha}_{zi}(t) = e^{-\mathbf{K}\mathbf{R}t}\boldsymbol{\alpha}(0) \tag{4.58}$$

where $e^{-\mathbf{K}\mathbf{R}t}$ is the matrix exponential, the fundamental matrix of the system.

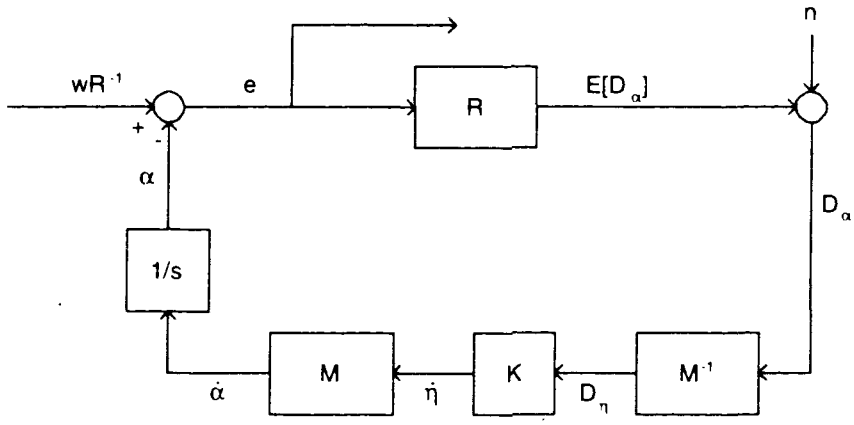

Figure 4.8: Linear model of adaptation of signal cancellation coefficients

4.6.2 Zero State Response to Noise Free Input

With no noise, (4.56) becomes

$$\dot{\boldsymbol{\alpha}}(t) + \mathbf{KR}\boldsymbol{\alpha}(t) = \mathbf{Kw} \tag{4.59}$$

with $\boldsymbol{\alpha}(0) = \mathbf{0}$ for the zero state solution. Again the solution can be verified to be

$$\boldsymbol{\alpha}_{zsl}(t) = (\mathbf{I} - e^{-\mathbf{KR}t})\mathbf{R}^{-1}\mathbf{w} \tag{4.60}$$

The final converged value at $t = \infty$ is given as

$$\boldsymbol{\alpha}_{zsl}(\infty) = \mathbf{R}^{-1}\mathbf{w}. \tag{4.61}$$

The fundamental matrix $\mathbf{e}^{-\mathbf{KR}t}$ can be represented in terms of the eigenvalues of $\mathbf{R}$, defined previously as $\lambda_1 = P_m + |R_m(\tau_1)|$ and $\lambda_2 = P_m - |R_m(\tau_1)|$, which provide the exponents of the fundamental set of solutions to the differential equation. Noting that

$$e^{\mathbf{\Lambda}t} = \begin{bmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{bmatrix} \tag{4.62}$$

then the fundamental matrix expressed in terms of the eigenvalues by transformation with the modal matrix $\mathbf{S}$ is

$$e^{-\mathbf{KR}t} = \mathbf{S}e^{-\mathbf{K}\Lambda t}\mathbf{S}^{-1} \tag{4.63}$$
$$= \frac{1}{2}\begin{bmatrix} \left(e^{-K\lambda_1 t} + e^{-K\lambda_2 t}\right) & \left(e^{-K\lambda_1 t} - e^{-K\lambda_2 t}\right) \\ \left(e^{-K\lambda_1 t} - e^{-K\lambda_2 t}\right) & \left(e^{-K\lambda_1 t} + e^{-K\lambda_2 t}\right) \end{bmatrix}$$

4.6.3 Zero State Response to Noise

The response to self noise alone is described by rewriting (4.56) as

$$\dot{\boldsymbol{\alpha}}(t) + \mathbf{KR}\boldsymbol{\alpha}(t) = \mathbf{Kn}(t) \tag{4.64}$$

with $\boldsymbol{\alpha}(0) = \mathbf{0}$. The solution is

$$\boldsymbol{\alpha}_{zs2}(t) = \int_0^t e^{-\mathbf{KR}(t-s)}\mathbf{Kn}(s)ds \tag{4.65}$$

4.6.4 Total Solution

The sum of the solutions given by (4.58),(4.60) and (4.65) gives the unique solution

$$\boldsymbol{\alpha}(t) = e^{-\mathbf{KR}t}\boldsymbol{\alpha}(0) + (\mathbf{I} - e^{-\mathbf{KR}t})\mathbf{R}^{-1}\mathbf{w} + \int_0^t e^{-\mathbf{KR}(t-s)}\mathbf{Kn}(s)ds \tag{4.66}$$

4.6.5 System Decomposition

Using the eigenvectors of the system matrix $\mathbf{R}$ as a new basis for the state space, the state vector $\boldsymbol{\alpha}(t)$ can be transformed into new coordinates $\boldsymbol{\eta}(t)$ using

$$\boldsymbol{\eta}(t) = \mathbf{S}^{-1}\boldsymbol{\alpha}(t) \tag{4.67}$$

Substituting $\mathbf{S}\boldsymbol{\eta}(t)$ into (4.56) and multiplying by $\mathbf{S}^{-1}$ transforms the state equation to

$$\dot{\boldsymbol{\eta}}(t) = \mathbf{KS}^{-1}\mathbf{w} - \mathbf{K}\Lambda\boldsymbol{\eta}(t) + \mathbf{KS}^{-1}\mathbf{n}(t) \tag{4.68}$$

with $\boldsymbol{\eta}(0) = \mathbf{S}^{-1}\boldsymbol{\alpha}(0)$

Equation (4.68) represents a set of two uncoupled first-order differential equations which can be solved independently since the components of $\boldsymbol{\eta}(t)$ are mutually orthogonal. In this case the step parameters K_1 and K_2 be varied to independently control the eigenvalues as

shown in (4.69). The value chosen for K_1 will have the most significant effect on the speed of convergence.

$$e^{-\mathbf{K}\Lambda t} = \begin{bmatrix} e^{-K_1\lambda_1 t} & 0 \\ 0 & e^{-K_2\lambda_2 t} \end{bmatrix} \tag{4.69}$$

4.7 Misadjustment error

The misadjustment error in the signal cancellation coefficients can be represented by the error vector $\mathbf{e}(t)$ as the difference between the value of $\boldsymbol{\alpha}(t)$ and the optimum value with no noise $E[\boldsymbol{\alpha}(\infty)]$

$$\begin{aligned} \mathbf{e}(t) &= \boldsymbol{\alpha}(t) - E[\boldsymbol{\alpha}(\infty)] \\ &= \mathbf{e}^{-\mathbf{KR}t}(\boldsymbol{\alpha}(0) - \boldsymbol{\alpha}(\infty)) \end{aligned} \tag{4.70}$$

Noting that $\dot{\boldsymbol{\alpha}}(\infty) = 0$ then (4.56) can be represented in terms of the error vector which leads to

$$\dot{\mathbf{e}}(t) = \dot{\boldsymbol{\alpha}}(t) - \dot{\boldsymbol{\alpha}}(\infty) = \mathbf{Kw} - \mathbf{KR}(\mathbf{e}(t) + \boldsymbol{\alpha}(\infty)) + \mathbf{Kn}(t) \tag{4.71}$$

Since $\boldsymbol{\alpha}(\infty) = \mathbf{wR}^{-1}$ (4.71) reduces to

$$\dot{\mathbf{e}}(t) = -\mathbf{KRe}(t) + \mathbf{Kn}(t) \tag{4.72}$$

which again is modelled in state space terms. The inverse for this system is

$$\mathbf{e}(t) = \int_0^t e^{-\mathbf{KR}(t-s)}\mathbf{Kn}(\mathbf{s})ds + \mathbf{e}^{-\mathbf{KR}t}\mathbf{e}(0) \tag{4.73}$$

The criterion of minimising the mean square error σ_e^2 of the error signal in signal cancellation circuit is useful for providing a metric for the amount of jitter in the coefficients

$$\sigma_e^2 = E[v_e(t)v_e^*(t)] \tag{4.74}$$

The error signal is given by

$$v_e(t) = \gamma_o v_m(t - \tau_a) + v_d(t - \tau_a) - \boldsymbol{\alpha}^T\mathbf{v}_m \tag{4.75}$$

and the variance is found to be

$$\sigma_e^2 = |\gamma_o|^2 P_m + P_d + 2Re[\boldsymbol{\alpha}\mathbf{w}^*] + \boldsymbol{\alpha}^*\mathbf{R}\boldsymbol{\alpha} \tag{4.76}$$

Substituting for $\boldsymbol{\alpha} = \mathbf{e} - \boldsymbol{\alpha}(\infty)$ in (4.76) yields

$$\sigma_e^2 = |\gamma_o|^2 P_m + P_d - \mathbf{w}^\dagger \mathbf{R}^{-1} \mathbf{w} - \mathbf{e}^\dagger \mathbf{R} \mathbf{e} \tag{4.77}$$

The component $|\gamma_o|^2 P_m - \mathbf{w}^\dagger \mathbf{R}^{-1} \mathbf{w}$ in (4.77) represents the minimum estimation error variance which is minimised by keeping the taps in the delay matching circuit small. P_d is the IM distortion power. The additional factor $\mathbf{e}^\dagger \mathbf{R} \mathbf{e}$ is the excess mean square error or jitter in the output. If the error vector $\mathbf{e}$ jitters due to self noise from the stochastic gradient signal then the criterion is the average value of $[\mathbf{e}^\dagger \mathbf{R} \mathbf{e}]$.

The autocorrelation matrix for the error vector is equivalent to

$$\mathbf{R}_e = \mathbf{e}\mathbf{e}^\dagger \tag{4.78}$$

and the criterion is expressed as

$$E[\mathbf{e}\mathbf{R}\mathbf{e}^\dagger] = tr[\mathbf{R}\mathbf{R}_e] = tr[\mathbf{R}_e \mathbf{R}] \tag{4.79}$$

where tr denotes the trace of the matrix equivalent to the sum of the diagonal components. Declaring $\mathbf{e}$ in terms of the error in the decorrelated system represented by $\boldsymbol{\nu}$

$$\mathbf{e} = \mathbf{S}\boldsymbol{\nu} \tag{4.80}$$

leads to the equivalent metric for the decorrelated system

$$\begin{aligned} E[\boldsymbol{\nu}^\dagger \mathbf{S}^\dagger \mathbf{R} \mathbf{S} \boldsymbol{\nu}] &= E[\boldsymbol{\nu}^\dagger \boldsymbol{\Lambda} \boldsymbol{\nu}] \\ &= tr[\boldsymbol{\Lambda} \mathbf{R}_\nu] \end{aligned} \tag{4.81}$$

where $\mathbf{R}_\nu = \boldsymbol{\nu}^\dagger \boldsymbol{\nu}$ is the autocorrelation matrix for the decorrelated error vector. Equation (4.81) is equivalently expressed as

$$tr[\boldsymbol{\Lambda} \mathbf{R}_\nu] = \lambda_1 \sigma_{\nu_0}^2 + \lambda_2 \sigma_{\nu_1}^2 \tag{4.82}$$

This simple expression represents the misadjustment error or jitter in the coefficients in terms of the fast and slow eigenvalue modes and the decorrelated error variances.

4.8 Noise Input

This section develops the equations which define the system's response to the self noise input from the stochastic gradient signals. The equivalent equations are also derived for the decoupled system.

The self noise from the stochastic gradient can be represented as white noise and approximated as flat since the noise bandwidth is much wider in comparison to the signal bandwidth. The autocorrelation of this noise input is expressed by

$$\mathbf{R}_n = E[\mathbf{n}(t)\mathbf{n}^\dagger(t)] \tag{4.83}$$

Near convergence the effect of the initial error in $\mathbf{e}(0)$ is negligible and by substituting $\beta = (t - s)$, (4.73) can be expressed as

$$\mathbf{e}(t) = \int_0^t e^{-\mathbf{KR}\beta}\mathbf{Kn}(t - \beta)d\beta \tag{4.84}$$

which is the convolution of the impulse response with an input noise vector. Substituting into (4.78) to find autocorrelation matrix for the error vector yields

$$\mathbf{R}_e = \int_0^t \int_0^t \mathbf{n}^\dagger(t - \beta_1)\mathbf{K}^\dagger e^{-\mathbf{KR}\beta_1} e^{-\mathbf{KR}\beta_2}\mathbf{Kn}(t - \beta_2)d\beta_1 d\beta_2 \tag{4.85}$$

Since the noise is assumed as white then the cross correlation is equal to zero. Thus the equation reduces to

$$\mathbf{R}_e = \int_0^t \int_0^t tr[\mathbf{Q}\mathbf{R}_n]\boldsymbol{\delta}(\beta_1 - \beta_2)d\beta_1 d\beta_2 \tag{4.86}$$

where $\mathbf{Q} = \mathbf{K}^\dagger e^{-\mathbf{KR}\beta_1} e^{-\mathbf{KR}\beta_2}\mathbf{K}$. Noting that $\beta_1 = \beta_2$ and $\mathbf{K}^\dagger = \mathbf{K}$ then this equation can be written as

$$\begin{aligned} \mathbf{R}_e &= \int_0^\infty tr[\mathbf{K}e^{-2\mathbf{KR}\beta}\mathbf{KR}_n]d\beta \\ &= tr[\frac{1}{2}\mathbf{R}^{-1}\mathbf{KR}_n] \end{aligned} \tag{4.87}$$

Substitution of (4.87) into (4.79) shows the criterion for the jitter as a function of the step parameters and the noise in the stochastic gradient

$$E[\mathbf{e}\mathbf{R}\mathbf{e}^\dagger] = tr[\frac{1}{2}\mathbf{KR}_n] \tag{4.88}$$

As a comparison the transformed error vector $\boldsymbol{\nu} = \mathbf{S}^{-1}\mathbf{e}$ is represented by

$$\dot{\boldsymbol{\nu}}(t) = -\mathbf{K}\boldsymbol{\Lambda}\boldsymbol{\nu}(t) + \mathbf{K}\mathbf{v}(t) \tag{4.89}$$

with initial condition $\boldsymbol{\nu}(0) = \mathbf{S}^{-1}\mathbf{e}(0)$, and $\mathbf{v}(t) = \mathbf{S}^{-1}\mathbf{n}(t)$. The inverse of equation is analogous to the criterion expressed in (4.73) and in the neighbourhood of convergence can be reduced to

$$\boldsymbol{\nu}(t) = \int_0^t e^{-\mathbf{K}\boldsymbol{\Lambda}\beta}\mathbf{K}\mathbf{v}(t-\beta)d\beta \tag{4.90}$$

Substituting into the autocorrelation matrix $\mathbf{R}_\nu$ gives

$$\mathbf{R}_\nu = \int_0^t \int_0^t \mathbf{v}^\dagger(t-\beta_1)\mathbf{K}^\dagger e^{-\mathbf{K}\boldsymbol{\Lambda}\beta_1} e^{-\mathbf{K}\boldsymbol{\Lambda}\beta_2}\mathbf{K}\mathbf{v}(t-\beta_2)d\beta_1 d\beta_2 \tag{4.91}$$

which reduces to

$$\mathbf{R}_\nu = tr[\frac{1}{2}\boldsymbol{\Lambda}^{-1}\mathbf{K}\mathbf{R}_\mathbf{v}] \tag{4.92}$$

thus the criterion for the jitter in the coefficients as given by (4.81) is shown to be

$$E[\boldsymbol{\nu}^\dagger\boldsymbol{\Lambda}\boldsymbol{\nu}] = tr[\frac{1}{2}\mathbf{K}\mathbf{R}_\mathbf{v}] = tr[\frac{1}{2}\mathbf{K}\mathbf{S}\mathbf{R}_\mathbf{n}\mathbf{S}^{-1}] \tag{4.93}$$

This equation relates the transformed self noise statistics to the transformed misadjustment error. It shows that the jitter in each of the sum and difference gradients (represented in $\boldsymbol{\nu}(t)$) can be controlled independently by selecting different values for K_1 and K_2.

4.9 Noise Statistics

Equation (4.93) expresses the jitter in terms of the self noise of the decoupled system. To complete the analysis the statistics of the self noise in the decoupled system must be determined. By substituting the result back into (4.93) it is easy to show the effect each step parameter has on the jitter in the coefficients.

As shown in [2] the variance of noise from the stochastic gradient in the neighbourhood of the convergence point where $\alpha = \gamma_o$ is given as

$$\sigma^2_{D_\alpha} = \int_{-\infty}^{\infty} x_m^2 |G(x_m) - \gamma_o|^2 p_m(x_m)dx_m \tag{4.94}$$

In a similar fashion the noise statistics for the present system are found by calculating the variance of the respective gradient signals. At the convergence point it was shown in

Section 4.6.1 that $\boldsymbol{\alpha} = \mathbf{R}^{-1}\mathbf{w}$. Assuming perfect interpolation of the coefficients such that $\alpha_0 v_m(t) + \alpha_1 v_m(t-\tau_1) \approx \gamma_o v_m(t-\tau_a)$ giving $v_e(t) = v_d(t)$ and representing the signals in amplitude form $r_{ma} = |x_m(t-\tau_a)|, r_{m1} = |v_m(t-\tau_1)|$ and $r_{m0} = |v_m(t)|$, the noise variance for the gradient of α_0 is defined as

$$\begin{aligned} \sigma_{n_0}^2 &= E[|v_d(t-\tau_a)|^2 |v_m(t)|^2] && (4.95) \\ &= \int\int_0^\infty r_{ma}^2 r_{m0}^2 \, |G(r_{ma}) - \gamma_o|^2 \, p(r_{ma}, r_{m0}) dr_{ma} dr_{m0} && (4.96) \end{aligned}$$

for $\tau_a \ll \frac{1}{W}$ where W is the bandwidth of the signal then it can be assumed that $r_{ma} \approx r_{m0}$ and (4.95) collapses down to equal the single branch variance given in (4.94). Similarly the variance of the stochastic gradient of $\alpha_1 is$ given as

$$\begin{aligned} \sigma_{n_1}^2 &= E[|v_d(t-\tau_a)|^2 |v_m(t-\tau_1)|^2] && (4.97) \\ &= \int\int_0^\infty r_{ma}^2 r_{m1}^2 \, |G(r_{ma}) - \gamma_o|^2 \, p(r_{m0}, r_{m1}) dr_{ma} dr_{m1} \end{aligned}$$

which again is similar to the single branch case. The covariance of the two gradient signals can also be found by

$$\begin{aligned} \sigma_{n_{01}}^2 &= E|v_d(t-\tau_a)|^2 v_m^*(t) v_m(t-\tau_1)| && (4.98) \\ &= \int\int\int_0^\infty r_{ma}^2 r_{m1} r_{m0} \, |G(x_{ma}) - \gamma_o|^2 \, p(r_{ma}, r_{m1} r_{m0}) dr_{ma} dr_{m1} dr_{m0} \end{aligned}$$

Combining these results to form noise covariance matrix $\mathbf{R}_n$ in terms of $\sigma_{D_\alpha}^2$ gives

$$\mathbf{R}_n = \sigma_{D_\alpha}^2 \begin{bmatrix} 1-\epsilon_0 & 1-\epsilon_{10} \\ 1-\epsilon_{10} & 1-\epsilon_1 \end{bmatrix} \qquad (4.99)$$

where ($0 < \epsilon_0, \epsilon_1 < \epsilon_{10} \ll 1$) are the errors in the approximation to the singular case.

Since $\boldsymbol{\nu}(t) = \mathbf{S}^{-1}\mathbf{n}(t)$ the noise statistics of the decorrelated gradients can simply be found by transforming (4.99) with the modal matrix inverse to produce

$$\mathbf{R}_v = \sigma_{D_\alpha}^2 \begin{bmatrix} 2 - \epsilon_{10} - \frac{\epsilon_1+\epsilon_0}{2} & \frac{\epsilon_1-\epsilon_0}{2} \\ \frac{\epsilon_1-\epsilon_0}{2} & \epsilon_{10} - \frac{\epsilon_1+\epsilon_0}{2} \end{bmatrix} \qquad (4.100)$$

Substituting (4.99) into (4.88) to find the expression for the criterion for jitter in the coefficients leads to

$$\frac{1}{2} tr[\mathbf{K}\mathbf{R}_n] = \frac{1}{2} K \sigma_{D_\alpha}^2 (1-\epsilon_0) + \frac{1}{2} K \sigma_{D_\alpha}^2 (1-\epsilon_1) \qquad (4.101)$$

which is approximately equal to $K\sigma^2_{D_\alpha}$.

For the case with decorrelation the criterion, substituting the noise statistics of (4.100) into (4.81) gives

$$\frac{1}{2}tr[\mathbf{KR}_v] = \frac{1}{2}K_1\sigma^2_{D_\alpha}(2 - \epsilon_{10} - \frac{\epsilon_0 + \epsilon_1}{2}) + \frac{1}{2}K_2\sigma^2_{D_\alpha}(\epsilon_{10} - \frac{\epsilon_0 + \epsilon_1}{2}) \tag{4.102}$$

Clearly the first term dominates so the step parameter K_1 will have more of an impact on the jitter. The jitter introduced by both methods with $K_1 = K$ is shown to be approximately equal, verifying that there is no degradation in system performance using the decorrelation method.

4.10 Summary of Analysis of Dynamics

This analysis has shown that the fast and slow modes in the coupled system, which require slow adaptation of the coefficients to maintain stability, can be separately controlled by decorrelating the gradient signals for the adaptation. The new gradients are equivalent to the sum of the original coupled gradients (the fast mode) and the difference between the gradients (the slow mode). In this way large step parameters can be used to control the fast mode which will increase the speed of convergence. The slow mode acts to fine tune the coefficients by taking small steps to separate them from the sum. Increasing the convergence speed will cause an increase in the jitter in the coefficients, however it is approximately equal to the jitter that is generated using the same step parameter in the coupled system.

Chapter 5

Circuit Implementation

The feedforward amplifier circuit hardware is primarily composed of Mini-circuits^{TM} SMA connectorised components. Each significant component in the circuit is described in the following sections and is followed by a full description of the circuit implementation.

5.1 Circuit Hardware Components

5.1.1 The Amplifiers

Figure 5.5 shows the schematic diagram of the signal cancellation circuit. A nonlinear but power efficient Class AB amplifier is used for the main amplifier and provides a gain of 28 dB with an output 1 dB compression point of 30 dBm. It is designed to operate over the frequency range of 1.7 -1.9 GHz with ±0.5 dB flatness. Over the 40 MHz linearisation bandwidth the delay variation with frequency changed by less than 0.1 ns.

All of the class A preamplifiers, except the one immediately preceding the main amplifier, have an output 1 dB compression point of 10 dBm. Each of the preamplifiers provides a gain of 20 dB with approximately 0.1 dB amplitude variation across the band of interest. The preamplifier preceding the main amplifier was chosen with a higher output power of 15 dBm to avoid operation near its region of saturation. Unfortunately, close to the completion of the project, this amplifier started producing unacceptable levels of distortion, in addition to significant gain loss, so had to be removed from the circuit. Instead another lower power preamplifier was added with 20 dB of attenuation to match the delays between the two branches. This may seem an unusual approach to add gain then cancel it immediately

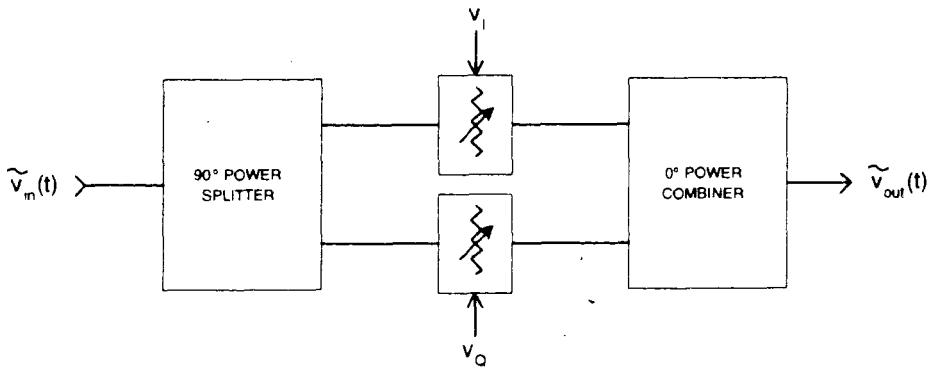

Figure 5.1: Schematic diagram of vector modulator

ahead of the PA. It would certainly have been more effective to have added some of this attenuation after the PA in the sampling coupler branch to increase the drive to the PA, but this change would have required considerable reworking of both cancellation circuits.

The error amplifier is a Class A amplifier with 40 dB gain and output 1 dB compression point of 33 dBm which ensures that the amplifier will be sufficiently backed off to operate in the linear region. Simulations showed that an input power of less than -15 dBm will ensure that the no significant IMD products are generated in the error amplifier. The frequency range of operation is 1 - 2 GHz and the amplitude variation with frequency over the 40 MHz bandwidth is 0.3 dB.

5.1.2 The Vector Modulators

The vector modulators (VM) controlling the attenuation and phase shift of the four complex coefficients are constructed using Mini-circuits power splitter/combiners and voltage controlled attenuators (VCA) as shown in Figure 5.1. The VCAs are designed to produce low harmonics and better IMD performance in comparison to a mixer used in the more conventional quadrature modulator. The first power splitter splits the input signal into two paths in phase quadrature which are then fed to the VCAs. The VCAs provide 0 or 180° phase rotation for positive or negative control voltages respectively and an associated attenuation which is a nonlinear function of the magnitude of the voltage applied. With these features controlling the input signal, the resulting vectors can be combined in the final power combiner to produce a resultant vector which can reach any phase shift from 0 to

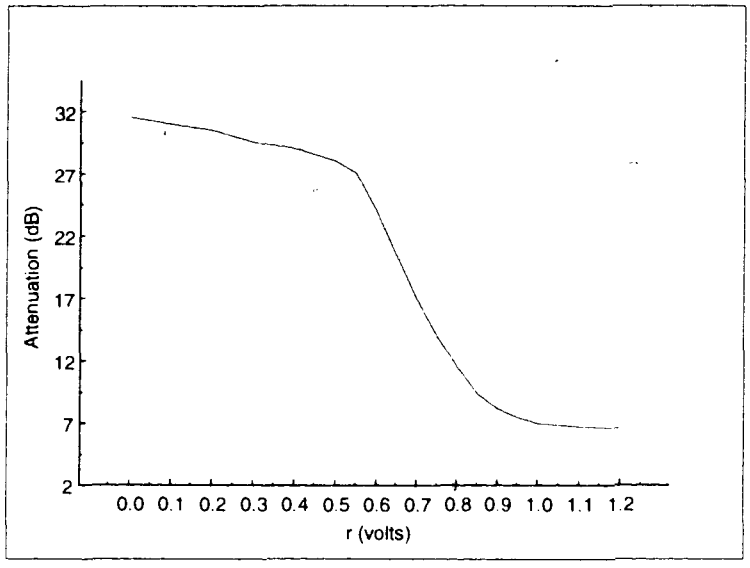

Figure 5.2: Vector Modulator measurements showing attenuation as a function of control voltage for $\phi = 45^o$

360^o. However, since the VCAs have a finite maximum attenuation value of 30 dB , the sum of the vectors will result in some unknown regions on the complex plane being unattainable. Fortunately this did not prove to be a problem. Measurements of attenuation provided by the vector modulator are plotted in Figure 5.2 as a function of $r = \sqrt{v_I^2 + v_Q^2}$ which shows that quick convergence of the coefficients will be achieved if the VM is operated at 15 dB nominal attenuation in the steepest part of the curve, corresponding to $r = 0.7$ V. Figure 5.3 show the measured values of phase shift provided by the vector modulator for varying the control phase $\phi = tan^{-1}(v_Q/v_I)$ using $r = 0.7$ V. This shows a monotonic increase in phase shift with increasing ϕ, changing rapidly in the mid-quadrants where $|v_I| = |v_Q|$. A plot of the attenuation in Figure 5.4 as a function of ϕ for $r = 0.7$ V shows the attenuation range for different values of ϕ.

Measurements across a 40 MHz span (1.78-1.82 GHz) with 0.7 V input voltage showed a 0.3 dB that was only roughly linear with frequency. To keep the levels of IMD generated by the vector modulator below -60 dBc the input power must be maintained below -5 dBm.

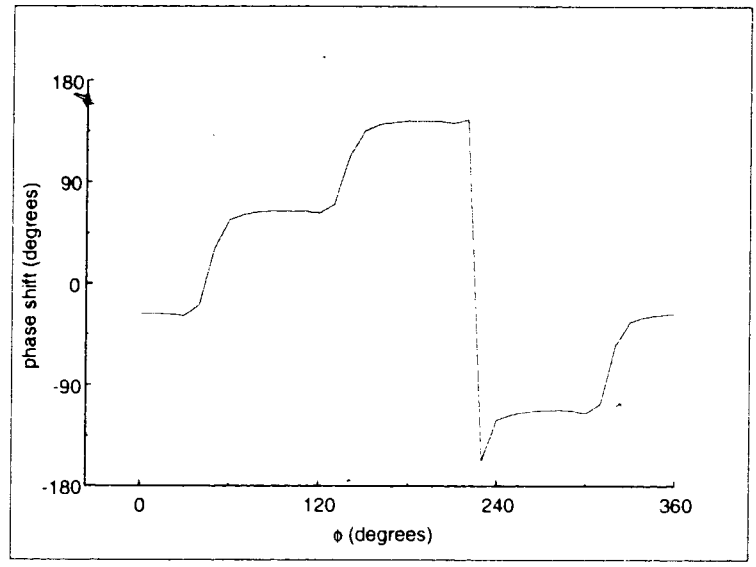

Figure 5.3: Measured phase shift through vector modulator as a function of ϕ for $r = 0.7$ V.

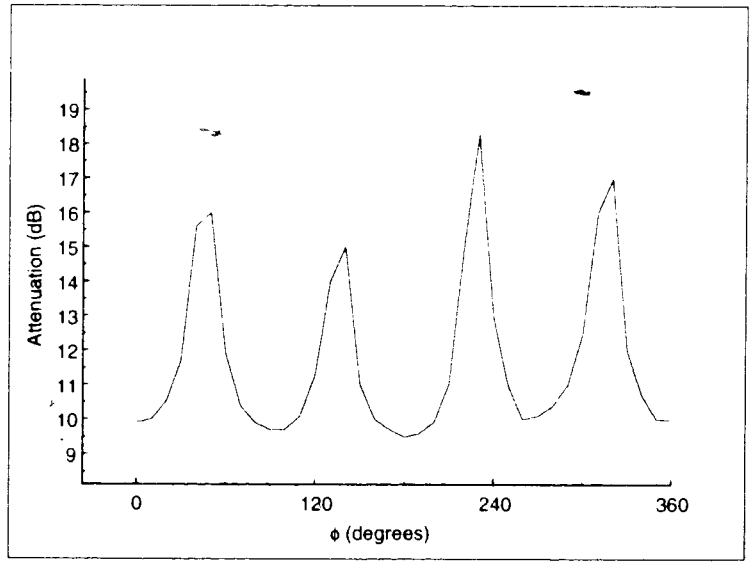

Figure 5.4: Measured attenuation through the vector modulator as a function of ϕ for $r = 0.7$ V.

5.1.3 The Downconversion Chain

The schematic diagram of the downconversion chain is shown in Figure 5.5. Power splitters are used to split the required signal for downconversion at appropriate points in the circuit. Each mixer shares the same local oscillators which are fed through two 3 way splitters to achieve a total of five signals for the five identical downconversion chains for $v_m(t)$, $v_m(t-\tau_1)$, $v_e(t)$, $v_e(t-\tau_2)$ and $v_o(t)$. The two outputs from the first splitter require 6 dB attenuators so that the levels from all five outputs will be approximately equal. Any relative phase shift between the oscillators will not bias the correlation as proven by Grant [1]. The first mixer performs a downconversion to 250 MHz using an LO frequency of 2.05 GHz.. A lowpass filter at 300 MHz removes the high frequency products before the second downconversion stage. This stage uses a local oscillator controlled by the PC via an HPIB interface to select various frequency subbands across the bandwidth of operation. Each band selected is downconverted to 45 MHz IF using an LO which steps around 205 MHz. A four pole monolithic crystal filter centered at 45 MHz $\pm$7.5 kHz cut-off frequency filters out the subband before the final downconversion stage to 9.5 kHz, one quarter of the maximum DSP sampling rate f_s. This frequency is chosen so that the subband can fit between DC and the Nyquist frequency $f_s/2$; however due to the low rolloff of the crystal filter (30 dB at $\pm$25 kHz) some signal frequencies can be expected to fall outside of this range, thus causing problems with aliasing and bias in the correlation due to the DC offset. It was shown in section 4.4.2 that aliasing does not effect the correlation but it will introduce self noise which will slow the adaptation time. The DC offset problem is resolved when the input signal is filtered with the complex bandpass filter in the DSP.

The 9.5 kHz subband is low pass filtered at 5 MHz and amplified to $\pm$1 V before being passed as an input to the DSP peripheral I/O board.

The downconversion chain posed a problem in selecting the first LO frequency high enough to avoid leakage back into the signal path due to poor isolation from the mixer. The level of 2.05 GHz LO leaking back was measured at -35 dBm. Since it is situated well away from the operating frequency of the circuit it is virtually transparent to the adapting coefficients. Proper design of a filter would be required in a practical implementation.

5.1.4 DSP Host Processor Board-PC/C32

The digital signal processing (DSP) board used to control the adaptation of the circuit coefficients is supplied by Loughborough Sound Images (LSI) and occupies a single 16 bit slot inside an Intel 80386 based PC. It uses the 50 MHz TMS320C32 floating-point DSP which achieves a performance of 25 Million Instructions Per Second (MIPS), or 40 ns instruction cycle time.

The PC/C32 DSP card has one bank of zero wait state 32k x 32 SRAM. It also provides a block of two wait state 2k x 16 DPRAM, used for fast data exchange between the PC and the C32 DSP without disrupting the processing of either device.

LSI provides software interface libraries which allow simple and easy access to the board's features from the PC. The board also comes equipped with a 16-bit parallel expansion interface called DSPLINK used for direct input/output to and from the DSP, avoiding the use of the I/O bus on the PC. Analog I/O channels are added to system using a standard multichannel analog I/O board connected via DSPLINK interface and is described in the following section.

Transfers between the two boards are interrupt driven. This will be explained in Section 5.3.

5.1.5 DSPLINK Multichannel I/O Board- PC/16IO8

A peripheral I/O board inserted inside the PC is used to provide the necessary interface from the DSP to the analog circuitry. It interfaces to the PC/C32 Host Processor Board via a 50 way shrouded DSPLINK connector, and an analog 50 way IDC connector at the endplate of the board is used for all analog input and output signals. Five out of the sixteen available Analog to Digital Converter (ADC) input channels are used for the signals from the downconversion chains; $\widetilde{v}_m(t)$, $\widetilde{v}_m(t-\tau_1)$, $\widetilde{v}_e(t)$, $\widetilde{v}_e(t-\tau_2)$ and $\widetilde{v}_o(t)$. All eight Digital to Analog Converter (DAC) channels were used to output the required voltages to the vector modulators, representing the real and imaginary parts of the adaptation coefficients α_0, α_1, β_0 and β_1.

Sampled data and control signals are passed over the 16 bit DSPLINK interface to the C32 processor board where correlations are performed and the coefficients updated. The maximum sampling frequency available was 38 kHz due to the number of input/output ports being utilised. Each channel has a resolution of 12 bits and a voltage range of ± 10

V. On board programmable gain amplifiers (PGA) and anti-alias/reconstruction filters are also provided prior to sampling by the ADC. The amplifiers are programmed with a gain of 10 to utilise the full ±10 V range of the ADCs. The Butterworth low-pass filters with -18 dB per octave roll-off in the stop band are set to a cutoff frequency of 17 kHz to provide a degree of anti-aliasing..

5.1.6 RF Input Signal Generation

Generation of desired test signals is made easy with the help of two phase locked HP33120A Arbitrary Waveform Generators (AWG). A Mathcad^{TM} worksheet was written to generate the required test signal at baseband and the real and imaginary components were written to separate files. Each of these files could be downloaded to the appropriate AWG via an HPIB interface for use as the quadrature modulator I and Q inputs to the HP ESG2000D signal generator. The signal generator was set to a carrier frequency of 1.8 GHz and the modulation bandwidth could be changed by setting output frequency of the AWGs. It is desirable to keep this value as large as possible so that the images at the sampling rate will be filtered by the AWGs anti-alias 10 MHz 7th order Bessel filter. However, the I/Q input bandwidth of the signal generator is limited to 15 MHz. An effect that could not be avoided was the tones produced at multiples of the output frequency from the carrier which caused unacceptable noise floors at certain frequencies.

A QPSK modulated input signal filtered with a root raised cosine filter at 35% roll-off was chosen to exercise the amplifier in the both nonlinear regions of cutoff and saturation. A 128 ksym/sec was used as a narrowband test signal and the output frequency was increased to test the operation at wider bandwidths.

5.2 Circuit Description

The circuit is described with reference to Figure 5.5 and Figure 5.6 which show the schematic diagrams of the signal and distortion cancellation circuits respectively. Unless otherwise specified, the measurements are made at 1.8 GHz centre frequency and the control voltage to the vector modulators is 0.7 V which corresponds to an attenuation of approximately 15 dB.

The RF test signal is applied to the input of the signal cancellation branch. This signal is split into two paths; the reference path in the lower branch and the main path in the upper

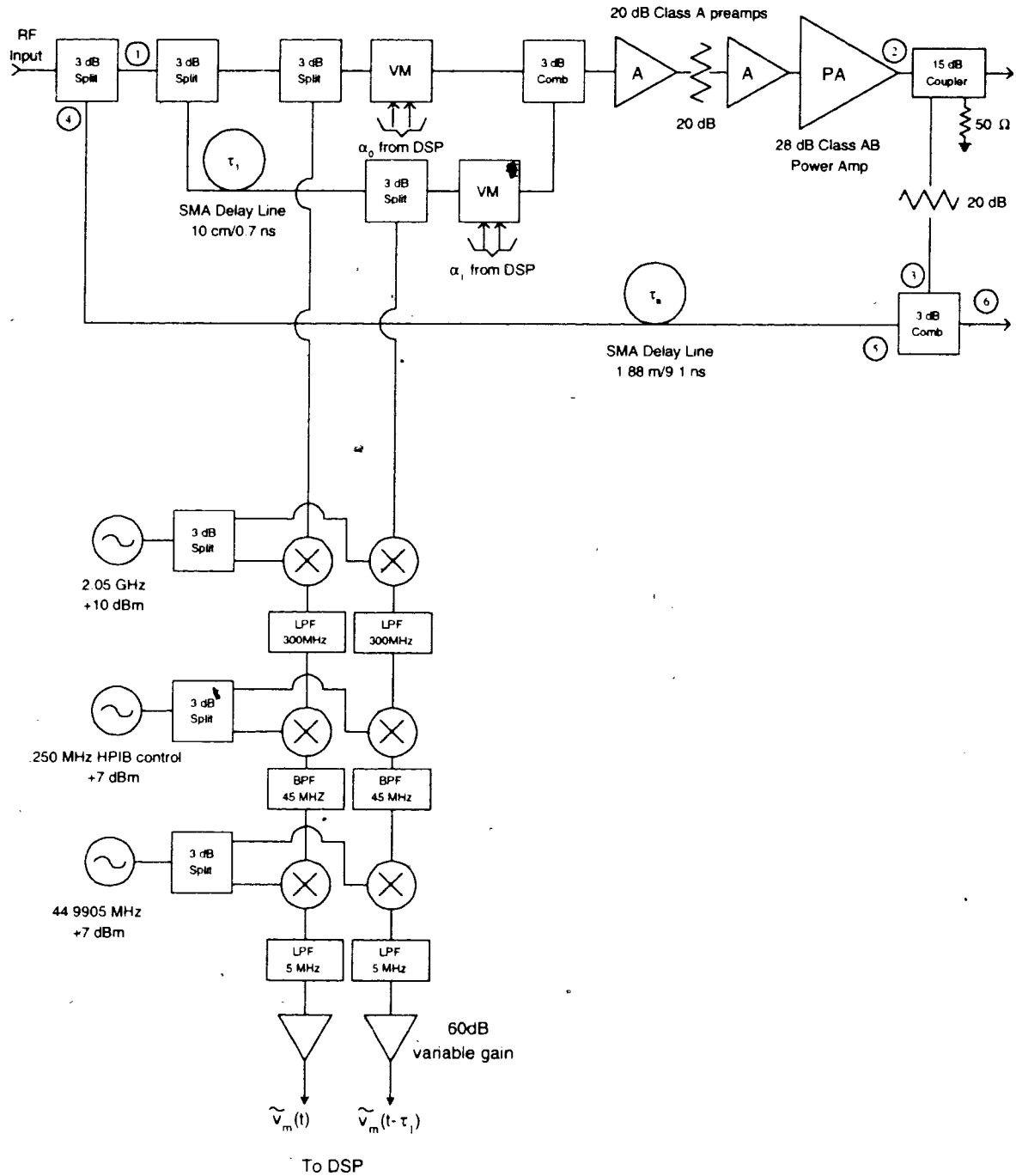

Figure 5.5: Schematic diagram of signal cancellation circuit

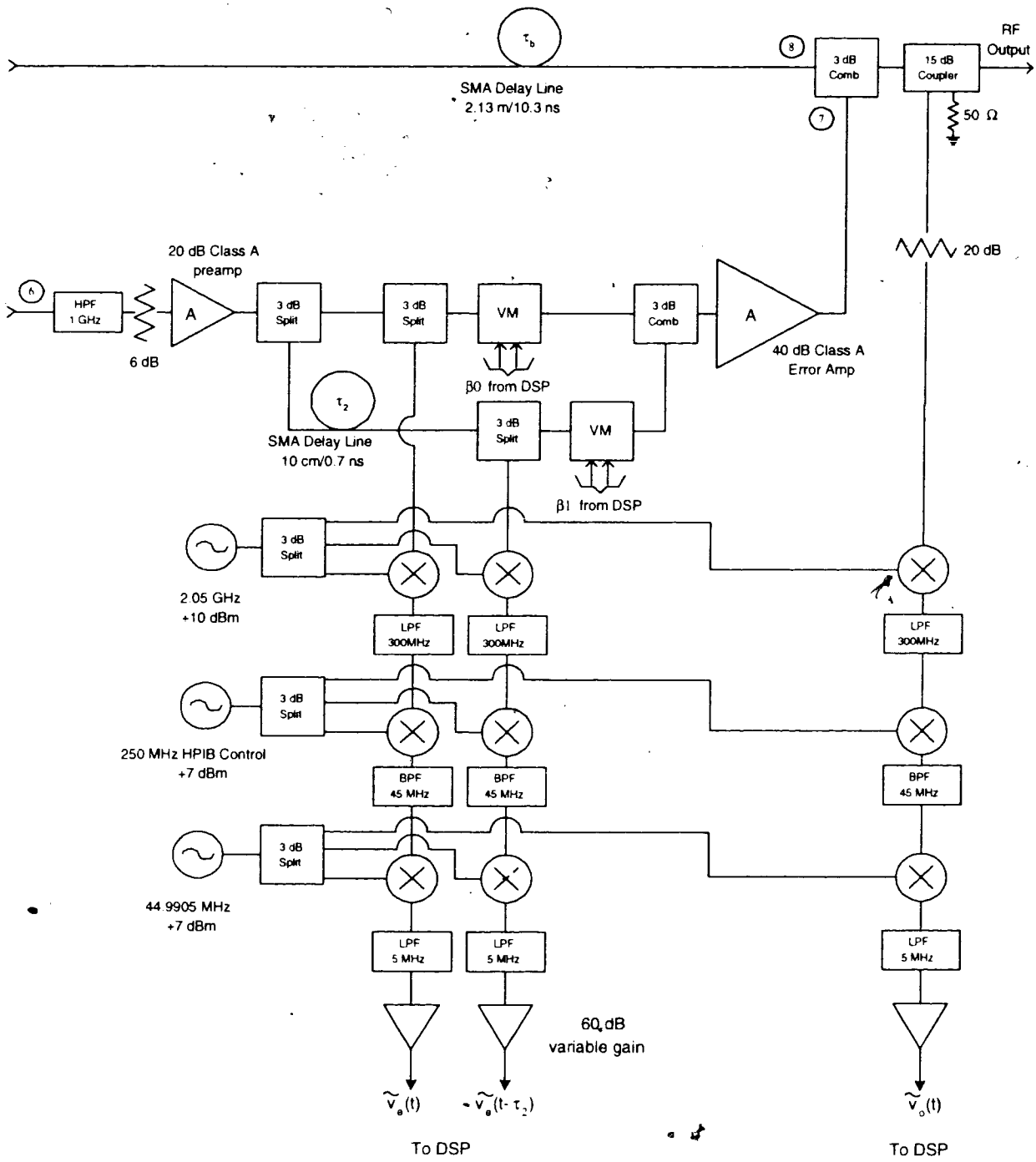

Figure 5.6: Schematic diagram of distortion cancellation circuit

Frequency (GHz)	α_0 path delay (ns)	τ_a path delay (ns)	α_1 path delay (ns)
1.78	8.9	9.1	9.6
1.79	8.9	9.2	9.7
1.80	8.8	9.1	9.6
1.81	8.7	9.3	9.5
1.82	8.7	9.1	9.4

Table 5.1: Measured delay in different paths of signal cancellation circuit

branch. The reference branch consists of a delay line with a measured delay of 9.1 ns. This is chosen to be approximately equal to the average group delay through the path 1-2-3. The main branch splits the input signal again for adaptation in the α_0 and α_1 vector modulators. The α_0 path first splits the signal for downconversion of $\tilde{v}_m(t)$ before the adjustment of the gain and phase in the VM. The α_1 path is delayed by a fixed delay, $\tau_1 = 0.7$ ns before it is split for downconversion of the signal $\tilde{v}_m(t - \tau_1)$. After undergoing attenuation and phase adjustment in the VM, the α_1 output signal is combined with the output from the α_0 path. The preamplifiers and attenuation are used to raise the power level to an appropriate value for input to the main amplifier. A portion of the output of the main amplifier is sampled by the coupler and attenuated such that the levels at the points 3 and 5 are approximately equal. The VMs will fine tune the levels in the main branch to match the level of the reference signal and when perfect adjustment is reached complete suppression of the signal will be achieved from the output of the first subtracting point.

Measurements of the group delay through the separate α_0 and α_1 paths 1-2-3 were made by disconnecting the VM in the other path and setting the control voltage to $r = 0.7$ V. Table 5.1 shows the varying delay measurements through each path across a bandwidth of 40 MHz. Clearly the reference delay τ_a falls within the delay measurements of α_0 and α_1.

The input power level of the1.8 GHz QPSK signal was limited by the signal generator which introduced distortion into the output signal when operated above an output power level of -1 dBm. Using this input power level ensures that the VMs are operated below -5 dBm input power to avoid introduction of additional intermodulation products. The PA output power is 26 dBm.

In the distortion cancellation circuit the upper reference branch contains the output of the amplifier $\tilde{v}_a(t)$ which is delayed by the fixed delay line of 10.3 ns such that the delay in path 2-8 is to approximately equal the average group delay through the path 2-3-6-7. Measurements of the delays through the separate β_0 and β_1 paths from point 2 to point 7

Frequency (GHz)	β_0 path delay (ns)	τ_b path delay (ns)	β_1 path delay (ns)
1.78	10.6	10.7	10.9
1.79	10.7	10.8	11
1.80	10.7	10.7	11
1.81	10.6	10.6	10.9
1.82	10.5	10.7	10.8

Table 5.2: Measured delay through different paths in distortion cancellation circuit

are compared in Table 5.2 with the measured delay through τ_b and the sampling coupler in path 2-8.

The lower branch passes the error signal output from the power combiner at point 6 through a high pass filter at 1 GHz. The 6 dB of attenuation is chosen so that the loss in path 2-8 measured at -3 dB is approximately equal to the loss through the path 2-3-6-7. With signal completely suppressed the error signal is at a much lower power level than the input signals and needs to be amplified before it is split for downconversion. The signal is again split to feed the β_0 and β_1 paths with a fixed delay line of 0.7 ns added to the β_1 path. The two paths are then recombined before amplification by the error amplifier to bring the level of distortion up to the same level as in the amplifier output signal at point 8. The error signal is then combined with the PA output to produce the final feedforward linearised output.

5.3 Software Design

5.3.1 FIR Filter Design

The complex bandpass FIR filter $\tilde{h}[n]$ used to filter the signals downconverted to 9.5 kHz is designed in the software package Matlab^{TM} using the Remez exchange algorithm. The passband of the filter is chosen as 10 kHz so that the 3 dB cutoff frequencies are of the same order as the crystal filter in the downconversion chain. High stopband attenuation is required to filter out any aliasing products above the Nyquist frequency of 19 kHz and to effectively attenuate any DC offsets which would bias the correlations. A filter length of $L = 32$ is required to meet the narrow passband requirements and provides 65 dB attenuation in the stopband. Figure 5.7 shows the magnitude response of the complex filter $\tilde{h}[n]$.

Note the same filter is used to filter all the signals from the downconversion chains. This is because the subbands are selected to downconvert bands which contain either the desired

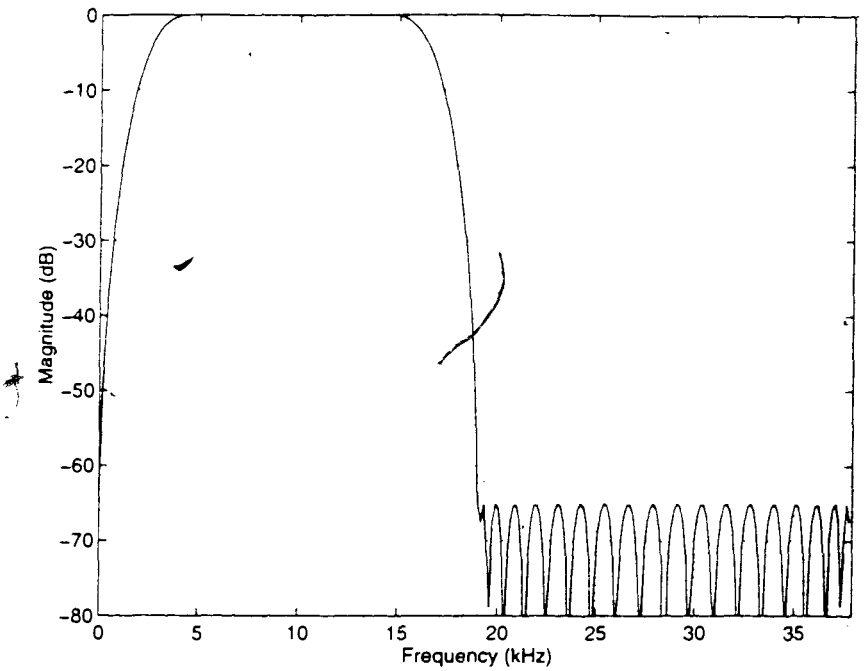

Figure 5.7: Complex bandpass filter used to filter $\widetilde{v}_m(t)$, $\widetilde{v}_m(t-\tau_1)$, $\widetilde{v}_e(t)$, $\widetilde{v}_e(t-\tau_2)$ and $\widetilde{v}_o(t)$ at 9.5 kHz.

signal or distortion only and subsequently a bandstop filter for the suppression of the signal in the beta adaptation is not required. This greatly reduces the length of the filter and the delays involved in performing the filtering in DSP.

5.3.2 TMS320C32 Assembly Code Design

The code written to control the adaptation of the coefficients is written using TMS320C32 assembly language. The code begins with a standard board initialisation procedure. The sampling rate of the ADC and DAC channels is set to 38 kHz. All analog input channels are sampled synchronously and the D/A channels are converted simultaneously at each clock pulse. Initial values are set for the DACs.

Transfers between the DSP board and the multichannel I/O board are interrupt driven. Level triggered interrupts are generated at the same frequency as the ADC sampling clock. This is equivalent to an interrupt every 26 μs or 657 machine cycles. Once the analog inputs are sampled the multi-channel board asserts an interrupt on the DSP chip's INT2 input pin which initiates the code contained in the interrupt service routine (ISR). The interrupt line

is released by accessing one of the ADC registers.

The ISR is used to read the current input samples from the ADC and output the computed values of the adaptation coefficients to the DACs. The ADC input samples are converted from 2's complement to IEEE floating point format and stored in a sample buffer for use in the main adaptation program.

The main program is used to filter the samples with the complex filter coefficients of $\tilde{h}[n]$ and generate the updates for the adaptation coefficients to be passed to the DACs. An FIR subroutine is used to calculate the filter outputs based on the previous L samples stored in memory. To reduce the amount of processing required by the DSP and to aid simplification of the code, the filter outputs are decimated by a factor equal to the length of the FIR filter. Once the sample input buffer is filled, the ISR raises a flag to the main program which takes the inner product of the buffer contents with the real and imaginary filter coefficients, thereby obtaining a new filter output every L samples. Decimation increases the convergence times by a factor of L. By applying more processing power decimation of the filter outputs would not be required and the speed of the adaptation increases by a factor of L.

Since the alpha coefficients use solely the desired signal band for adaptation and the beta coefficients require only the distortion bands to adapt, the code only needs to adapt one coefficient at each frequency subband selected. In a practical implementation an algorithm could search the frequency band for power spectra and make a decision based on power levels as to whether it is signal or distortion and adapt the correct coefficient accordingly. For a CDMA signal where the signal spectrum spans the entire bandwidth, the beta coefficients would be adapted on either side of the signal bandwidth.

A simple approach was taken to adapting the coefficients across the band. Knowing the frequency and bandwidth of the input signal and the width of the generated IMD bands on either side of the desired signal the selection of the subbands for downconversion is simplified. The linearisation bandwidth is divided into three sectors as shown in Figure 5.8, where the low and high frequency sectors encompass the IMD bands and the centre band spans the signal bandwidth.

The main program begins by adapting the signal cancellation coefficients at the centre frequency. Once a certain number of updates have been performed, chosen such that the signal cancellation coefficients have time to adapt, the program sends a request to the PC that it is ready to change frequency by setting the semaphore flag to 1. A windows program written as an interface between the PC and the DSP waits for the flag to be set and upon

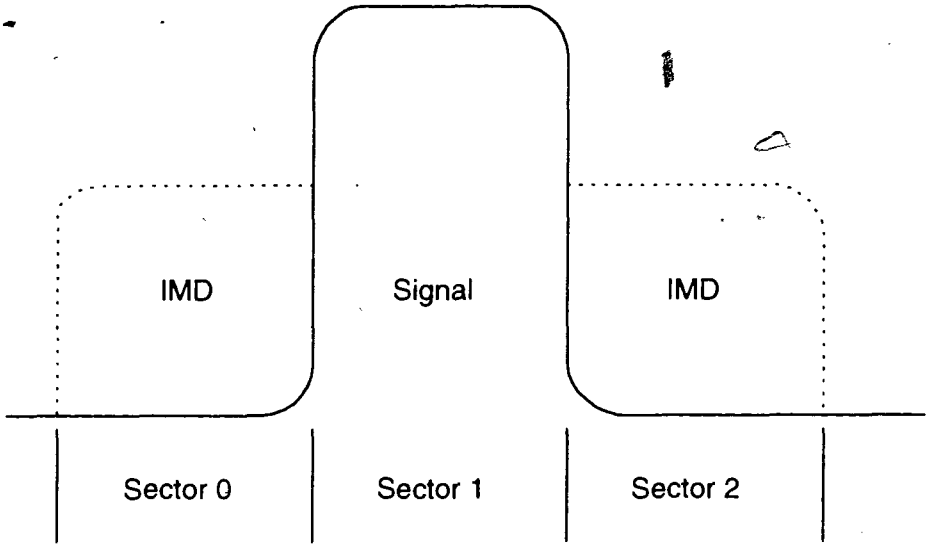

Figure 5.8: Representation of the dividing of the linearisation band for the coefficient adaptations.

acknowledgment decrements the frequency of the signal generator by a value specified by the user, such that it steps into the upper IMD band for adaptation of the beta coefficients. The PC also resets the semaphore flag to 0 and awaits a new request to change frequency. The coefficients have now adapted and linearisation at the feedforward amplifier output is achieved.

To maintain convergence of the coefficients the code begins an infinite cycle of stepping across the band, first adapting the beta coefficients in the upper band of IMD, then the alpha coefficients at the centre frequency, then to the beta coefficients in the lower IMD band. The PC returns the signal generator to the starting frequency specified in the upper band of IMD after each sweep across the band.

The DSP code waits until the semaphore flag is reset to 0 by the PC before continuing the code execution. Due to the delay between the PC sending the instruction to the signal generator to change frequency until the frequency is output, a timing delay was placed within the DSP code to count 2000 buffers before proceeding with the new coefficient adaptation. A delay of 2000 buffers corresponds to a wait of approximately 1.7 sec, which is more than sufficient. Use of separate LO's in the downconversion chains would reduce the need to change frequency between the alpha and beta adaptations and would allow for simultaneous

updates of the coefficients.

Originally a double buffering technique was used so that calculations could be performed on a full sample buffer while another buffer was being filled. However since the code is required to change the frequency of the signal generator it simplified the code to lock out the ISR while the correlations were being performed. Each of the five input channels has an input sample buffer of length L which are arranged contiguously in memory. The top buffer which contains samples of $v_m(t)$ is aligned on an 128-bit boundary. The individual channel buffers for $v_m(t-\tau_1)$, $v_e(t)$, $v_e(t-\tau_2)$ and $v_o(t)$ are filled by adding L, 2L, 3L and 4L respectively to the pointer of the first sample buffer.

The ISR fills the individual buffers and sets a flag to indicate when the buffer is full. The main program which takes care of filtering and the correlations, polls this flag until it is set then proceeds to perform the calculations on the full buffer. Since the ISR is locked out until the calculations are completed, context saves and restores are not required.

Each ISR takes approximately 105 cycles to complete and the buffers take 0.85 msec to fill. The primary functions of the ISR are as follows:

- Reads the input sample data from the ADCs and converts it from 2's complement integer format to floating point format
- Stores the input samples in the input sample buffers
- Outputs the current values of the alpha and beta coefficients to the DACs
- Checks for full buffer
- When buffer is full, signals main program using IOF flag

The main program polls the buffer flag until it is set then begins a rather complicated set of steps to perform frequency stepping across the band. To begin the ISR is disabled and a bit is tested which controls the first alpha adaptation at the centre frequency. When this is set to 1 the code performs the alpha correlations, clearing the full buffer flag and enabling the interrupts at the completion of the calculations then returning to poll for the next full buffer until the specified number of updates is reached (set to 10000 to minimise the number of frequency changes). Once this occurs the frequency semaphore flag is set and the value of the bit is then decremented to 0 to allow for normal operation of stepping across the frequency band. The code then executes the timing loop until adequate time has been

given for the signal generator to output the change of frequency. A flag is set so that the code will skip directly to the adaptations once the timing loop has terminated. The value of the sectors is then tested to determined which adaptation will be performed. When it is set to the value of 2 or 0, the code will branch to the beta adaptation using data sampled from the bands on either side of the desired signal. If it is set to 1 then the alpha coefficients will adapt at the centre frequency of the band. After each update the program checks to see if all the updates have been performed. If so, the code will request a change of frequency and waits for the semaphore flag to be reset to 0 before proceeding. The updates counter and flags are reset and the sector value is decremented. At the completion of this code the ISR will be enabled and the flag indicating the full buffer will be cleared. The program will continue polling for the next full buffer flag. When the program has cycled through all three adaptations across the band, it will return to the beta adaptation in the upper IMD band and continue frequency stepping until the PC stops the program.

The DSP algorithm for the adaptation of the alpha coefficients is described with reference to Figure 5.9. The bandpass signals $\widetilde{v}_m(t), \widetilde{v}_m(t-\tau_1)$ and $\widetilde{v}_e(t)$ are filtered with the real and imaginary coefficients of the complex filter $\widetilde{h}(n)$ using an FIR subroutine. The diagram shows the decimation factor L of the filter outputs. After filtering, the complex multiplication to form the gradient estimates $D_{\alpha_0}(t) = v_m^*(t)v_e(t)$ and $D_{\alpha_1}(t) = v_m^*(t-\tau_1)v_e(t)$ is performed using the real and imaginary parts of the rotating signal constellations. The resulting derotated gradient estimates for α_0 and α_1 are then transformed using the sum and difference transformation matrix and the decorrelated gradient estimates are multiplied by the step parameters K_1 and K_2. Inverse transformation of the scaled gradient estimates is then performed and the values are accumulated using the familiar LMS algorithm given by

$$\alpha_0(n) = \alpha_0(n-1) + K_1 v_e(n)[v_m(n) + v_m(n - f_s\tau_1)]^* + K_2 v_e(n)[v_m(n) - v_m(n - f_s\tau_1)]^* \quad (5.1)$$

$$\alpha_1(n) = \alpha_1(n-1) + K_1 v_e(n)[v_m(n) + v_m(n - f_s\tau_1)]^* - K_2 v_e(n)[v_m(n) - v_m(n - f_s\tau_1)]^* \quad (5.2)$$

to produce the updated adaptation coefficients. The values are converted from floating point format to 2's complement integer format and stored in an output buffer to be output to the DACs.

The algorithm for the beta coefficients follows in a similar fashion, using $\widetilde{v}_e(t), \widetilde{v}_e(t-\tau_2)$ and $\widetilde{v}_o(t)$ as the filter input signals and specifying the step parameters as K_3 and K_4.

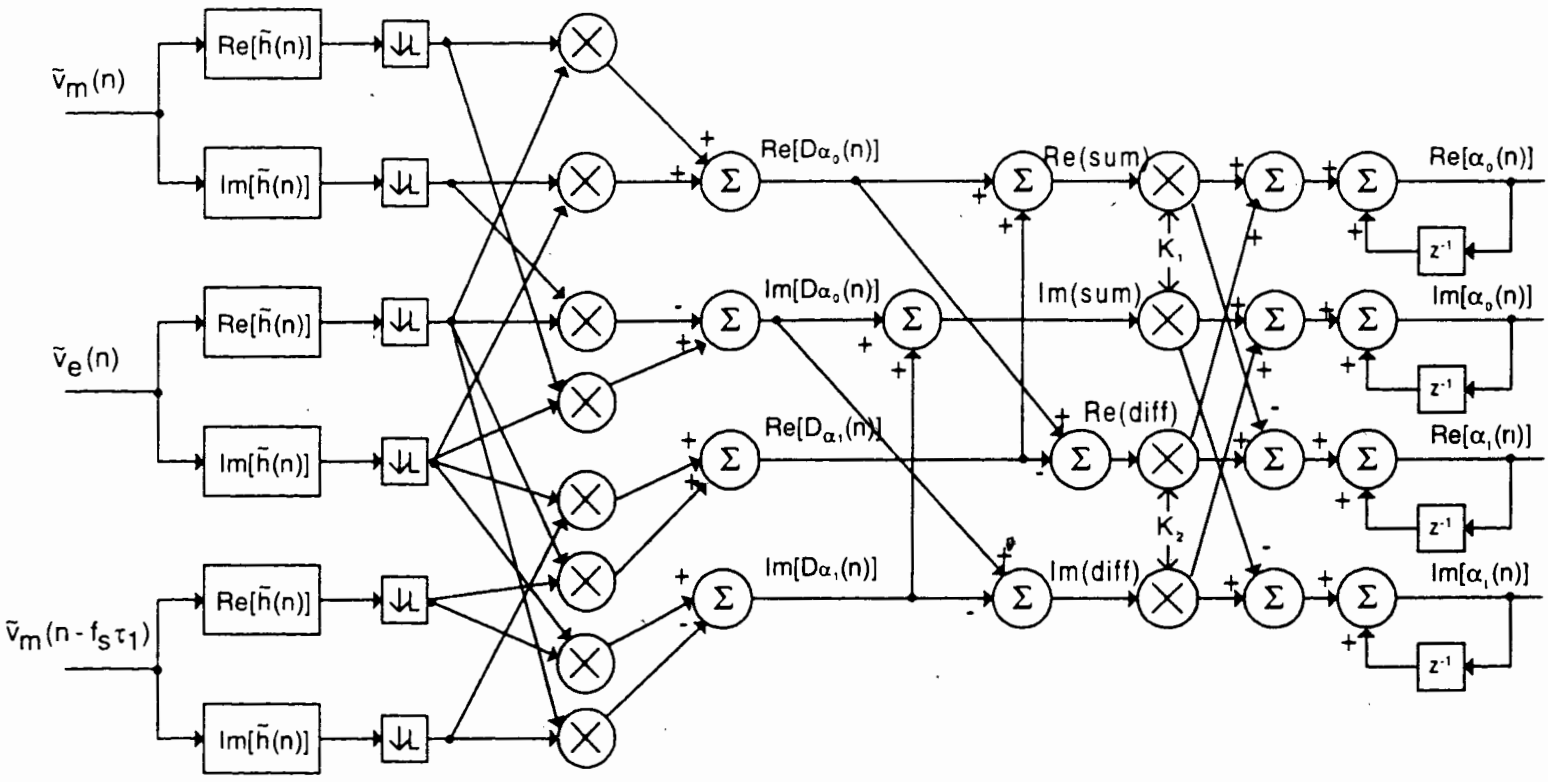

Figure 5.9: DSP algorithm for the adaptation of the signal cancellation coefficients

$$\beta_0(n) = \beta_0(n-1) + K_3 v_o(n)[v_e(n) + v_e(n - f_s\tau_2)]^* + K_4 v_o(n)[v_e(n) - v_e(n - f_s\tau_2)]^* \quad (5.3)$$

$$\beta_1(n) = \beta_1(n-1) + K_3 v_o(n)[v_e(n) + v_e(n - f_s\tau_2)]^* - K_4 v_o(n)[v_e(n) - v_e(n - f_s\tau_2)]^* \quad (5.4)$$

Chapter 6

Results

6.1 Comparison of Convergence Behaviour

Following through from the analysis performed in Chapter 4 which proposed the decorrelation of the gradient signals to speed the convergence time, a variety of measurements were carried out on the feedforward circuit to obtain a measure of the degree of improvement this new method of adaptation has on the overall performance. Choosing appropriate step-size parameters, the behaviour of the coefficients for the case where they adapt independently can be compared to the adaptation with decorrelated gradients. The tests were performed using a QPSK modulated input signal at 128 ksym/sec with 35% rolloff root raised cosine filtering. For each adaptation using a particular set of step parameters the coefficients were saved in dualport memory on the DSP board, then uploaded to a file for analysis.

6.1.1 Signal Cancellation Circuit Adaptation

To find an appropriate value for the sum step-size parameter K_1 and to analyse its effect on the convergence speed of the coefficients the sum of the gradients for α_0 and α_1 is adapted while the difference step-size parameter K_2 is set to zero. Plotting the imaginary part of α_0 against the real part forms a polar plot of the convergence path of α_0 as given in Figure 6.1. The path starts at the zero origin and is plotted for two different values of the step-size parameter K_1. Both paths converge on the same point, however the larger step-size parameter shows a more direct path toward the optimum value. The path for both coefficients is the same because they are incremented by the same gradient step and

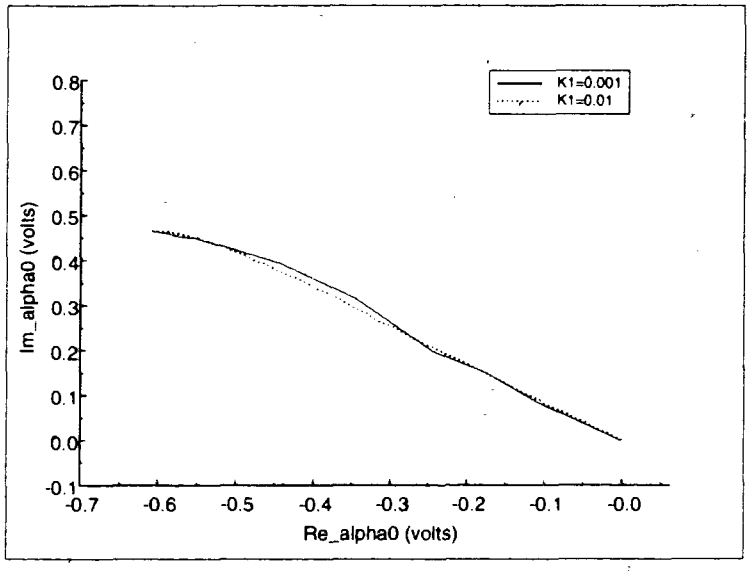

Figure 6.1: Path of the signal cancellation coefficients for different K_1 adapting the sum only ($K_2 = 0$).

therefore a plot of α_0 shows all the information required.

Figure 6.2 shows the convergence behaviour of the real part of α_0 with respect to time. Using the smaller step parameter $K_1 = 0.001$ results in a much longer convergence time (approximately 1.5 sec). Plotting both the real and imaginary parts of α_0 for $K_1 = 0.01$ on a much smaller time scale in Figure 6.3 shows the convergence time of α_0 is approximately 0.15 sec. Factoring in the decimation factor $L = 32$, the initial convergence time for the signal cancellation coefficients can be estimated at 4.7 msec, an excellent result.

Another interesting result relates to Figures 6.4 and 6.5 which show the convergence behaviour of the sum of the coefficients for varying the added attenuation in the upper branch of the signal cancellation circuit. The ideal case is when the levels at the input to the 180^o signal cancellation combiner are equal which is the case when the additional attenuation is equal to 20 dB. Reducing this value causes the adaptation coefficients to step blindly in the direction of the optimum attenuation then return to a lower voltage level to compensate for mismatch in attenuation between the branches. Ideally the vector modulators should be operated around 0.7 V to obtain fast convergence times and matching

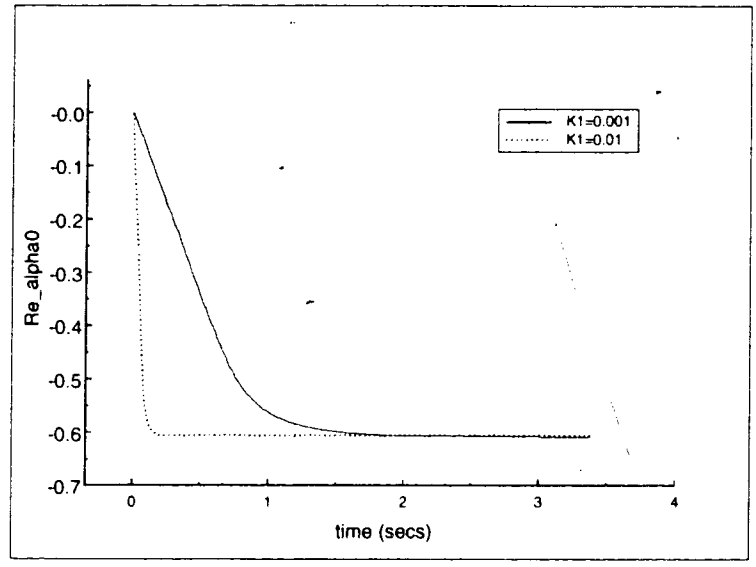

Figure 6.2: Convergence of the real part of α_0 over time for different values of K_1 ($K_2 = 0$).

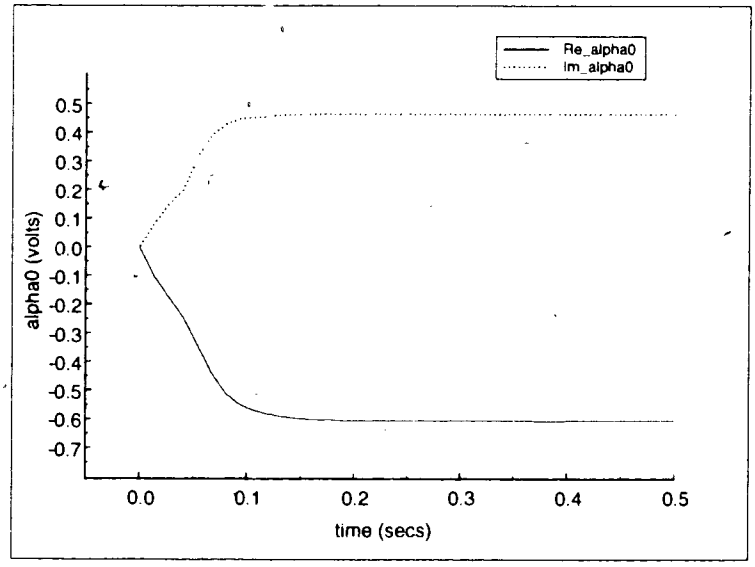

Figure 6.3: Convergence of α_0 for $K_1 = 0.01$ and $K_2 = 0$.

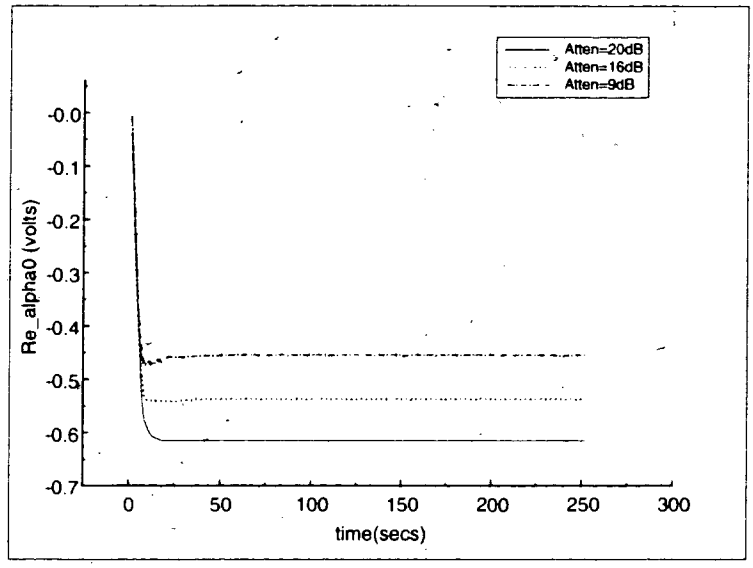

Figure 6.4: Convergence behaviour of the real part of α_0 with varying attenuation in the upper branch ($K_1 = 0.01$, $K_2 = 0$)

these levels as closely as possible is important.

To compare with the case where α_0 and α_1 converge independently, the step-size parameters K_1 and K_2 are set to the same value. The gradient step for the sum only is given as $K_1(D_{\alpha_0}(t) + D_{\alpha_1}(t))$ which increments both coefficients in unison and is approximately equal in magnitude to $2K_1 D_{\alpha_0}(t)$. For the case where $K_1 = K_2$ the gradients for α_0 and α_1 respectively are $2K_1 D_{\alpha_0}(t)$ and $2K_1 D_{\alpha_1}(t)$. Since the magnitude of these gradients in both cases are very similar for the same K_1 values, the convergence times can be compared without introducing a bias to the result. However Figures 6.7 and 6.8 show that for $K_1 = 0.01$ convergence does not occur due to the interactions cause by the large step taken by each adapting coefficient. The signal does in fact get cancelled but the values drift off over time. Therefore a comparison of convergence times can only be made on the coefficients with $K_1 = 0.001$ which is equal to 1.5 sec in both cases. Notice that by scaling the step-size parameter by a factor of 0.1 the convergence time has increased by a factor of 10, a very logical result. From these results it can be concluded that the decorrelation of the gradient signals speeds the convergence time by eliminating the interactions between the coefficients

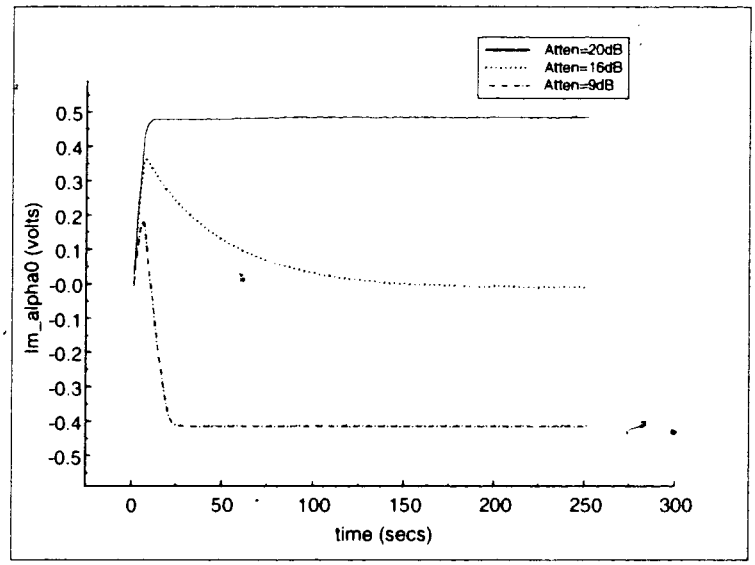

Figure 6.5: Convergence behaviour for the imaginary part of α_0 with varying attenuation in the upper branch. ($K_1 = 0.01$, $K_2 = 0$)

and thus allowing the use of larger step parameters.

The next step is to find the best value for the difference step-size parameter so that the coefficients may diverge from the sum to new optimal values. Setting $K_1 = 0.01$ the circuit coefficients were adapted for three selected values of K_2. Figure 6.9 shows the path of α_0 as it sets out from the origin. Clearly the most appropriate step-size parameter for the convergence of the difference is $K_2 = 0.001$ as it draws a path directly toward the optimal value. Using this value Figure 6.10 shows how the coefficients diverge from the sum to a new optimal value when both sum and difference are adapting together. The adaptation using the difference gradient has lengthened the adaptation time slightly as expected to approximately 0.17 sec, which scales to 5.3 msec.

The amount of suppression obtained from adapting the sum of the gradients is approximately 40 dB. This result is achieved for both values of K_1 and also when the difference gradients are allowed to converge with $K_2 = 0.001$. Since the output of the two VMs are summed together there are a number of possibilities for the values of α_0 and α_1 depending on the chosen step-size parameters K_1 and K_2. The adaptation is quite forgiving in this

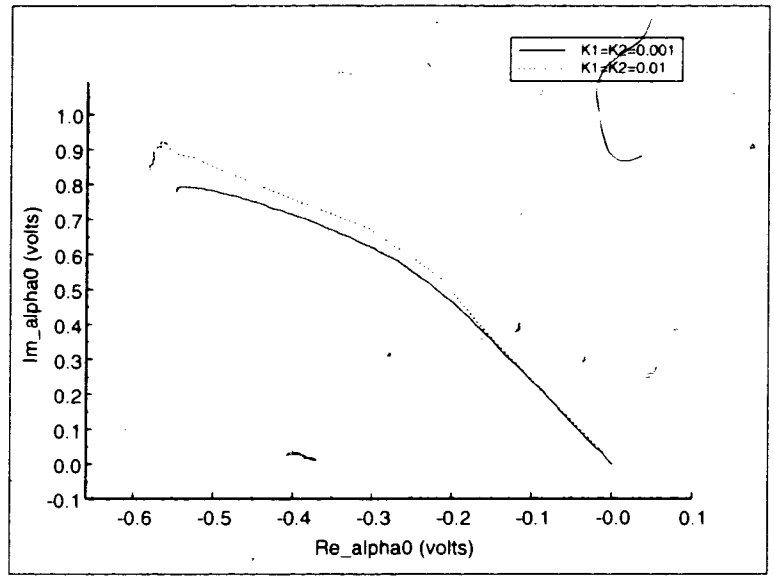

Figure 6.6: Path of α_0 coefficient for different step-size parameters $K_1 = K_2$.

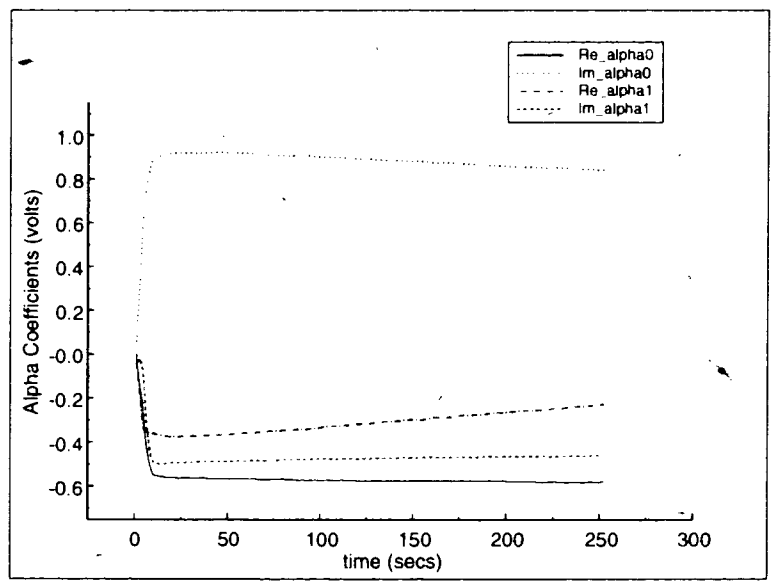

Figure 6.7: Adapting both alpha coefficients together with $K_1 = K_2 = 0.01$.

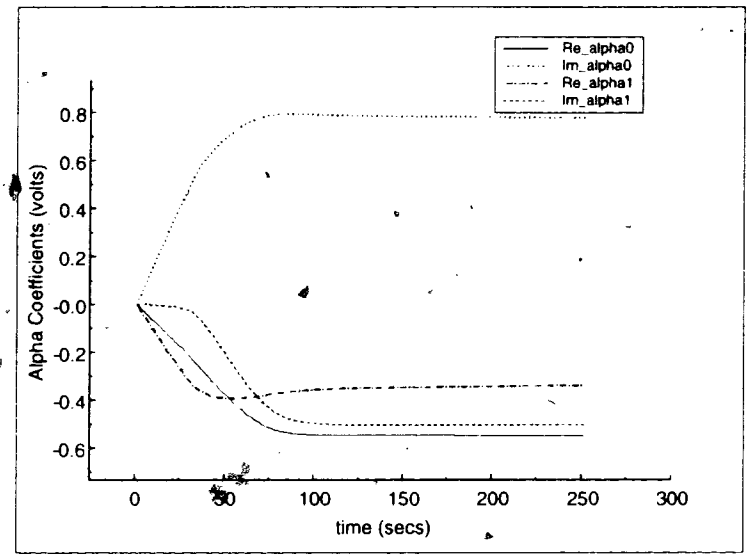

Figure 6.8: Adapting both coefficients together with $K_1 = K_2 = 0.001$

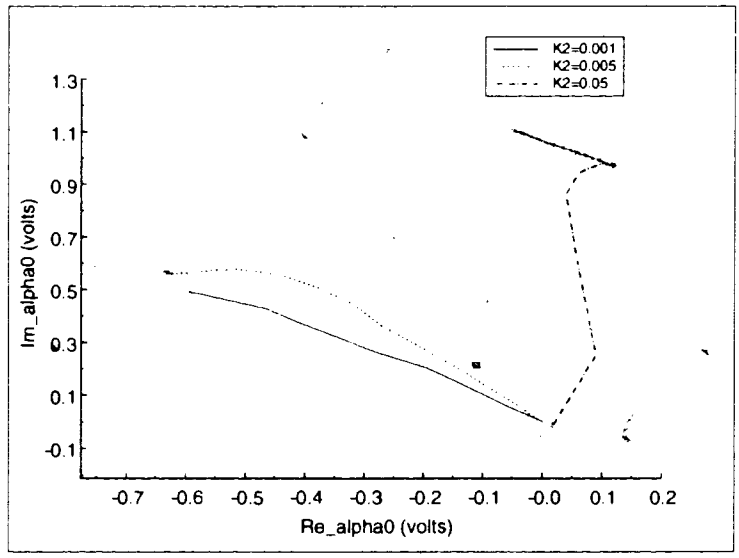

Figure 6.9: Path of α_0 for different values of K_2 ($K_1 = 0.01$).

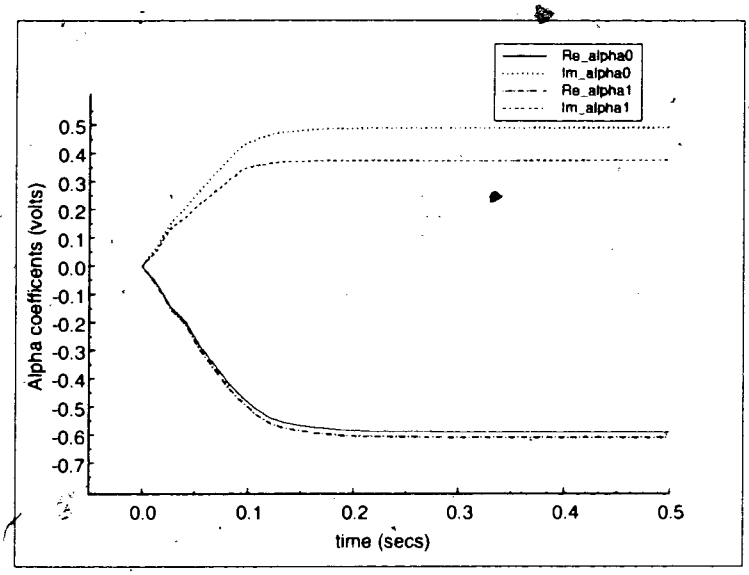

Figure 6.10: Sum and difference adaptation of the alpha coefficients for $K_1 = 0.01$ and $K_2 = 0.001$

respect.

6.1.2 Distortion Cancellation Circuit Adaptation

To achieve fast convergence in the distortion cancellation circuit the magnitude of the step parameters must be much greater than the step parameters in the signal cancellation circuit due to the reduction in power level for the correlations performed in the band of distortion. However, since the gradient estimates will be noisier for the beta coefficients the suppression will be more noticeably compromised with larger step parameters.

Similar measurements were performed as above on the distortion cancellation circuit with the alpha coefficients adapted using $K_1 = 0.01$ and $K_2 = 0.001$. Due to phase rotations caused by delays in the circuit and in the downconversion chains, the step-size parameters for the beta coefficients are negative in value to produce a 180^o phase rotation to the calculated gradients. If the region of convergence is known then the adaptation can be helped along by multiplying the gradient estimates by some phase rotation $e^{j\theta}$ such that the gradient step is taken in the direction toward the point of convergence. This will increase the convergence

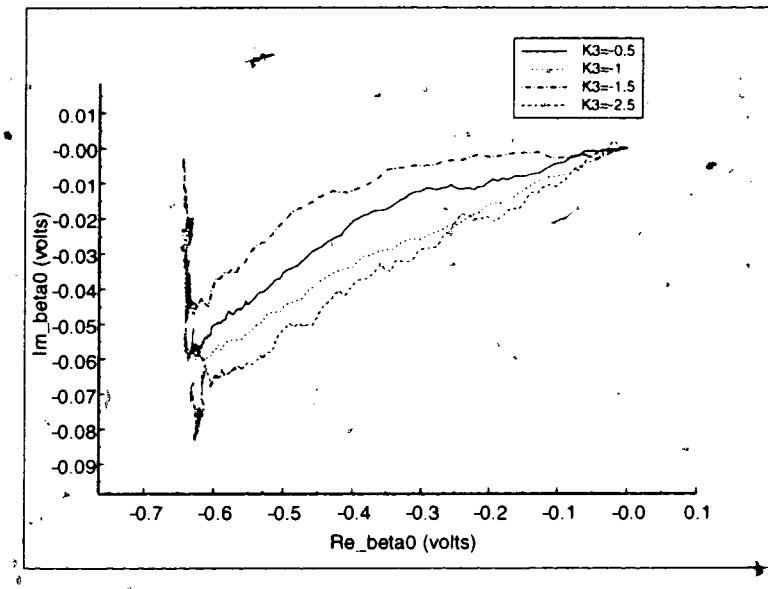

Figure 6.11: Path of β_0 for various sum step parameters with the difference $K_4 = 0$.

time by avoiding a path which spirals in toward the optimum value.

An appropriate choice for the sum step parameter is found by adapting only the sum of the gradients and plotting the convergence behaviour of the coefficients for different values of K_3. The path traced out by the coefficient β_0 is plotted in Figure 6.11 for each step parameter. Since the imaginary part of β_0 is small, the jitter produced by the larger step parameters is more visible. The convergence behaviour of the real and imaginary parts of β_0 is shown in Figures 6.12 and 6.13. Due to the limited size of the dualport memory, every 64^{th} update of the beta coefficients was stored to obtain these results. Convergence speed is increased at the expense of jitter, however the distortion is cancelled perfectly at the output for all cases so the largest value of K_3 can be selected.

The difference step-size parameter K_4 is chosen by setting $K_3 = -2.5$ and plotting the convergence of β_0 for different values of K_4. Obviously the larger the choice for K_4 the further the coefficients will separate from the sum. Plots of convergence behaviour and convergence paths for the beta coefficients using $K_4 = -1$ are given in Figures 6.14 and 6.15. Convergence time for these parameters is approximately 2 sec, which corresponds to approximately 60 msec when taking into account the decimation factor $L = 32$. The jitter

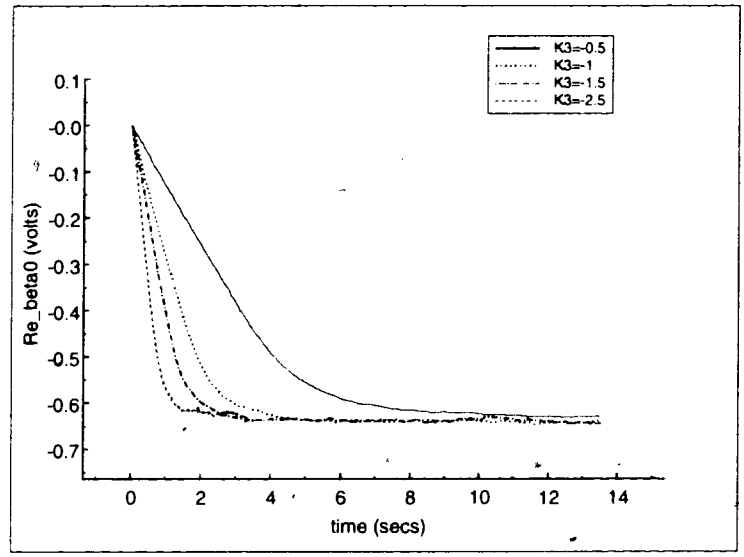

Figure 6.12: Convergence behaviour of the real part of β_0 for different sum stepsize parameters ($K_4 = 0$).

in these plots is less noticeable due to storing every 16^{th} update of the beta coefficients.

To determine the limits to the speed of the adaptation requires performing a series of tests to find the best match for the step-size parameters. A final attempt to improve the convergence speed is shown in Figure 6.16. By increasing K_3 the convergence time has been improved to 1 sec corresponding to approximately 30 msec without decimation.

Finally, to form a comparison between the convergence times for the decorrelation method using the sum only ($K_4 = 0$) and adapting the coefficients independently ($K_3 = K_4$) plots were combined to form Figures 6.17 and 6.18 using $K_3 = -2.5$. Figure 6.18 clearly shows that decorrelating the gradients has increased the convergence speed by a factor of two. Both methods achieve distortion cancellation at the output; however, the increased convergence speed will enhance the circuit's ability to quickly adapt to any changes in operating conditions.

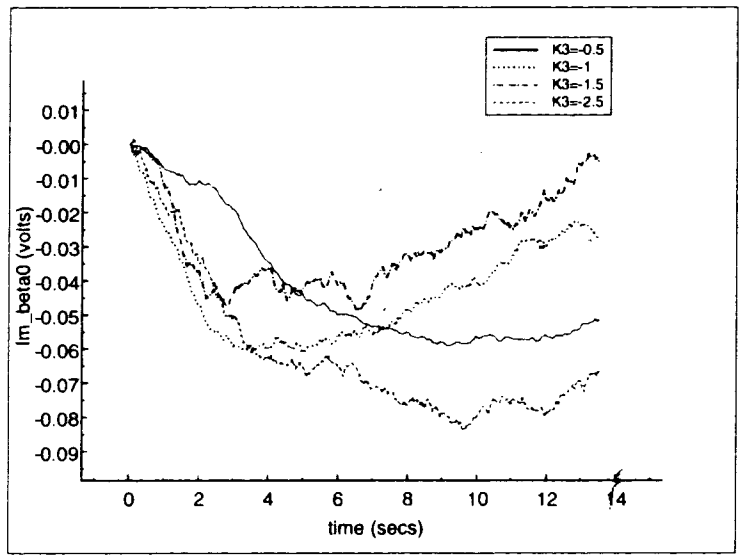

Figure 6.13: Convergence behaviour of the imaginary part of β_0 with different sum step parameters ($K_4 = 0$)

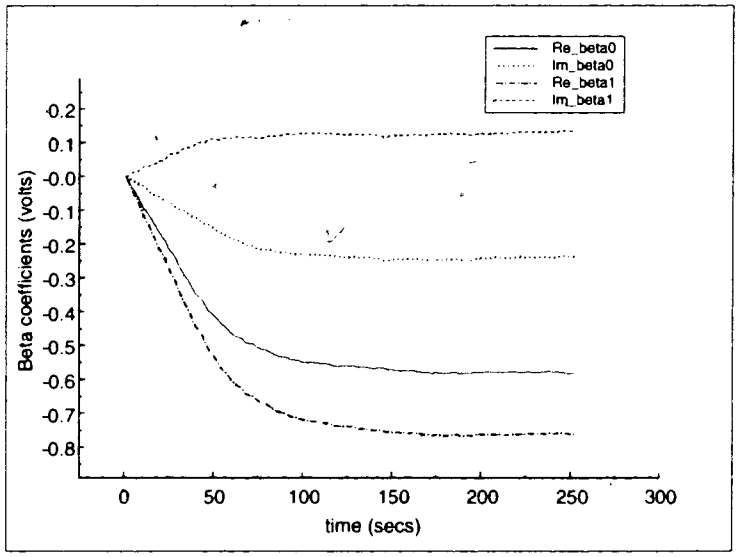

Figure 6.14: Convergence behaviour of beta coefficients when $K_3 = -2.5$ and $K_4 = -1$.

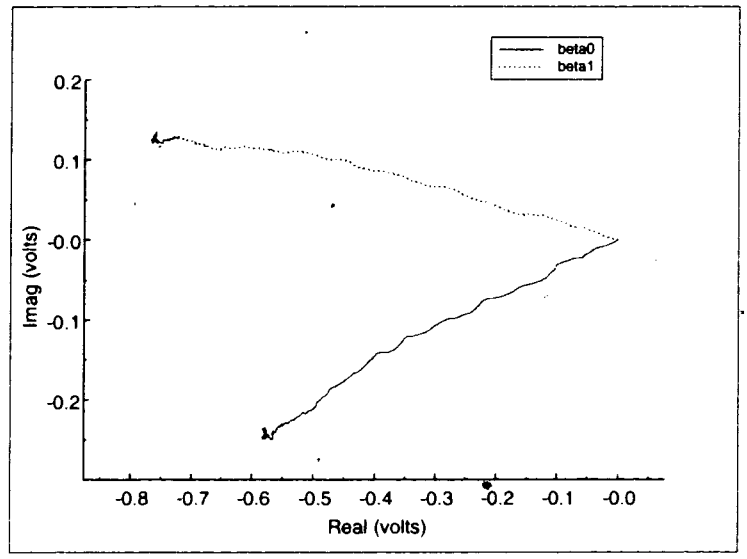

Figure 6.15: Path of beta coefficients for $K_3 = -2.5$ and $K_4 = -1$.

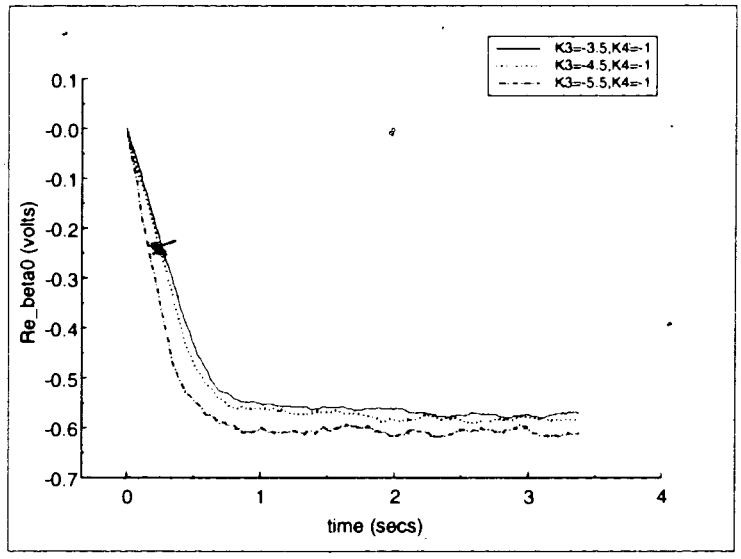

Figure 6.16: Convergence of the real part of β_0 for increasing values of K_3 ($K_4 = -1$).

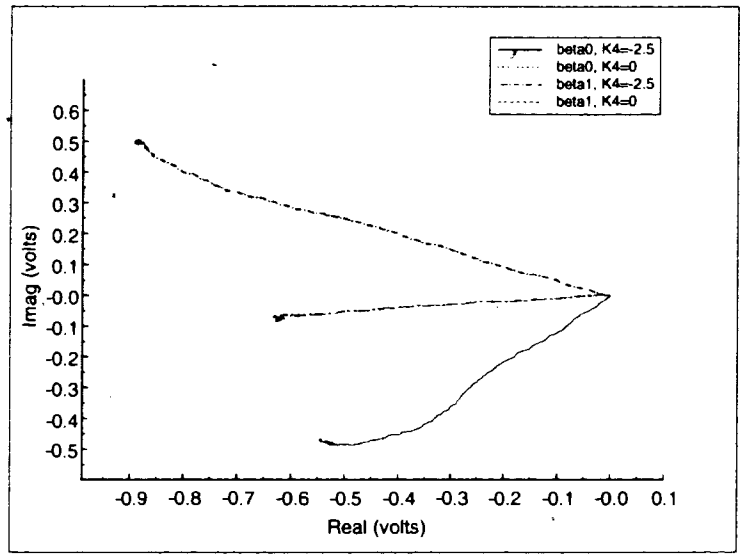

Figure 6.17: Comparison of path traced by beta coefficients for adapting the sum and adapting coefficients independently ($K_3 = -2.5$).

6.2 Suppression of Intermodulation

It was initially thought that a number of selected subbands would be required to obtain good distortion suppression across wide bandwidths and much effort was put into developing an effective method of stepping the adaptation across the band. However Figure 6.19 illustrates a very interesting feature of the circuit. The signal cancellation coefficients are adapted at the centre band of the narrow bandwidth signal (128 ksym/sec) with a distortion bandwidth of approximately 400 kHz and the coefficients are saved when they reach their optimum values. The symbol rate is increased by a factor of 100 to 12.8 Msym/sec and the span on the spectrum analyser is increased to 40 MHz. The output of the signal cancellation circuit is plotted to illustrate that the adaptation at the centre frequency provides adequate signal suppression over a very wide bandwidth. As long as the taps span the delay in the reference path over the whole bandwidth of interest then only one correlation for the alpha coefficients is required. This simplifies the adaptation procedure immensely.

Figure 6.20 shows the narrowband QPSK input signal and the associated distortion produced by the PA. The noisy floor of the input signal is due to the output frequency of

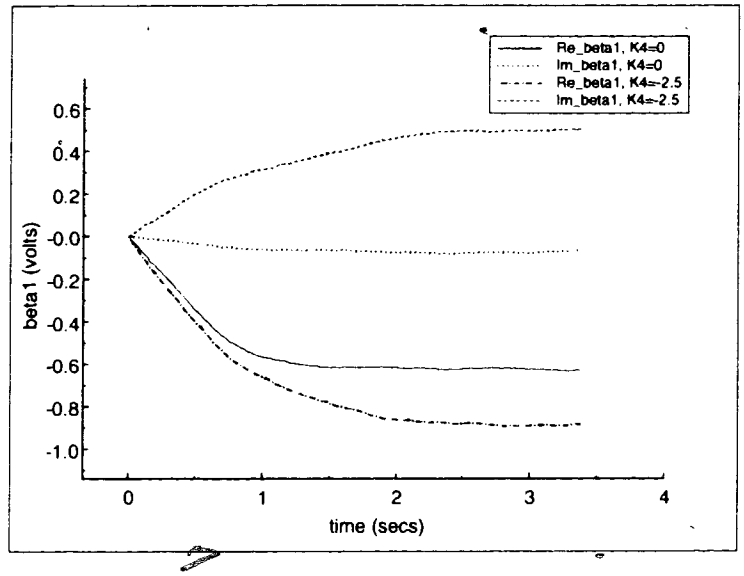

Figure 6.18: Comparison of convergence behaviour for adapting the sum only and adapting β coefficients independently ($K_3 = -2.5$).

the waveform generators used to modulate the I /Q inputs of the 1.8 GHz signal generator. Due to the loss of the high output power preamplifier in the final stages of the project, the PA could not be driven close to its 1 dB compression point and therefore high levels of distortion could not be generated. The output power of the PA is 26 dBm and all plots shown are derived from this power level.

The distortion signal output from the signal cancellation circuit after the convergence of the alpha coefficients is shown in Figure 6.21 compared with the spectra of the input signal that produced it. This clearly shows the complete cancellation of the reference signal from the error signal.

The degree of linearisation achieved for the narrowband signal is shown in Figure 6.22. The alpha coefficients are first adapted at centre band then the LO is stepped to 125 kHz away from the carrier to adapt the beta coefficients. The output signal is almost identical to the input signal with a peak IMD suppression of 35 dB.

Difficulty arose when trying to generate a wideband QPSK signal. Since the input power is distributed across the bandwidth of the input signal the voltage levels at the output of the

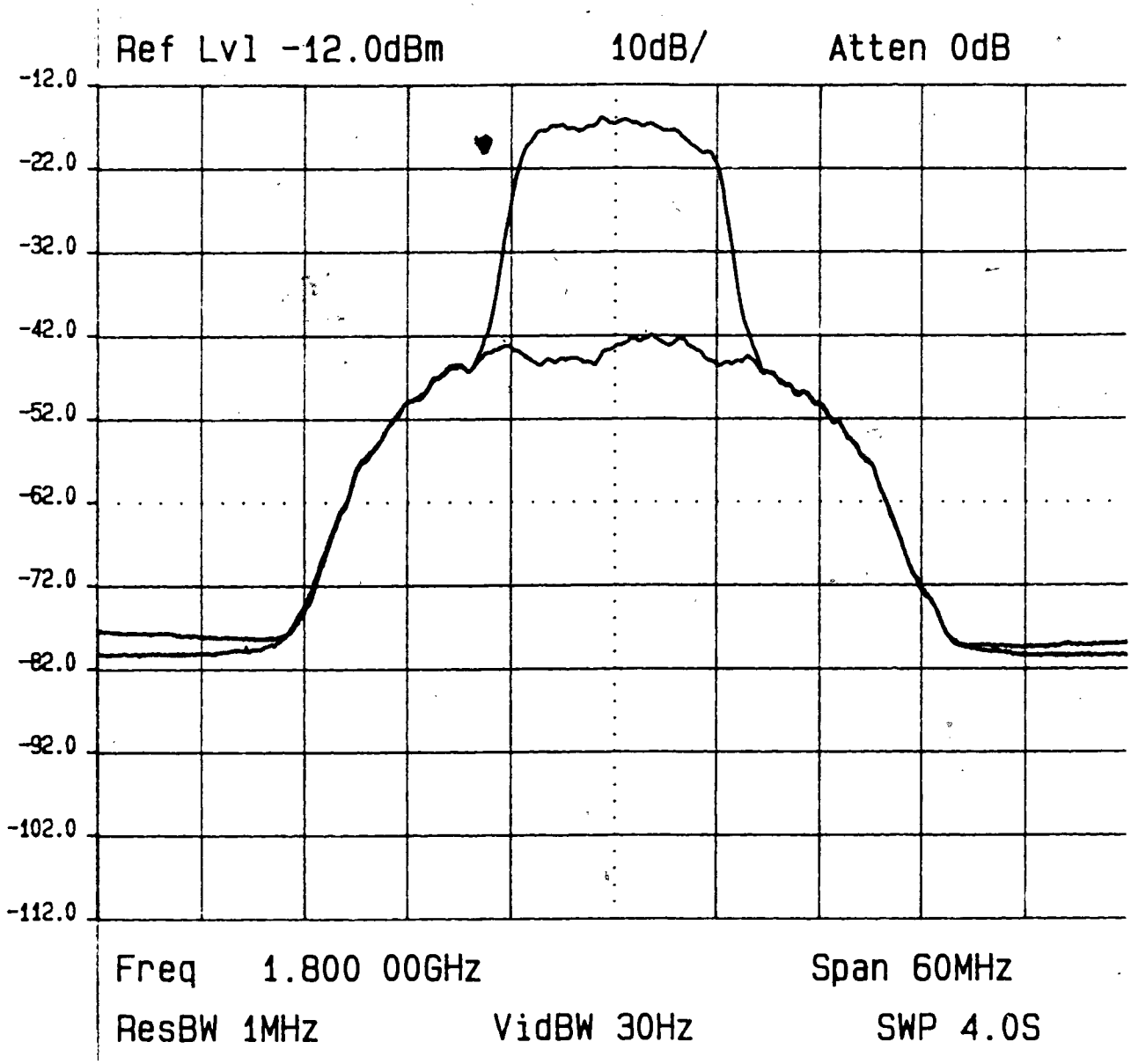

Figure 6.19: Signal suppression of 12.8 Msym/sec QPSK signal for coefficients adapted at centre frequency of narrowband signal.

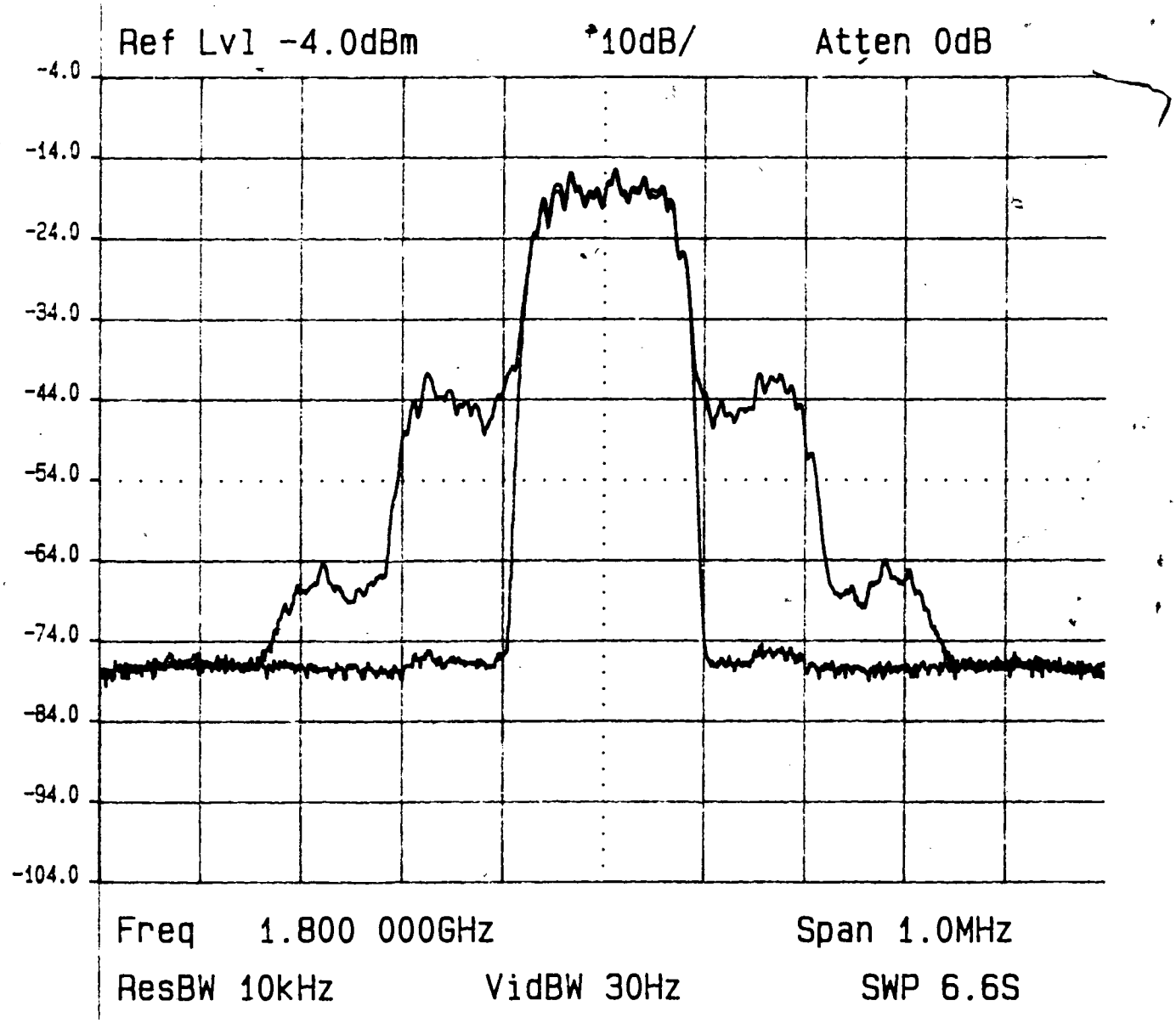

Figure 6.20: Spectrum of narrowband QPSK input signal and the resulting PA output signal.

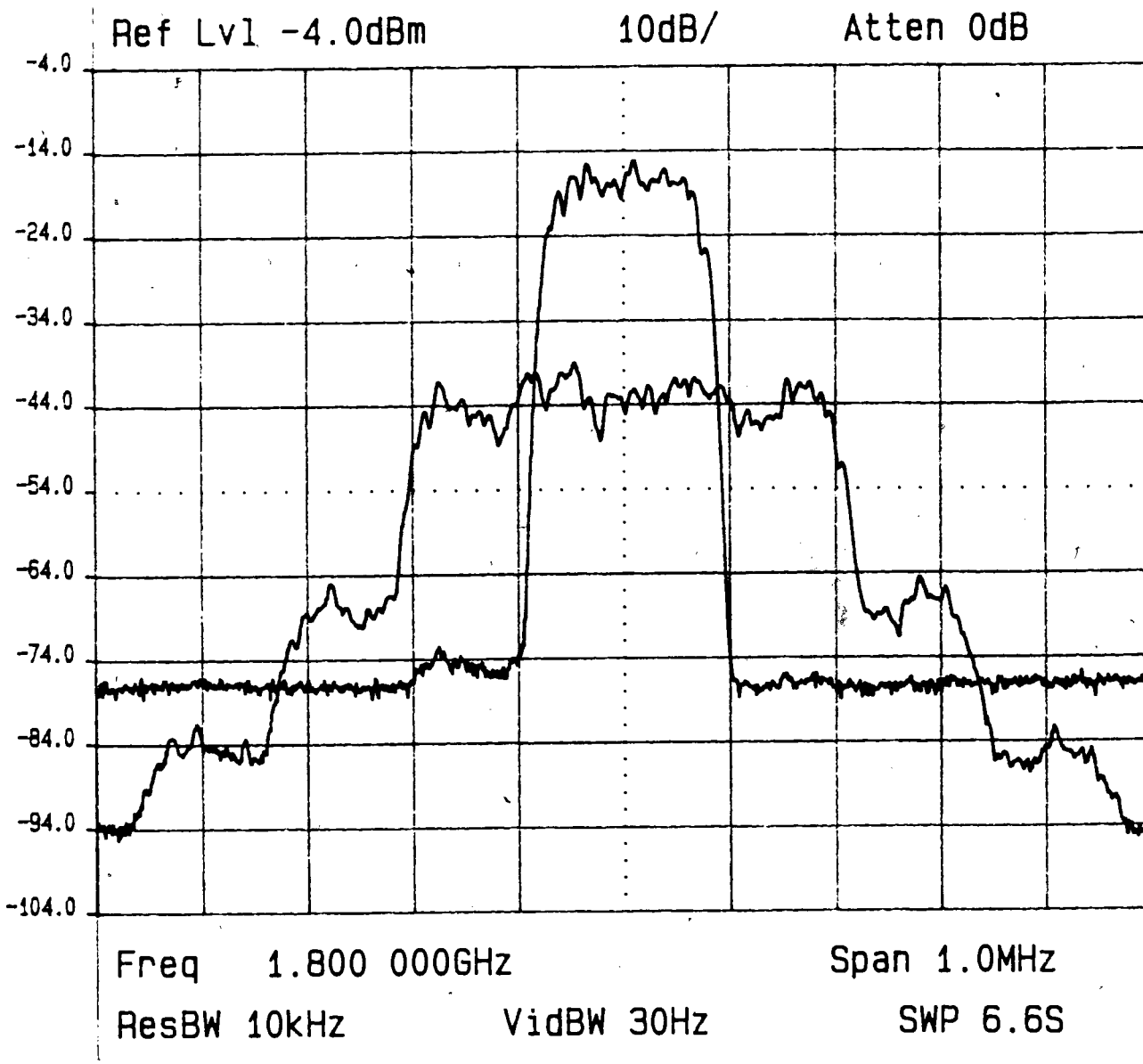

Figure 6.21: Spectra of narrowband QPSK input signal and error signal at signal cancellation circuit output.

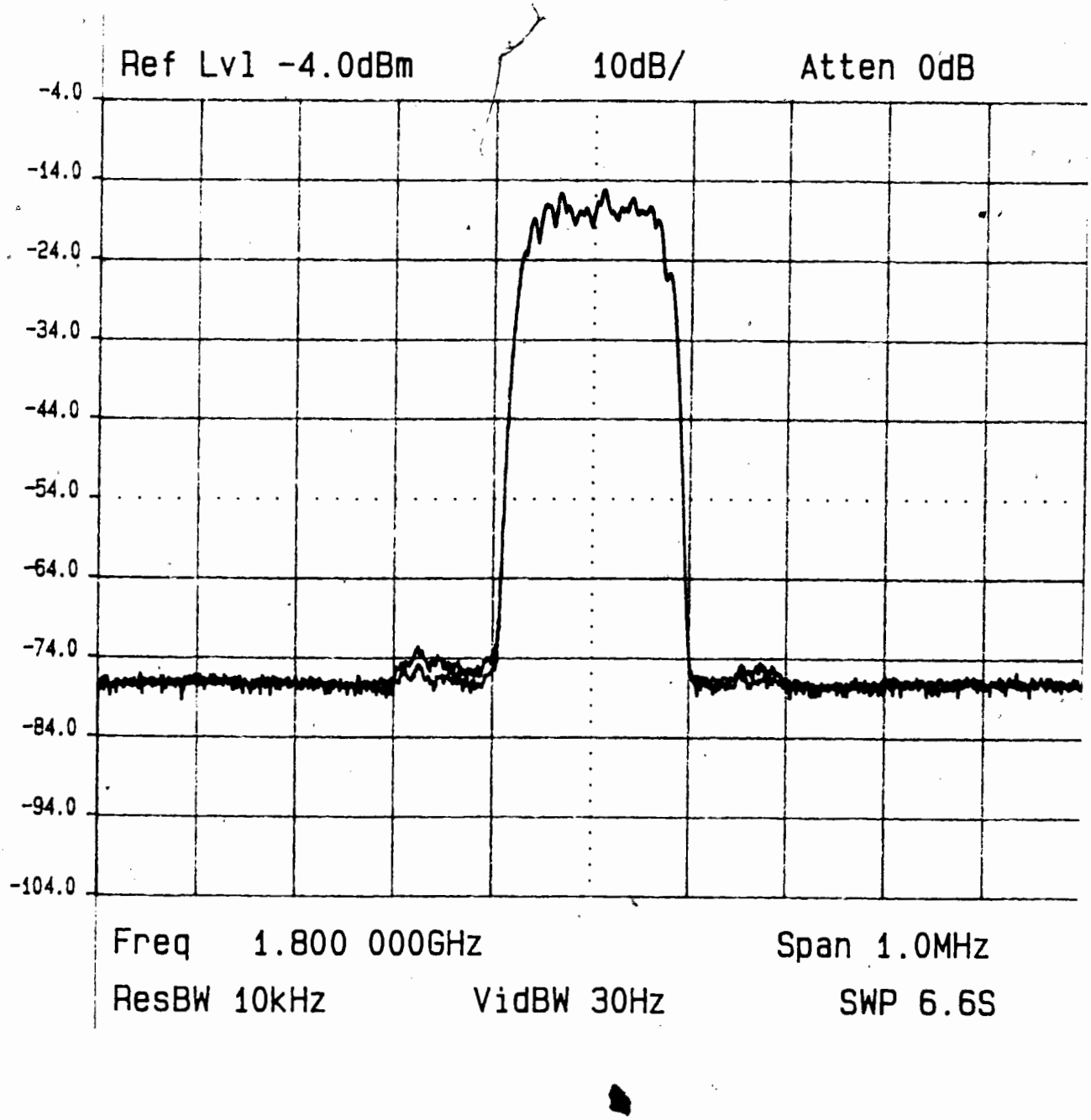

Figure 6.22: Lineariser output spectrum compared with that of the input spectrum.

downconversion chains are significantly lower and the existing baseband amplifiers were not capable of amplifying the signal up to a reasonable level. To compensate for the decrease in power level the programmable gain amplifiers on the DSP were set to a gain of 100. A 12.8 Msym/sec QPSK signal was generated using the arbitrary waveform generators and the signal cancellation circuit had no difficulty in cancelling the input signal from the error signal. The distortion cancellation circuit, however, had problems dealing with the noisy correlations at the low power levels and small values of K_3 were required to avoid large fluctuations in the suppression at the feedforward output. Covergence time was increased by a factor of 10. A solution to this problem would be to split the error signal $v_e(t)$ in the downconversion chain so that the level for the beta correlations could be increased thus improving the SNR in the ADC. Figure 6.23 shows the wideband input signal and amplifier output before linearisation. Figure 6.24 compares the linearised output with the input signal showing almost complete cancellation of the IMD.

The wide bandwidth capability of the circuit is measured by adding a single tone, offset from the original modulated carrier signal by 20 MHz, at the input of the feedforward circuit using a power combiner. The addition of the combiner to the input results in a reduction of the level of the narrowband QPSK signal of 4 dB and leads to a PA output power of approximately 22 dBm. When observing the input signal over very wide bandwidths the modulation detail in the 1.8 GHz signal is lost and the input can be considered equivalent to a two tone test. Figure 6.25 shows the original QPSK signal modulated with 128 ksym/sec with a tone at 1.78 GHz. The intermodulation products (IMPs) caused by the nonlinear PA are shown in Figure 6.26 with the signal cancellation coefficients adapted. The tone at 1.78 GHz and all IMP products include replications of the distortion produced by the QPSK signal. When the adaptation is performed, the alpha coefficients adapt at the centre of the band then the beta coefficients are allowed to adapt on the IMP at 1.82 GHz using $K_3 = -1$ and $K_4 = -0.1$. The magnitude of K_3 is smaller due to the higher power level of the distortion at the 1.82 GHz image frequency. The distortion improvement is illustrated in Figure 6.27. The IMP at 1.82 GHz has been suppressed by more than 40 dB which equates to a 40 MHz linearisation bandwidth- a very nice result. This figure also shows that at ± 40 MHz the intermodulation products are suppressed by 25 dB.

If the gain and phase characteristics of each of the paths for the vector modulator pairs had identical gain and phase characteristics, then the adaptation of only the sum of the gradients would be equivalent to the original single coefficient case and no improvement for

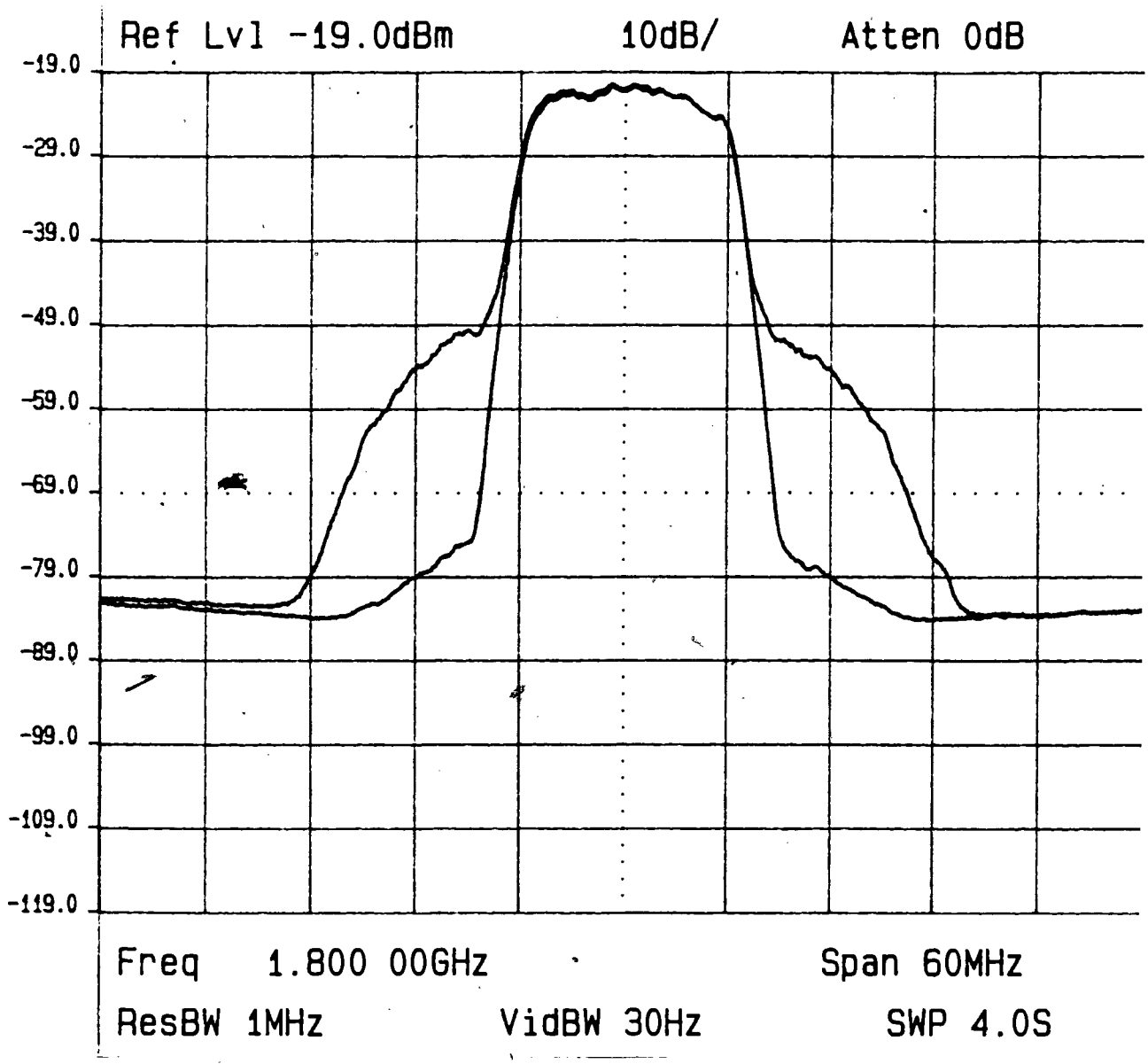

Figure 6.23: Spectra of wideband QPSK input signal and amplifier output signal before linearisation.

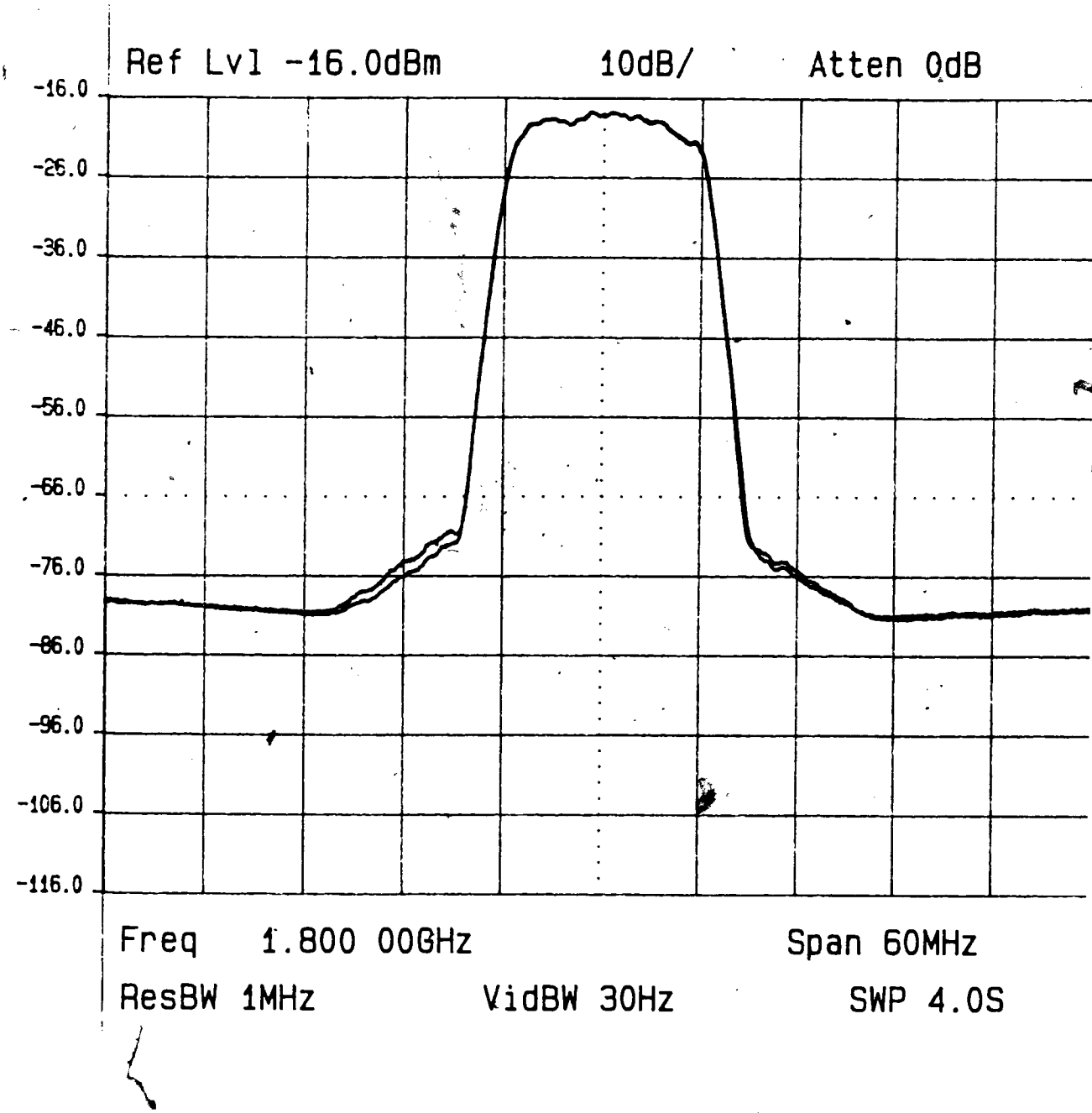

Figure 6.24: Lineariser output spectra of wideband QPSK signal compared with input signal.

wide bandwidth linearisation would be expected. However the added delay in each of the paths for α_1 and β_1 has an associated phase rotation and the components within the vector modulator have varying gain and phase characteristics. For this reason, adapting the sum only can actually find an optimum value for the coefficients which achieves the wide band linearisation results as above. In other words, K_2 and K_4 can be set to zero while still obtaining 40 dB suppression over 40 MHz bandwidth.

Extensive testing proved that adapting the signal cancellation coefficients at the centre of the band and the distortion cancellation coefficients at the band edges produced the best result. Adapting the beta coefficients in the IMD band 125 kHz from the 1.8 GHz narrowband QPSK signal results in only 30 dB suppression of the IMP at 1.82 GHz shown in Figure 6.28. Figure 6.29 shows how a further 10 dB is achieved when the beta coefficients are instead adapted on the 1.82 GHz IMP. Some residual distortion now exists in the upper band of the IMD as illustrated by Figure 6.30. To optimise the suppression obtained at both the centre band and at the band edges, partial gradients could be calculated on the subbands at each point and the coefficients updated with the combined sum of the scaled gradients. In this way an average value across the band could be obtained. Due to the time delay in changing the frequency of the LO, this method of optimisation was not performed.

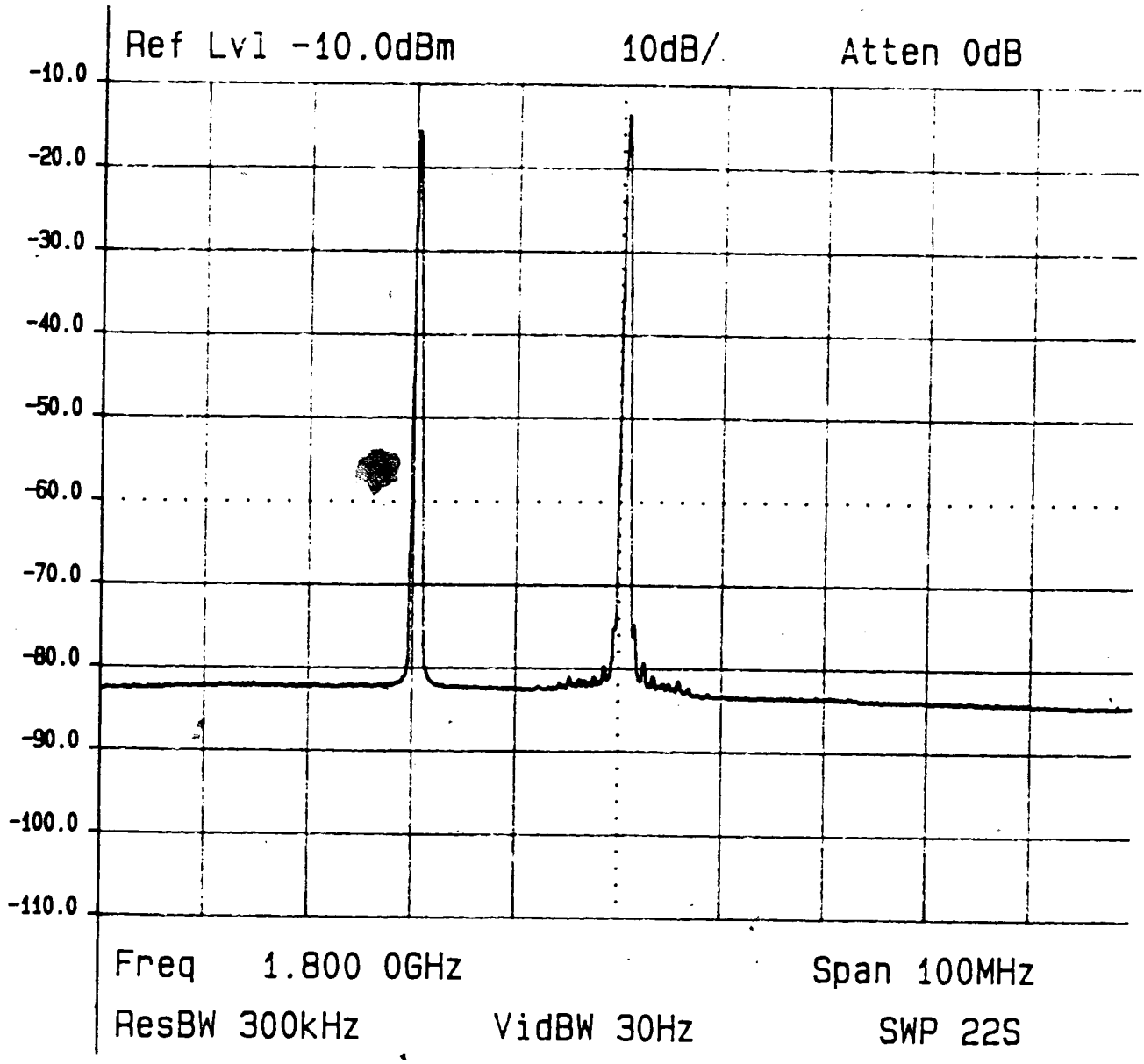

Figure 6.25: Input signal spectrum of narrowband QPSK input signal at 1.8 GHz and 1.78 GHz tone.

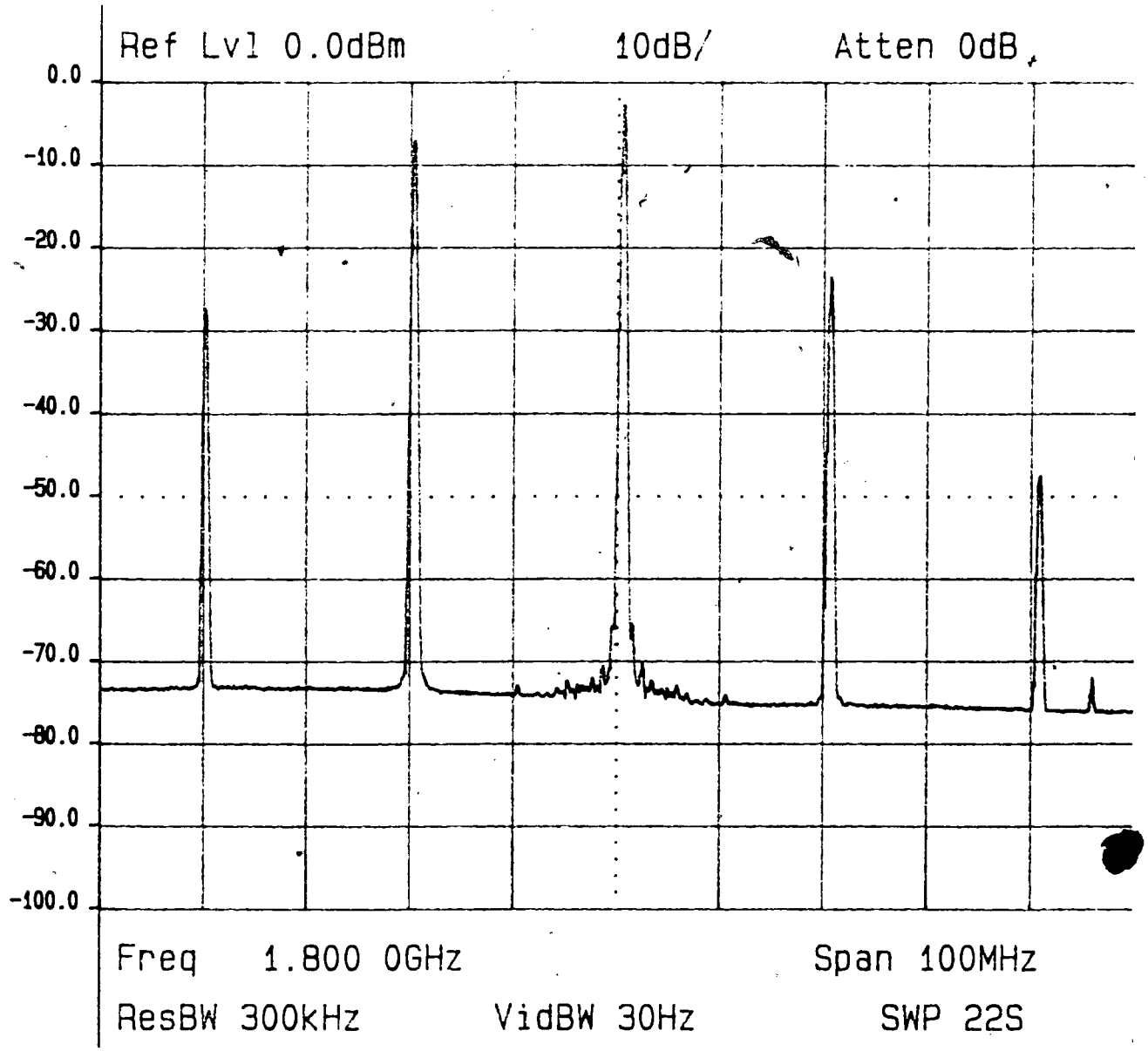

Figure 6.26: Amplifier output spectrum for narrowband QPSK input signal and 1.78 GHz tone.

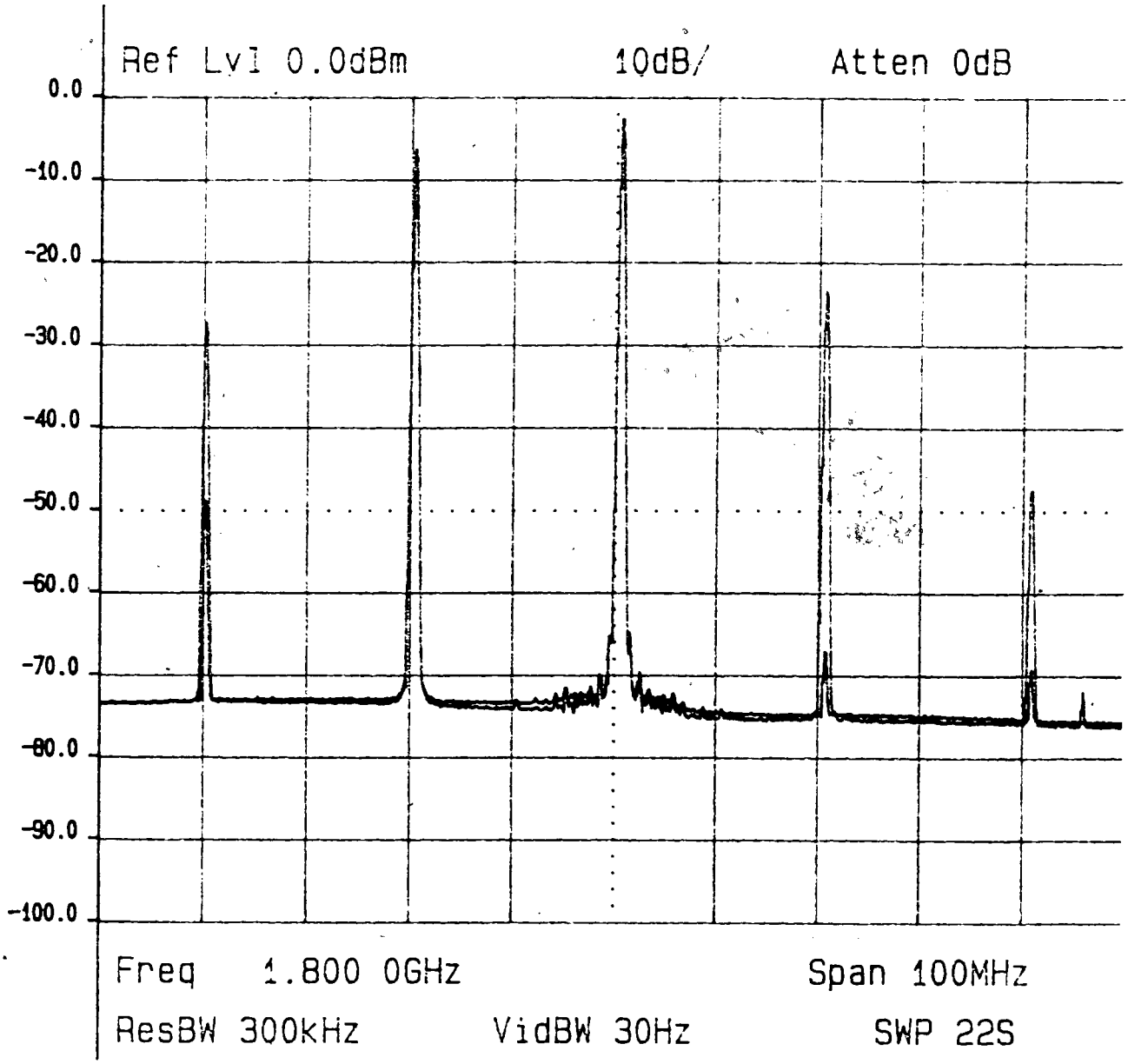

Figure 6.27: Feedforward amplifier output spectra before and after linearisation for narrow-band QPSK with tone at 1.78GHz.

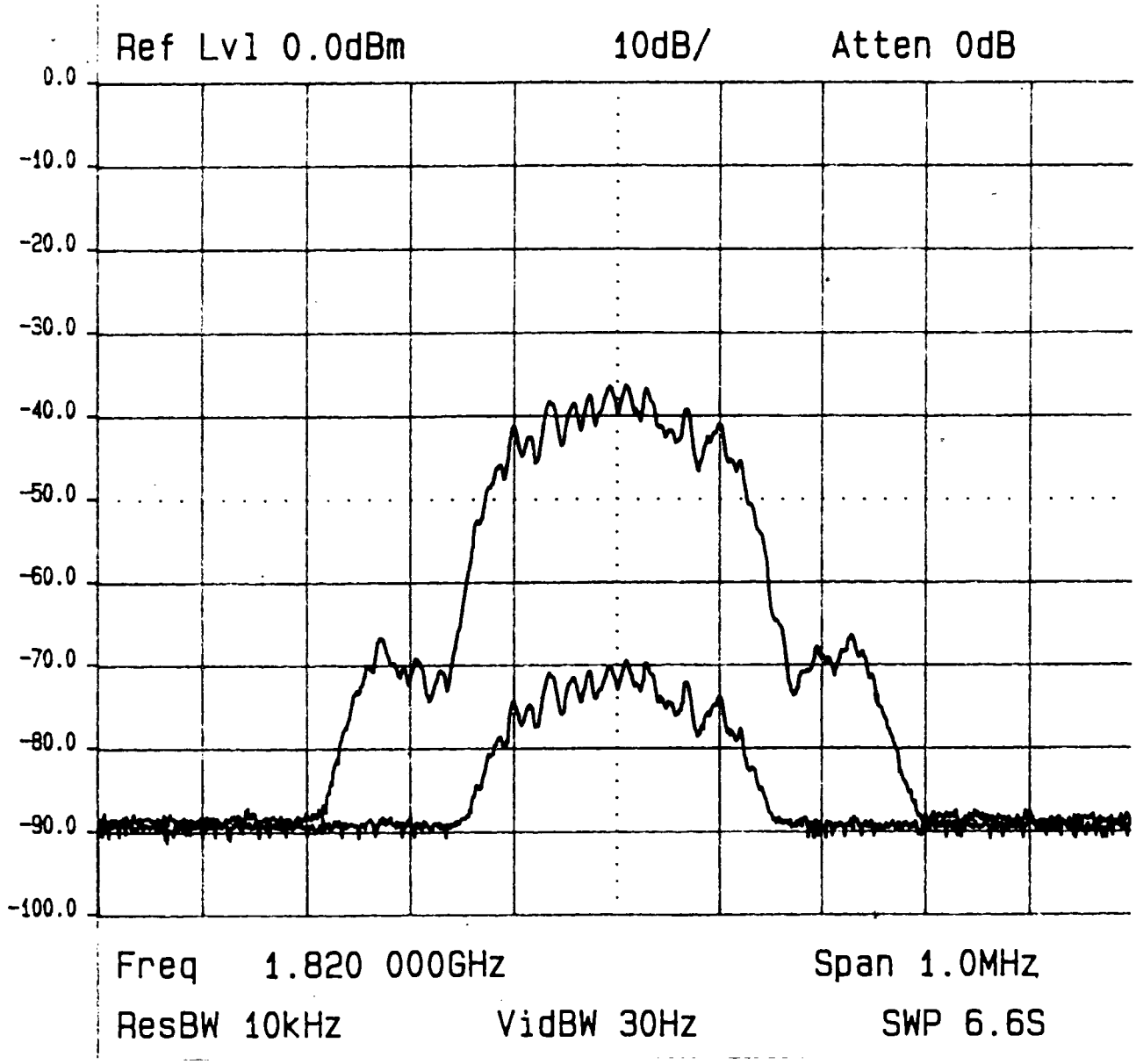

Figure 6.28: Spectra of 1.82 GHz IMP showing suppression after beta coefficients have adapted on IMD, 125 kHz from band centre.

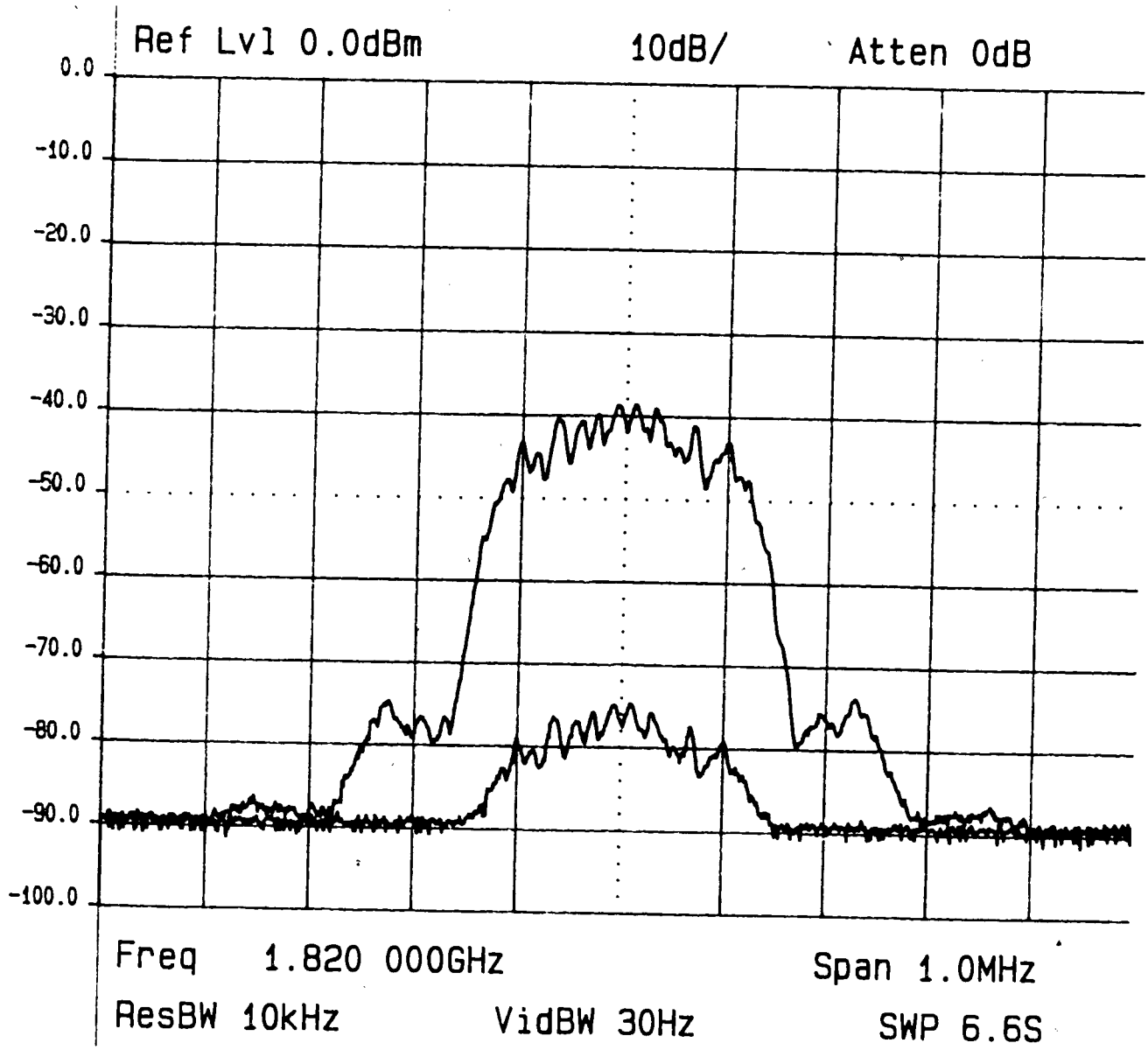

Figure 6.29: Spectra of 1.82 GHz IMP showing suppression after beta coefficients adapted on the 1.82 GHz IMP.

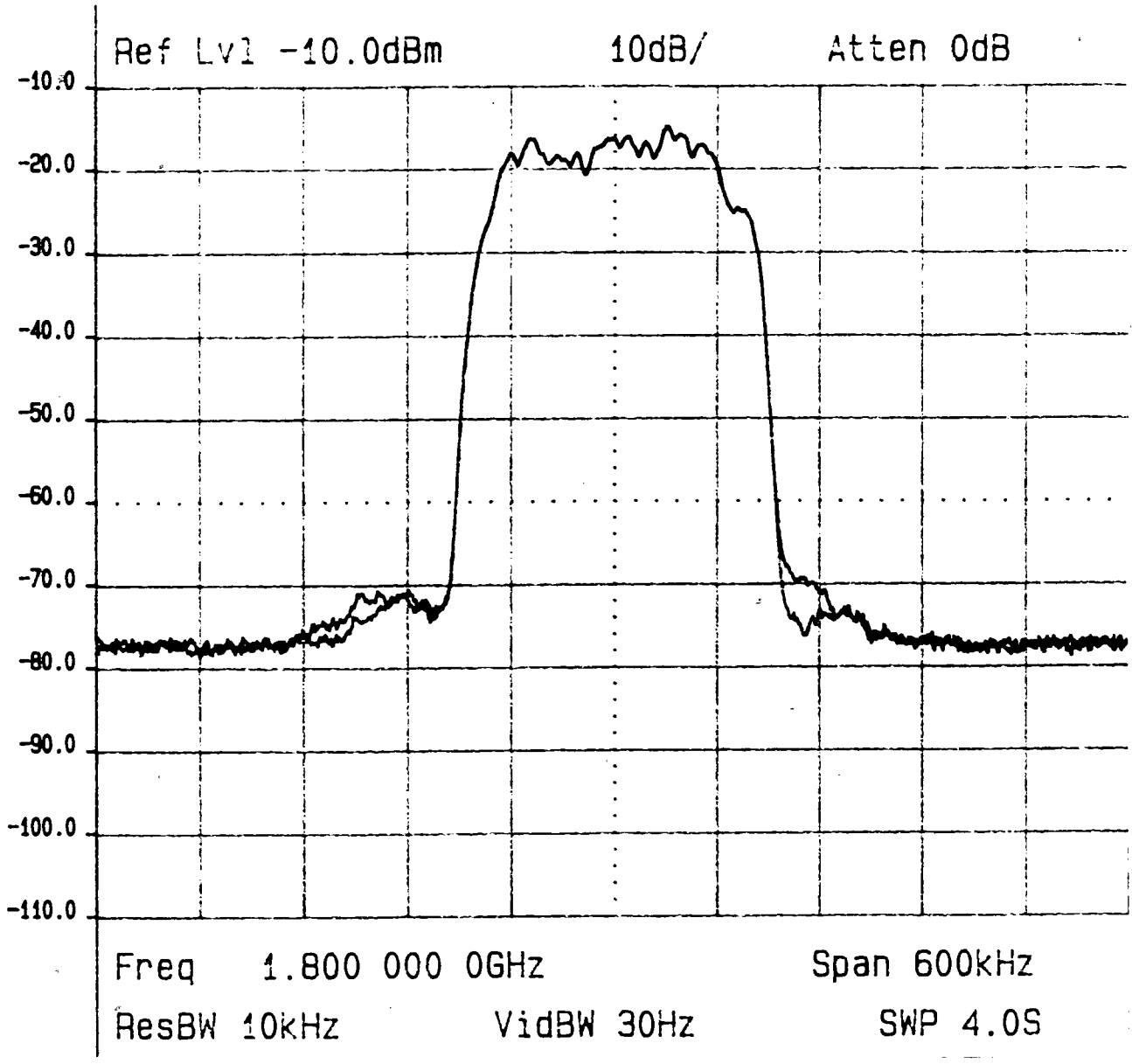

Figure 6.30: Spectra of QPSK input signal showing residual distortion introduced by adapting beta coefficients on 1.82 GHz IMP.

Chapter 7

Conclusions

The addition of the delay matching circuitry to the original single coefficient adaptive feed-forward amplifier has presented some quite remarkable results. The analysis showed that not only will the circuit compensate for the mismatches in the delays between the branches but it also has the ability to compensate for linear variations of amplitude with frequency. Moreover, since the measurements on the actual circuit showed a nonlinear variation of amplitude across the 40 MHz linearisation bandwidth this suggests that there are less stringent requirements on the passband ripple than expected.

The ability to select appropriate subbands for the separate adaptations of the alpha and beta coefficients overcame the masking problem which causes slow convergence of the distortion cancellation coefficients. The implementation showed that only one subband at the centre frequency was required to adapt the signal cancellation coefficients across a 40 MHz bandwidth. The best method of adapting the distortion coefficients for wide bandwidth operation was to perform the correlations using the intermodulation products at the edge of the band. This resulted in a very small trade-off of the linearisation achieved close to the centre frequency. Best suppression was achieved when only adapting the sum of the gradients for the beta coefficients. Weak signals introduce a lot of noise to the adaptation and adequate amplification to utilise the full range of the ADC should be designed.

Decorrelation of the gradient signals proved an effective method to increase the convergence time over the standard method, achieving initial convergence times for the signal cancellation circuit in the order of 5 msec. For the distortion cancellation circuit initial convergence times were in the order of 30 msec. The convergence times are controlled by the choice of step parameter K_1 which determines the steps taken toward the optimum

value using the higher SNR gradient sum. Since only one cancellation circuit was adapting at a given subband, reconvergence times after a change in input power level were not tested. The effect on the jitter of the coefficients caused by the sum and difference step parameters was proved by analysis to be approximately equal to the jitter introduced by the adapting coefficients in the standard configuration for equivalent step parameters $K_1 = K$.

Gradient adaptation of the coefficients has proved to be a very reliable method of adaptation. It was shown that the effect of aliasing caused by sampling below the Nyquist frequency does not bias the correlation but does introduce a degree of self noise to slow the convergence time.

Although the main amplifier was not run near its 1 dB output compression point the circuit's ability to suppress intermodulation products at very wide bandwidths was shown while still maintaining good suppression of the narrowband signal. The results show that the linearisation circuit was able to achieve 40 dB cancellation of intermodulation products across a bandwidth of 40 MHz. The degree of cancellation decreased with increasing distance from the centre frequency, but was still 25 dB at $\pm$ 40 MHz away from centre. Since the IMD itself decreases with distance from centre, the net effect is suppression of all IMD to a roughly constant level. Limitations on the bandwidth of the circuit will most likely be due to the spacing of the taps chosen between the paths of the adapting coefficients. It would be expected that the circuit will also have limits on the amount of ripple it can accept and at wider bandwidths these effects will be more significant.

Bibliography

[1] S.J. Grant, "A DSP Controlled Feedforward Power Amplifier Lineariser," M.A.Sc. Thesis, School of Engineering Science, Simon Fraser University, July 1996.

[2] J.K. Cavers, "Adaptation Behavior of a Feedforward Amplifier Lineariser," *IEEE Transactions on Vehicular Technology*, vol. 44, no. 1, pp 31-40, February 1995.

[3] S.J. Grant and J.K. Cavers, "A DSP Controlled Feedforward Amplifier Lineariser", *IEEE ICUPC*, Campbridge Ma. Sept 1996.

[4] J.K. Cavers, "Adaptive Feedforward Lineariser for RF Power Amplifiers," U.S. Patent 5,489,875, February 6,1996.

[5] H. Seidel, "A Microwave Feedforward Experiment", *The Bell System Technical Journal*, vol.50 , no.9, pp 2879-2916, November 1971.

[6] W. Bosch and G. Gatti, "Measurement and simulation of memory effects in predistortion linearisers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 37, no.12, pp1885-1890, December 1989.

[7] S.P. Stapleton, G.S. Kandola, and J.K. Cavers, "Simulation and Analysis of an Adaptive Predistorter Utilizing a Complex Spectral Convolution," *IEEE Transactions on Vehicular Technology*, vol. 41, no. 4, pp 387-394, November 199707-719, November 1996.

[8] L. Sundstrom, M. Faulkner and M. Johansson, "Quantisation Analysis Design of a Digital Predistortion Lineariser for RF Power Amplifiers," *IEEE Transactions on Vehicular Technology*, vol. 45, no. 4, pp 707-718, November 1996.

[9] A.S. Wright and W.G. Durtler, "Experimental Performance of an Adaptive Digital Linearised Power Amplifier," *IEEE Transactions on Vehicular Technology*, vol. 41, no. 4, pp 395-400, February 1995.

[10] J.K. Cavers, "Amplifier Linearisation using a Digital Predistorter with Fast Adaptation and Low Memory Requirements," *IEEE Transactions on Vehicular Technology*, vol. 39, no. 4, pp 374-382, November 1990.

[11] M. Faulkner and M. Johansson, "Adaptive Linearisation Using Predistortion- Experimental Results", *IEEE Transactions on Vehicular Technology*, vol. 43, no. 2, pp 323-332, May 1994.

[12] M. Johansson and L. Sundstrom, "Linearisation of RF Multicarrier amplifiers using Cartesian Feedback," *Electronic Letters*, vol.30 , no.14, pp1110-1112, 7th July 1994.

[13] M. Johansson and T. Mattson, "Linearised High-Efficiency Power Amplifier for PCN," *Electronic Letters*, vol.27 , no.9 pp 762-1112, 25th April 1991.

[14] M. Johansson and T. Mattson, "Transmitter Linearisation using Cartesian feedback for Linear TDMA Modulation," Proceedings 41st IEEE Vehicular Technology Conference, pp 439-444, May1991.

[15] M. Johansson, "Linearisation of Wideband RF Power Amplifiers using Modulation Feedback", Doctoral Dissertation, Dept. of Applied Electronics, Lund Institute of Technology,1995.

[16] Y. Akaiwa and Y. Nagata, "Highly Efficient Mobile Communications with a Linear Modulation Method," *IEEE Journal in Selected Areas in Communications*, vol. SAC-5, no.5, pp 890-895, June 1987.

[17] A.M. Brown and V. Petrovic, "Phase Delay Compensation in HF Cartesian-Loop Transmitters", *Fourth International Conference on HF Communication Systems and Techniques* , London,11-14 , pp 200-204, April 1988.

[18] E. Bedrosian and S.O. Rice, "The Output Properties of Volterra Systems (Nonlinear Systems with Memory) Driven by Harmonic and Gaussian Inputs", *Proceedings of the IEEE*, vol.59, no.12, pp1688-1706, December 1971.

[19] A. Papoulis, "Probability, Random Variables and Stochastic Processes," McGraw Hill 3rd ed, pp 412-416, 1991.